



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. Ellis

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 ($1/\text{variance}$) 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 DiBAC ₄ (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.

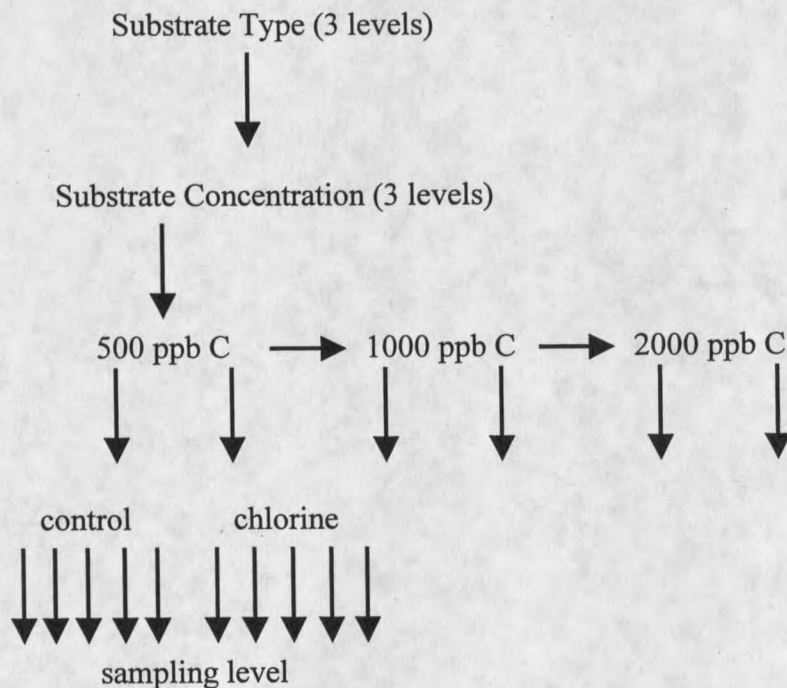


Figure 1.1 – Schematic of the experimental design used to evaluate substrate type, substrate concentration, and chlorination in this dissertation. Samples can be divided into control and chlorinated which can be nested within the next higher level of substrate concentration which contained three treatment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

water bath maintained at $20\pm1^{\circ}\text{C}$ utilizing an external cooling device. The influent solution to each reactor consisted of dilution water from the biologically activated carbon (BAC) filter, plus the substrate and buffer solution (see Figure 2.1).

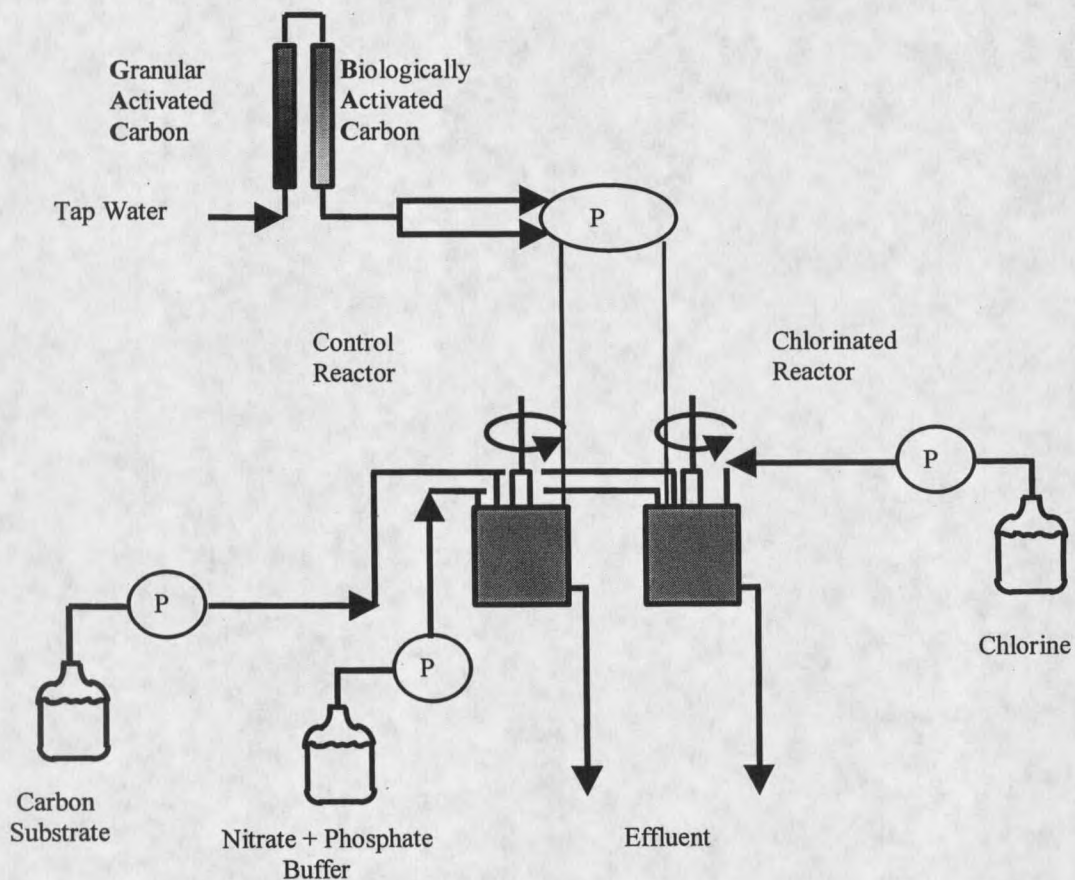


Figure 2.1 - Schematic of experimental setup. The circles with P inside refer to pumps. See text for further details.

City of Bozeman, Montana, USA, drinking water was treated in two columns containing granular activated carbon (GAC) media to provide the dilution water at a flow rate of 4.5 mL/minute. The first column removed chlorine from the water and contained GAC (12 x 40 mesh Nuchar, Westvaco). The second column contained biologically

activated carbon (BAC) originally obtained from a full-scale water treatment plant (City of Laval, Quebec, Canada), and had been receiving Bozeman drinking water for over three years prior to the experiments. The assimilable organic carbon (AOC) level of the effluent from the BAC column was typically less than 25 micrograms per liter (μg acetate equivalents/L) as determined by the method of van der Kooij (van der Kooij, et al., 1982) as modified by LeChevallier et al. (LeChevallier, et al., 1993). The BAC column effluent contained a mixed population of heterotrophic bacteria (ca. 10^4 to 10^5 total cells/mL as determined by epifluorescence microscopy) that colonized the reactor surface and provided a continuous supply of nutrient-limited bacteria to the reactor. Substrate and nitrate/phosphate feed solutions were made up in glass carboys using Nanopure reagent grade water that had been autoclaved prior to use. All glass carboys were combusted at 500°C for 4 hours prior to adding Nanopure water and autoclaving. Sterile substrate or nutrient solutions were added by syringe to the carboys after the water had been autoclaved and cooled to room temperature.

Prior to each experiment the reactors were cleaned and disinfected using a 50 mg/L free chlorine solution placed in the reactor for at least 24 hours. The reactor top assembly/inner-drum and all tubing (Masterflex L/S silicone) were autoclaved prior to reactor assembly and chlorine disinfection. After 24 hours with the inner drum rotating, the chlorine solution was drained and a sterile feed of Nanopure water was initiated for 24 hours. Following the 24-hour flushing period the reactor effluents were sampled to insure there was no remaining chlorine residual. A sterile sodium thiosulfate solution was added to remove any remaining free chlorine prior to inoculation.

Nutrient Amendments

Three different carbon source mixtures used in separate experiments were composed of amino acids, carbohydrates, or humic material and were supplied at a rate of 0.25 mL/minute from concentrated stock solutions. The amino acid mixture was composed of L-glutamic acid, L-aspartic acid, L-alanine, and L-serine. The carbohydrate mixture was composed of D(+)-glucose, D-galacturonic acid, D(+)-galactose, and D(-)-arabinose. All substrates were HPLC grade purchased from Sigma. Amino acids and carbohydrates were supplied as equimolar carbon concentrations of each individual compound in the group. The humic substrates were extracted from Elliot Silt Loam Soil, BS102M, purchased from the International Humic Substances Society (IHSS). The humics concentration was based on total organic carbon (non-purgeable) analysis of a concentrated stock solution. A concentrated stock solution was prepared by adding IHSS humic material (100 g) to 1 L of 0.1 N NaOH solution, mixed for 24 hours, and then centrifuged at 10,000 x g and 4° C to remove particulates. The concentrated humic stock solution (approx. 1000 mg C/L) was not autoclaved and was stored in the dark at 4° C. After addition of the humic stock solution to the sterile Nanopure water, the pH was adjusted to approximately 7 (± 0.2) using 2N HCl. Nitrate (as KNO_3) and phosphate (as $\text{KH}_2\text{PO}_4 + \text{K}_2\text{HPO}_4$) were supplied at a rate of 0.25mL/ minute at concentrations of 612 and 591 μg /L (9.9 and 6.2 μM), respectively. Carbon feed was controlled in order to provide nominal influent concentrations at three different levels: (1) 500 ppb, (2) 1000 ppb, and (3) 2000 ppb. The reactors were carbon limited under all substrate feed concentrations.

Reactor Operation

Figure 2.2 shows the substrate and chlorine feed protocols used during each experiment. Initially each reactor was fed a substrate concentration of 500 $\mu\text{g C/L}$ until it reached a pseudo steady state based on total and culturable cell counts (14 to 30 days). The substrate concentration was then increased to 1,000 $\mu\text{g C/L}$, the reactors were operated until a new steady-state was attained, and then chlorine addition to one of the two reactors was started. The chlorine dose was slowly increased until a free chlorine residual of 0.15 to 0.2 mg/L was measured in the effluent. The reactors were operated until another pseudo steady-state was reached, then the substrate concentration was again increased to 2,000 $\mu\text{g C/L}$, the reactors operated until pseudo steady-state was again reached, and then the chlorine feed dose was slowly adjusted until the effluent residual of the chlorinated reactor was 0.15 to 0.2 mg/L. Triplicate experiments (duplicate for the humics) for each of the three substrate groups were performed. Experiments were typically run for 90 to 150 days.

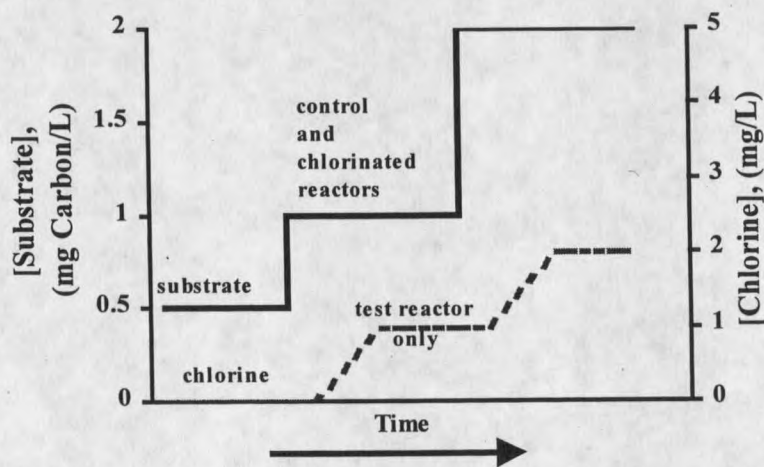


Figure 2.2 – Carbon and chlorine dosing protocol for all experiments.

Bacterial Counts

Bacteria in effluent samples were counted during reactor operation, especially during pseudo steady-state, prior to increasing the carbon feed and/or chlorination. Bacteria were fixed with 37% filter (0.2 μm nylon syringe filters, Fisher Scientific) sterilized formalin (final concentration 0.7%), filtered onto Poretics filters (0.2 μm , prestained with Irgalan black), stained with 0.5 mL of 100 mg/L of 4',6-diamidino-2-phenylindole (DAPI, Sigma) in autoclaved and filtered nanopure water and counted by epifluorescence microscopy. Both the dilution water and the reactor effluent were monitored in this manner. A minimum of 30 fields or greater than 400 cells were counted. An Olympus BH-2 microscope and a Nikon PlanFluor objective (N.A., 1.30) at magnifications of 1000-1500X were used for counting and biovolume determinations.

Biovolume Determinations

Samples stained for direct counts were also used to determine the biovolumes. Image analysis was performed with the use of a charge coupled device (CCD) camera (Optronics, 470T) connected to Image Pro (ver 3.0, Media Cybernetics) computer software for archiving of images on CD-ROM. A minimum of 20-30 random images were collected and analyzed for morphological parameters (length, width, area, perimeter, feret box, roundness, etc.) with the use of the freeware image analysis program Image Tool (ver. 1.27) from the University of Texas Health Sciences School at San Antonio (UTHSCSA, available at <http://ddsdx.uthscsa.edu/dig/itdesc.html>). Since bacteria were predominantly rods and cocci, biovolume was calculated as follows: Vol =

$\pi/4 \times w^2 (l - w/3)$ where l is length and w is width (Bjornsen, 1986; Norland, 1993). Biovolume determinations were performed on 2 to 5 separate samples from the reactors during pseudo steady state operation at the 1000 and 2000 ppb carbon feed levels. Calibration for morphometric analysis was performed with a calibration slide etched with 10 μm wide lines. Comparison of calibration lines to the measurement of 1 μm latex spheres (Polysciences, Ltd.) showed good agreement (measured diameter = $1.0 \pm 0.1 \mu\text{m}$ S.D.). Color images were converted to grayscale and thresholding was performed manually. Images were analyzed at a final magnification of 2000 to 6000X depending on cell sizes. Microscopic magnifications of 1000 to 1500x resulted in approximately 10 to 14 pixels per μm .

Determination of Cell Carbon

Samples for total organic carbon (TOC) were processed immediately upon sampling. Water was filtered through sterile 0.2- μm pore-size nylon syringe filters (Fisher Scientific), prewashed with 35 mL of 0.1N HCl followed by 4 separate 35 mL washes with sterile Nanopure reagent grade water using a 30 mL sterile, disposable syringe that was also used for the sample. The first 10 to 20 mL of the filtered sample was discarded. All total organic carbon samples were acidified prior to analysis using 0.2 mL of 2N HCl in 40 mL of sample. All samples were analyzed for non-purgeable organic carbon (NPOC) using a Shimadzu TOC 5000A carbon analyzer with high sensitivity platinum catalyst by the high temperature combustion method and an auto-sampler. Purge times were 5 minutes minimum with medical grade oxygen. Standard curves for the TOC

5000A were developed using 4 concentrations of potassium hydrogen phthalate (Shimadzu), and the resulting linear calibration curve typically had a coefficient of determination (R^2) of 1.00 and never less than 0.98. An external standard (from the standards used to develop the standard curve) was always run with each group of samples to correct for any variations in carrier gas flow rate. All glassware was precombusted at 500 °C for at least 4 hours, whenever possible. Auto-sampler vials were washed for at least 12 hours in a solution of concentrated sulphuric acid + NoChromix (Godax Laboratories, Inc.), rinsed 6 times with Nanopure reagent grade water, and precombusted at 300 to 400°C for 6 hours, minimum. Glass volumetric flasks and pipettes were acid washed (as above) and rinsed 6 times with Nanopure water. Parallel filtered and unfiltered samples from the reactor influent and effluent were used to determine the amount of carbon (non-purgeable organic carbon, NPOC) in the particulate and fluid phase. The difference between unfiltered and filtered samples represents the particulate organic carbon in the sample and is due to the bacterial cells. This value can be divided by the total cell count to give a population average for C_c in the sample. This C_c value can then be divided by the average biovolume of the cells in the sample to give a population average for C_v . System blanks comprising the reagent grade Nanopure water were run at the beginning and end of each sampling period. The coefficient of variation (CV) for all samples averaged 1.8% (range, 0.07 to 7.4%, $n = 215$). A minimum of 4 measurements was performed on each sample.

Statistical Analyses

Analysis of variance (ANOVA) was the main comparative statistical test. All ANOVAs were performed with the statistical program Minitab (ver. 12.1, Minitab, Inc.) using a general linear model (GLM). Prior to ANOVA, the data were tested for normality and homogeneity of variance. In those cases where transformation was required, the method of Box-Cox was utilized (where $Y' = Y^\lambda$, λ being a number which minimizes the standard deviation of the standardized transformed variable, if $\lambda = 0$ it is a Log_e transformation). Since the transformation is based on what factor (treatment group) is used and more than one can provide homogeneity of variance, the best transformation was chosen based on a visual inspection of the residuals resulting from the GLM ANOVA. Non-parametric tests were performed when no applicable transformation could be found. Probability distributions were analyzed with a statistical program, WQHYDRO, designed for water quality analysis (Aroner, 1994) by the method of unbiased probability weighted L-moments (Hosking, 1990; Vogel and Fennessey, 1993; Vogel, et al., 1993; Hosking and Wallis, 1987). L-moment diagrams were constructed for visual inspection of the data and estimation of the underlying probability distribution. See Appendix I, Part I for an in depth description of moments, L-moments, probability weighted moments (PWMs), their relationships with each other and how they apply to the generalized Pareto distribution (GPD).

Distributions included in the software, WQHYDRO, are generalized Pareto, generalized logistic, lognormal, generalized extreme value, Weibull (EV 3), Pearson 3 (gamma), Gumbel (EV 1), uniform, normal, and exponential. Probability weighted (L-)

moments (unbiased) were used for parameter estimation of the 3 parameter generalized Pareto. The probability density function (pdf) for the GPD is

$$(1) f_X(x) = \alpha^{-1} [1 - \kappa(X - \beta)/\alpha]^{1/\kappa - 1}$$

and the cumulative distribution function (cdf) is

$$(2) F_X(x) = 1 - [1 - \kappa(X - \beta)/\alpha]^{1/\kappa}$$

where α , β , and $\kappa \neq 0$. κ is referred to as the shape parameter, α is the location parameter, β is the origin parameter and X is the measured biovolume. The program constrains the value of κ between the bounds of -0.7 and +0.7, which did not affect its applicability for any of the datasets encountered in these experiments. The modified Anderson-Darling (AU^2) test statistic and r^2 values determined for the distributions arise from linearization of the GPD. This can be done for any cdf which is expressible in an invertible manner and results in a quantile function for the GPD that is equal to:

$$(3) Q_p = \beta + (\alpha/\kappa) * [1 - p^\kappa]$$

where p is the probability that $x \geq X$ and Q_p is that value of biovolume (x) which is exceeded with probability p . A plot of Q_p (y-axis) vs. $1 - p^\kappa$ (x-axis) will be linear with y-intercept equal to β and slope equal to (α/κ) if the distribution follows a GPD. Although this plotting point method can be used to estimate the parameters of the distribution, the unbiased probability-weighted L-moment method was used to estimate the parameters and the linearization was used to estimate the goodness of fit to the theoretical distribution. Some properties of this distribution are given below in A-D (also see Appendix I).

A) For $\kappa < 0$ biovolume is bounded below so that $\beta \leq X \leq \infty$. For $\kappa > 0$, biovolume is bounded both above and below, so that $\beta \leq X \leq (\beta + \alpha/\kappa)$. If both $\beta = 0$ and $\kappa = 0$, the GPD reduces to the exponential distribution. When $\beta = 0$ and $\kappa = 1$ the GPD reduces to the uniform distribution.

B) The GPD distribution parameters can be related to conventional product moments, PWMs, or L-moments. If they exist, the mean, variance, and skewness are,

$$\mu = \beta + \alpha/(1+\kappa)$$

$$\sigma^2 = \alpha^2/[(1+\kappa)^2(1+2\kappa)]$$

$$\gamma = 2(1+2\kappa)^{1/2}(1-\kappa)/(1+3\kappa)$$

Thus, the distribution has infinite variance when $\kappa \leq -1/2$, infinite skewness when $\kappa \leq -1/3$, and infinite kurtosis when $\kappa \leq -1/4$ (Hosking and Wallis, 1987; Hosking, 1990). Hosking (1990) showed that the L-moments λ_r , $r = 1, 2, \dots$, of any real-valued random variable X exist, if and only if X has a finite mean. This implies that a distribution may be specified by its L-moments even if some of its conventional moments do not exist, as is the case for several of the datasets collected in this study. The parameters of the distribution in terms of L-moments are

$$\kappa = (1-\tau_3)/(1+\tau_3)$$

$$\alpha = \lambda_2(1+\kappa)(2+\kappa)$$

$$\beta = \lambda_1 - \alpha/(1+\kappa)$$

where λ_r , $r = 1, 2, 3$, are the first three L-moments, τ_3 is the ratio λ_3/λ_2 and is termed the L-skewness.

C) For estimation of parameters, the cdf of the GPD can be inverted to obtain the quantile function

$$X_p = \beta + (\alpha/\kappa)[1-p^\kappa]$$

where p is the exceedance probability and X_p is that value of the biovolume which is exceeded with probability p . Substitution of sample estimates of λ_r into the above equations (B) produces L-moment estimators of the quantile function. These L-moment estimators always produce feasible values for the estimated parameters.

D) Derivation of the relationship between L-moments and the GPD model parameters was performed by Hosking (1990) such that

$$\lambda_1 = \beta + \alpha/(1 + \kappa)$$

$$\lambda_2 = \alpha/[(1 + \kappa)(2 + \kappa)]$$

$$\tau_3 = (1-\kappa)/(3-\kappa)$$

$$\tau_4 = (1-\kappa)(2-\kappa)/[(3 + \kappa)(4 + \kappa)].$$

where τ_3 was defined above and τ_4 is termed the L-kurtosis and is equal to λ_4/λ_2 . Hosking also derived an approximate relationship between L-kurtosis and L-skewness, which is useful for comparing the theoretical L-moment relationships with sample estimates

$$\tau_4 = 0.20196\tau_3 + 0.95924\tau_3^2 - 0.20096\tau_3^3 + 0.05061\tau_3^4$$

A second type of Pareto (Type II) distribution that has the probability density function

$$(3) \text{ pdf } (x) = c(K + D)^c (x + D)^{-(c+1)}$$

and a cumulative distribution function

$$(4) \log[\text{prob}(x \geq X)] = c \log(K + D) - c \log(X + D)$$

was described by Vidondo et al. (1997) and a similar form can be found in Johnson, et al.(1987). This equation can be used to estimate parameters of the distribution by non-linear regression. This function was used to solve for the parameters c , K , and D by the curve-fitting program TableCurve-2D (ver. 4, SPSS Inc.). While the parameters of this distribution are different from the first GPD, it contains similar information. The shape parameter is c , the location parameter is K , and D is an additive constant. This function was used to compare the distribution of the average biovolume (μm^3) in the samples and provide different estimates for the distribution that can be compared across functional groups by using the two different forms for a Pareto distribution. Due to the extreme trends in the residuals when performing non-linear regression on the biovolume probability distributions, this method could not be used for a comparison of the datasets and only the (L-)PWM results are reported for the biomass distributions. Curve fits for the biovolume vs. C_v were performed in TableCurve 2D in order to derive an empirical relationship for predictive purposes. Table 2.1 contains the parameters and method of estimation for the different Pareto distributions.

Distribution	Shape	Location	Origin/Constant	Estimation Method
GPD1	κ	α	β	L-moments
GPD2	c	K	D	non-linear regression

Table 2.1 – Comparison of distribution type, parameters, and method of analyses.

Results

Biovolumes were determined for bacterial communities subjected to three different treatment protocols. These measurements were made for the calculation of bacterial carbon production, observed yields, and growth rates as shown in Chapter 4. The treatment factors included substrate type, substrate concentration, and chlorination. Control measurements were made from parallel reactors not receiving chlorine. Effluent cell volumes were determined at two influent substrate concentrations of 1000 ppb C and 2000 ppb C. Concurrent organic carbon (particulate and dissolved) analyses allowed me to calculate average C_c and C_v values for the individual datasets. The GPD was found to fit 68 of 72 biovolume datasets as well as the C_v data. Distribution identification performed by the method of L-moments, provided a fast and easily interpretable diagram of potential candidate distributions. Both biovolume and C_v probability distributions determined from the samples allowed the construction of probability plots that described the datasets very well and provided a predictive capability for both biovolume and C_v .

Distribution Identification

The distribution of a variable such as biovolume can be of considerable biological interest. Asymmetrical distributions can tell us that there may, perhaps, be selection for or against organisms falling in one of the tails of the distribution. They can also tell us about the appropriateness of specific statistical tests for the data in question. For example, biovolume can only take on values greater than zero, limiting the suitability of an

assumption of a normal distribution, which by definition is unbounded and ranges from negative infinity to positive infinity. Assuming that a variable such as biovolume follows a specific distribution is a mathematical convenience that would allow us to make predictions based on the parameters of the distribution.

See Figure 2.3 for an example of the percent frequency distributions of 3 datasets (one from each of the amino acids, carbohydrates, and humics experiments) from both the chlorinated and unchlorinated reactors. These histograms are the individual cell volumes compiled from three separate samples taken over the course of three days from each reactor. Figure 2.4 shows two individual datasets from the carbohydrates experiment fitted to the GPD using unbiased L-moments. The mean observed biovolume for Figure 2.4 panel A was $0.123 \mu\text{m}^3$ and the mean expected (i.e. from the parameters describing the distribution) biovolume was $0.123 \mu\text{m}^3$. For panel B, the observed value was $0.391 \mu\text{m}^3$ and the expected value was $0.391 \mu\text{m}^3$. This good fit was seen for all datasets that conformed to the GPD. The correlation between expected theoretical means (where $\text{mean} = \beta + [\alpha/(1 + \kappa)]$) calculated from the fitted parameter values and the observed, measured means was 0.9996 ($p < 0.0005$). Both figures 2.3 and 2.4 show the impact that chlorination has on the distribution of cell volume. In order to identify the probability distribution that best represented the cell biovolume data, L-moment diagrams were constructed and inspected visually. These diagrams are formed by plotting the probability weighted L-moments for the skewness ($\tau_3 = \lambda_3/\lambda_2$) versus kurtosis ($\tau_4 = \lambda_4/\lambda_2$) on the abscissa and ordinate, respectively.

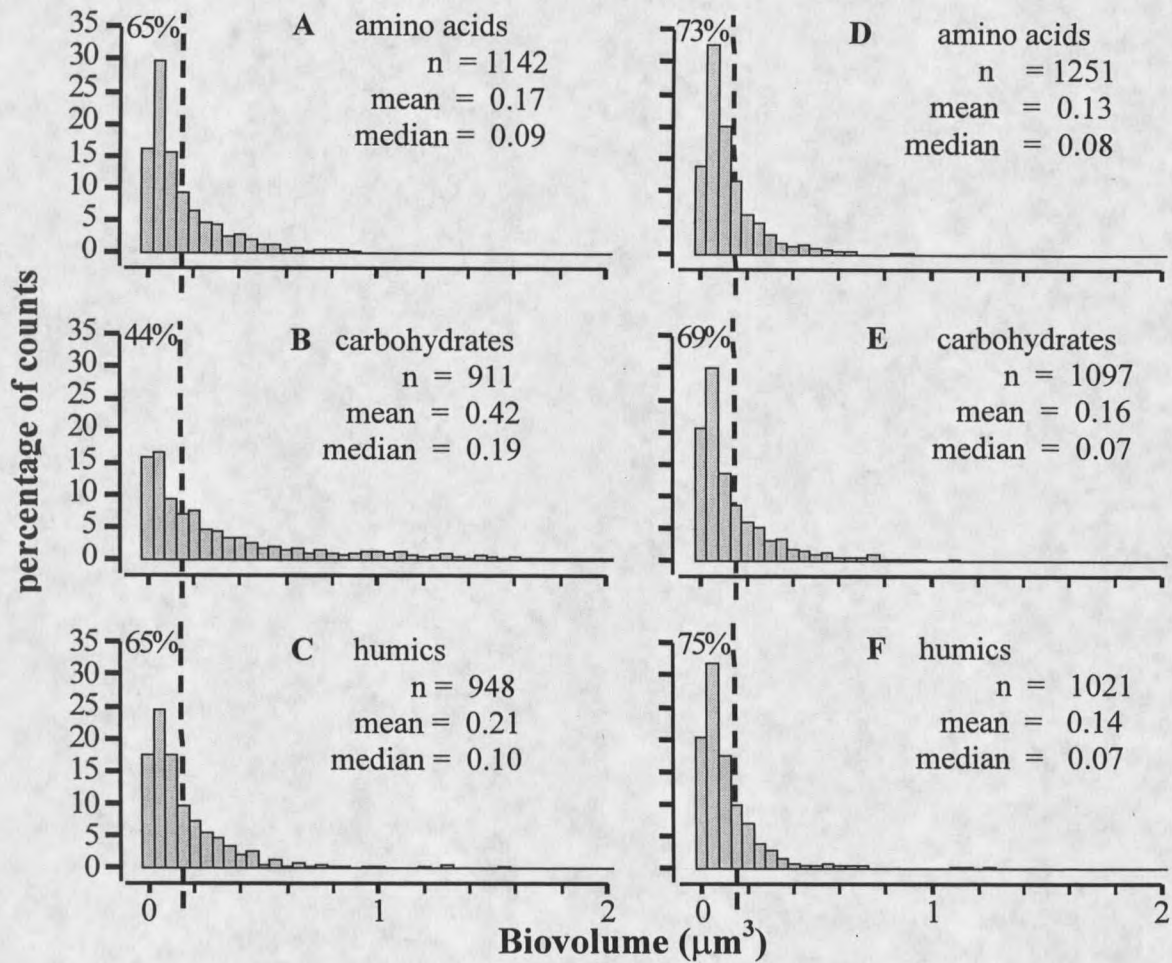


Figure 2.3 – Frequency distribution of cell volume as a percentage of the total number of cells measured. Note the highly skewed nature of the data. Panels A-C are for the control reactor and panels D-F are for the chlorinated reactors. Vertical reference line was arbitrarily set at $0.15 \mu\text{m}^3$ 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 $0.05 \mu\text{m}^3$.

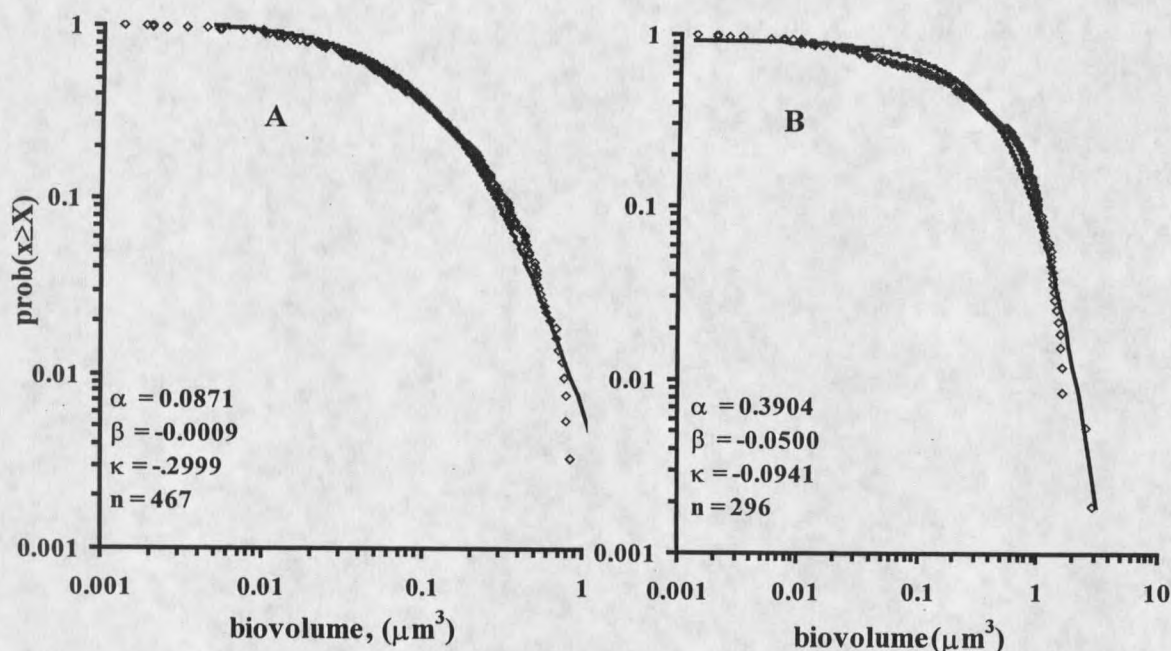


Figure 2.4 – GPD probability plots. α , β , and κ were determined by probability weighted L-moments. Solid line is the fit and symbols are the data points. Panel A dataset is from a carbohydrate experiment at a carbon concentration of 1000 ppb C, chlorinated reactor. Panel B is data from the corresponding 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 Quantile (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.

Figure 2.5 panel B shows the distribution fit for the converted C_v in units of fg $C/\mu m^3$. These values were obtained by dividing the difference for the amount of carbon between filtered (no cells) and unfiltered samples. The observed average was 268.15 fg $C/\mu m^3$ and the expected average was 268.15 fg $C/\mu m^3$, with 57% of the sample values below this average and 90% of the sample values below 560 fg $C/\mu m^3$. The sample standard deviation was 194 fg $C/\mu m^3$ while an estimate of the parametric standard

deviation based on the parameters of the fitted distribution was $24 \text{ fg C}/\mu\text{m}^3$. Figure 2.5 panel A is an L-moment diagram showing representative results from seventy-two data sets comprising an average of 382 (range = 112 to 1862) individual volume estimations on bacteria from the influent and effluent of the reactors. Four of the seventy-two datasets gave poor r^2 values and high modified Anderson-Darling (AU^2) test statistics and were not included in the subsequent ANOVAs for the Pareto parameters. Two of these datasets followed a lognormal distribution and the other two followed an extreme value distribution. All of the relevant values from the distribution analysis of the 68 datasets are provided in Appendix I.

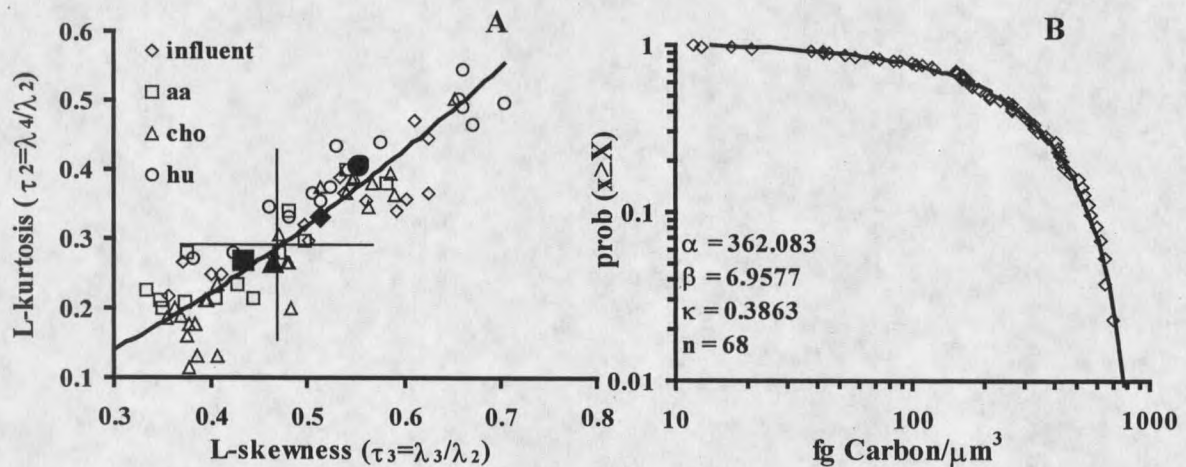


Figure 2.5 – Panel A is an L-moment diagram compiled from 72 datasets showing the overall trend of L-kurtosis vs. L-skewness for bacterial biovolume measurements. The solid line is the theoretical distribution of L-skewness vs. L-kurtosis for the GPD. Open symbols refer to individual datasets. Solid 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 Quantile plot.

Another indication that the GPD was adequately describing the probability distribution of biovolume can be seen through use of the equation given by Vogel, et al (1993) and shown in the Methods (statistics section D). The correlation between the measured sample L-kurtosis and theoretical sample L-kurtosis (of the GPD) calculated from the measured sample L-skewness was 0.8983 ($p < 0.0005$). Inclusive of the sixty-eight datasets the average r^2 value was 0.963 (± 0.0316 S.D.), and the average AU² statistic was 0.965 (± 0.874 S.D.).

Biovolume and GPD Parameter Analysis

As stated above, fitting a probability distribution to observed data does not provide for a mechanistic model of the data. Rather, it is a method that allows one to predict events based on the parameters of the distribution. Assuming that a variable such as biovolume follows a specific distribution is a mathematical convenience based on the approximate resemblance of the data to the distribution. If this assumption is appropriate, it would, however allow us to make predictions based on the parameters of the distribution. Table 2.2 contains mean and median volumes separated according to carbon source, carbon concentration, and reactor. For both the amino acids and humics control reactors mean and median volumes were significantly different at the two different carbon feed levels ($p < 0.03$ for all measures). Conversely, in the presence of chlorine no significant difference was detected in either the mean or median volume for any of the substrates when the carbon concentration was doubled (average $p = 0.445$, range = 0.177 – 0.825, $n = 6$, see Table 2.1). Both mean and median cell volume were not significantly

affected by increasing the carbon concentration in the carbohydrate experiments ($p = 0.506$ and $p = 0.471$ respectively).

Parameter	Treatment	Amino Acids		Carbohydrates		Humics	
		1000	2000	1000	2000	1000	2000
mean volume [#]	chlorinated	0.14 (0.03)	0.17 (0.03)	0.3 (0.08)	0.41 (0.08)	0.12 (0.02)	0.15 (0.02)
	control	0.23 (0.03)	0.34 (0.03)	0.6 (0.2)	0.8 (0.2)	0.14 (0.04)	0.31 (0.05)
median volume [#]	chlorinated	0.10 (0.02)	0.10 (0.02)	0.13 (0.05)	0.23 (0.05)	0.07 (0.01)	0.06 (0.01)
	control	0.14 (0.02)	0.22 (0.02)	0.3 (0.2)	0.5 (0.2)	0.07 (0.02)	0.15 (0.02)
n*	chlorinated	3 (6)	3 (7)	2 (7)	2 (7)	2 (6)	2 (3)
	control	3 (6)	3 (6)	2 (6)	2 (6)	2 (6)	2 (3)

Table 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.

Table 2.3 contains the Pareto parameters for carbon type and concentration according to the chlorinated or control reactor resulting from individual analysis. Both the shape parameter κ and location parameter α were found to be dependent on the type of carbon supplement added to the system. Not surprisingly, the mean and the median were highly correlated with the location parameter α ($r^2 = 0.981$ and 0.990 , respectively, $p < 0.0005$ for both) and a partial correlation analysis indicated that this was not due to the correlation between the mean and the median ($r^2 = 0.98$, $p < 0.0005$). The origin

parameter β was negatively correlated with the mean and median ($r^2 = -0.663$ and -0.591 , respectively, $p < 0.0005$ for both). The shape parameter κ was negatively correlated with carbon type ($r^2 = -0.536$, $p < 0.0005$), indicating that as the substrate progressed in the order - amino acids, carbohydrates, humics - the shape parameter became more negative. The shape parameter κ was also negatively correlated with both the L-skewness and L-kurtosis ($r^2 = -0.658$ and -0.485 respectively, $p < 0.0005$ for both).

Parameter		Amino Acids (ppb)		Carbohydrates (ppb)		Humics (ppb)	
		1000	2000	1000	2000	1000	2000
α	chlorinated	0.12 (0.02)	0.13 (0.02)	0.19 (0.07)	.32 (0.07)	0.09 (0.01)	0.07 (0.01)
	control	0.21 (0.04)	0.29 (0.04)	0.6 (0.2)	0.7 (0.2)	.09 (0.02) ¹	0.20 (0.02) ¹
β	chlorinated	0.009 (0.006)	0.013 (0.005)	-0.011 (0.005)	-0.014 (0.005)	0.0111 (0.0006)	-0.001 (0.0009)
	control	0.007 (0.006)	0.013 (0.006)	-0.05 (0.02)	-0.04 (0.02)	0.001 (0.003)	-0.007 (0.004)
κ	chlorinated	-0.04 (0.04)	-0.12 (0.04)	-0.37 (0.06)	-0.27 (0.06)	-0.30 (0.03) ²	-0.56 (0.05) ²
	control	-0.09 (0.07)	-0.11 (0.07)	-0.14 (0.04)	-0.22 (0.04)	-0.26 (0.08)	-0.3 (0.1)

Table 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.

One-way ANOVAs were performed to compare the distribution parameters of the different communities at the different treatment levels. Comparison of reactors at the two carbon levels indicated that carbon concentration had relatively little effect on the parameters of the distribution, whereas chlorination had a significant impact in the

presence of each carbon source (Tables 2.2 and 2.3). Both individual comparisons and a nested analysis (reactors nested within carbon levels) support this conclusion. Carbon feed concentration only played a significant role in parameter estimates for α in the humic control reactor ($p = 0.004$) and κ in the humic chlorinated reactor ($p = 0.003$). All other p -values averaged 0.429 (range 0.102 – 0.813, $n = 16$). Using the average Pareto parameters in Table 2.3 for determination of the theoretical average biovolume and comparing these to the observed average biovolumes from Table 2.2 shows good agreement. For example, the correlation between the theoretical and observed volumes at 1000 ppb carbon was 0.930 ($p = 0.007$) and at 2000 ppb carbon the correlation was 0.993 ($p < 0.0005$). Irrespective of the mean volume, the percentage (abundance) of the population biovolume below the mean was consistent at $68 \pm 4\%$ ($n = 12$) based on the average Pareto parameters.

Overall analyses of the data by a general linear model (GLM) ANOVA with carbon level (1000 and 2000 ppb carbon) nested within carbon type (amino acids, carbohydrates, and humics) and with reactors (control and chlorinated) nested within carbon type and carbon level were also used to analyze the parameters of the Pareto distribution. This supported the conclusions of the separate analysis shown in the preceding paragraphs. However, the origin parameter β could not be analyzed by this method due to the inability to induce homogeneity of variance between carbon sources. The analysis indicated that the carbon source was significantly different for all parameters ($p < 0.001$). In order to satisfy the assumptions inherent to an ANOVA, the location parameter α , medians, and means were transformed by the method of Box-Cox

to induce homogeneity of variance. This analysis corroborated the preceding statements concerning reactor, carbon level, and carbon type. The same analysis on the measures for L-kurtosis and L-skewness, which were normally distributed and had homogeneous variances, supported the conclusion of the above analysis on the parameters of the distribution. Different carbon sources produced distributions that had significantly different ($p < 0.001$) L-kurtosis and L-skewness whereas carbon levels within carbon source were not significantly different ($p = 0.667$ for τ_3 and $p = 0.590$ for τ_4). No significant difference was found for reactors within carbon levels within carbon types ($p = 0.085$ for τ_3 and $p = 0.071$ for τ_4). See Figure 2.5 panel A for a graphical representation of the L-kurtosis vs. L-skewness data as determined from the PWMs.

Table 2.4 contains a similar analysis on the parameters of the GPD for the normalized biovolume data (all volumes in a dataset divided by the smallest volume within that dataset). As stated above, the origin parameter β could not be analyzed by this method, however, non-parametric tests indicated that there is a significant difference in the medians and p-values are reported in the legend for Table 2.4. The shape parameter κ for the humics test reactor was significantly different for the two carbon levels, see Table 2.3 for these estimates. Probability distribution analysis for the averaged volumes from each dataset by non-linear regression [according to equation (4) in the Methods section] was significant and since the residuals showed no trend, the estimates for r^2 and the standard error of the fit are reliable. This analysis is shown in Table 2.5 and Figure 2.6 for both the raw and normalized probability distribution data for cell biovolume. The shape parameter is independent of normalization and is the same for both graphs. Since

the GPD1 (equation 1 and 2) has been much more extensively studied, there is more information that can be derived from this distribution and the parameters used for its estimation. The observed average volume was $0.267 \mu\text{m}^3$ and the average volume expected from the fitted parameters of the distribution was $0.267 \mu\text{m}^3$ while the average normalized volume was 2.29. Sixty-eight percent of the biovolume measurements were below this mean value as determined from the fitted parameters, similar to the value determined for the individual datasets. Since the data fit the GPD, the parametric variance can be calculated from the formula $\sigma^2 = \alpha^2 / [(1 + \kappa)^2 (1 + 2\kappa)]$, $\kappa > -1/2$ (Hosking, 1990), and is equal to 0.078 so that an estimate of the parametric standard deviation would be $0.034 \mu\text{m}^3$, where $n = 68$ (i.e. population), as compared to the sample S.D. of $0.25 \mu\text{m}^3$.

Parameter	Amino Acids		Carbohydrates		Humics	
	mean	p-value	mean	p-value	mean	p-value
α	chlorinated 34 (12)	0.518	95 (61)	0.022	71 (15)	0.494
	control 43 (10)		208 (34)		87 (15)	
β	chlorinated 0.9 (0.06)	0.877	-5 (4)	*	0.3 (0.9)	0.507
	control 0.9 (0.06)		-17 (4)		-0.7 (0.9)	
κ	chlorinated -0.10 (0.04)	0.723	-0.32 (0.04)	0.019	na	na
	control -0.08 (0.04)		-0.18 (0.04)		-0.28 (0.06)	

Table 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.

Distribution Parameter	α	k	β	D	κ	c	r^2	AU^2 SE
Biovolume (μm^3)	0.153	0.065	0.0659	0.496	-0.24	4.05	0.952 0.993	0.227 0.035
Normalized Biovolume	2.01	0.849	0.865	6.519	-0.24	4.05	0.952 0.993	0.227 0.035

Table 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 and c.

The distribution of C_c fit a 3 – parameter Weibull (EV 3, where the cdf = $1 - \exp[-((X - \epsilon)/(\beta - \epsilon))^\alpha]$) probability distribution at an α -level of 0.01 ($r^2 = 0.9773$, $AU^2 = 0.348$) while the distribution of C_v followed the GPD. A comparison between the distribution suggested by Vidondo et al. (1997), represented by the GPD2 (equations 3 and 4) and the GPD1 (equations 1 and 2) was performed for the biovolume distribution and is presented in Figure 2.6 and Table 2.5 for normalized and raw data. Both the table and figure are included for reference and comparison to non-bacterial biomass distributions.

The parameters describing the probability distribution of biovolume for the individual substrates without regard to carbon levels are graphed in Figure 2.7, Panel A (control) and B (chlorinated). Both graphs are predictive in that the probability of measuring a cell equal to or greater than the specified biovolume can be read off of the probability (y) axis. Some examples are provided for each substrate at probabilities equivalent to 1 in 10, 1 in 100 and 1 in 100,000 cells respectively. Probabilities for the

amino acids are provided below the horizontal line to the left of the curves with the humics above the corresponding line to the left of the curves. Values for the carbohydrates are provided to the right of the curves. For reference, the equivalent spherical diameter (ESD) is supplied on the biovolume axis. As an example, the probability of finding a cell with a biovolume greater than or equal to $1.46 \mu\text{m}^3$ with amino acids as the substrate is 0.01. Conversely, the probability of measuring a cell with a volume less than $1.46 \mu\text{m}^3$ is 0.99. Note that both axes are plotted on a $\log_{10}$ scale. The probability scale was extended to include a biologically relevant sample of 1×10^6 cells.

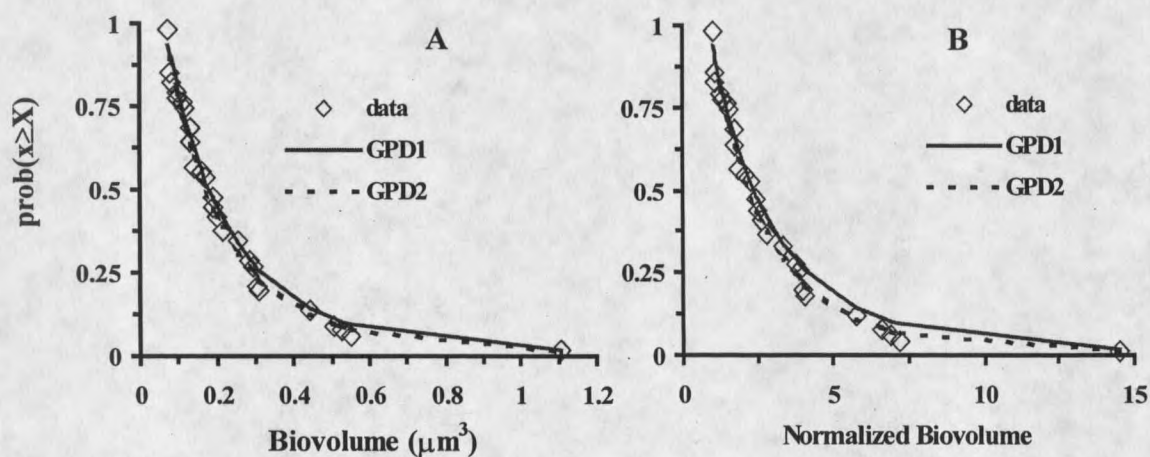


Figure 2.6 - Comparison of the raw and normalized data fits for the GPD1 (see equation 2) and the GPD2 (see equation 4). Refer to Table 2.5 for parameters and fits for the 2 graphs.

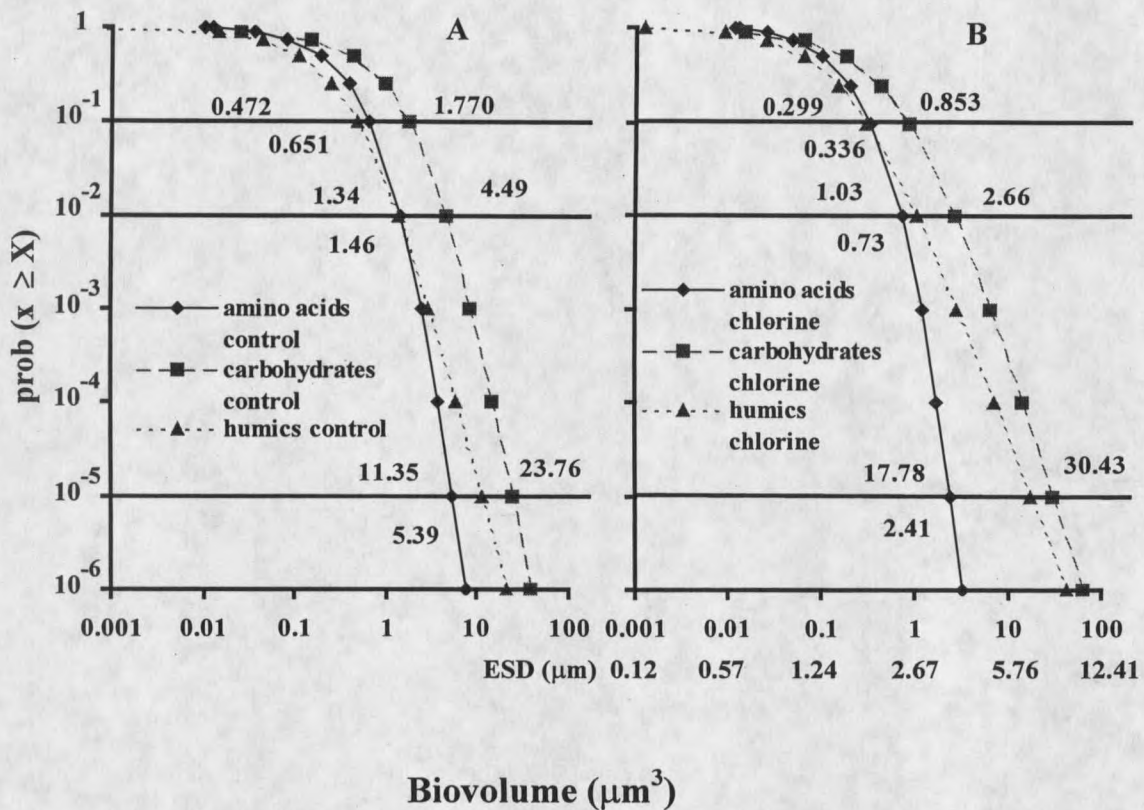


Figure 2.7 – Probability plots on a log-log scale from the average Pareto parameters for the control reactor (Panel A) and chlorinated reactor (Panel B). Horizontal lines at 10^{-1} , 10^{-2} , and 10^{-5} indicate the probability of finding a cell with a volume equal to or greater than the indicated 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.

Analogous to Figure 2.7, irrespective of carbon source, carbon level, and chlorination, the parameters of the Pareto distribution derived from the biovolume data of

Figure 2.6 (see Table 2.5) and C_v data of Figure 2.5 panel B were used to construct probability plots assuming that the parameters effectively describe the underlying population distribution. This data is shown in Figure 2.8. The probability scale is extended to include the probability of selecting a cell from a population of 1×10^6 cells. Note that this graph does not indicate any relationship between biovolume and C_v . Rather, it indicates that as the biovolume increases, the probability of detecting a cell with the indicated (or greater) volume decreases. Similarly, as the C_v increases the probability of detecting or measuring a cell with the indicated C_v (or greater) decreases.

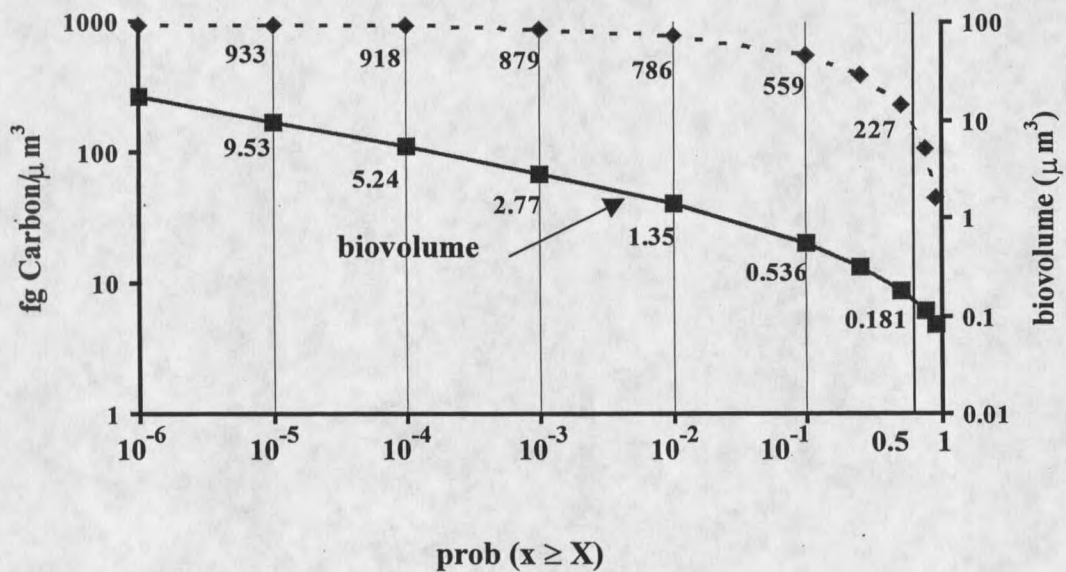


Figure 2.8 – Probability plots for biovolume and fg carbon/ μm^3 for the Pareto distribution derived from Table 2.5 and Figure 2.6. Plot is derived from all of the experimental data irrespective of treatments. Upper curve is fg carbon/ μm^3 and the lower curve is biovolume. Note that all axes are scaled 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 fgC/ μm^3 (268) and for biovolume, the mean was 0.316 μm^3 (0.298).

Relationship between Biovolume and Carbon per Biovolume

A number of other investigators have found a relationship between biovolume and C_v (or C_c). Identification of a relationship between biovolume and C_c or C_v would enable a more reliable conversion between biovolume and biomass. In our datasets we found no relationship between biovolume and carbon/cell, however a significant relationship was found between biovolume and C_v . Panel A of Figure 2.9 shows the log-log transformed data as normally presented in the literature ($C_v = bVol^a$ or $\log[C_v] = \log b + a\log Vol$). The values of for a and b conform to the power law relationship normally used to describe this type of data, where a is the location parameter and b is a constant that relates the amount of carbon present at unit volume. Since this is a Model II regression where both the x and y variable are subject to error, the ordinary least squares method for determining the regression line is inappropriate, although it is included for reference. While a number of alternative methods exist, no one method has been agreed upon by statisticians. As such, Sokal and Rohlf (1995) provide 3 alternative methods for estimation in Model II cases. These methods include the geometric mean (GM) regression, Bartlett's three-group method, and the slope of the major (principal) axis method. Bartlett's three-group method provides a slope of -1.083 (95% C.I. is -1.223, -0.993), not significantly different from ordinary least squares regression where the slope is -1.086 (-1.23, -0.934). The GM method provides a slope of -1.24 (-1.392, -1.087) and the major axis method has a slope of -1.27 (-1.407, -1.114). The y -intercepts are 1.529, 1.543, 1.434, and 1.407 respectively. Backtransformation of the intercepts provides an estimate of fg carbon at unit volume (i.e. $1 \mu m^3$) and are equivalent to 33.8, 34.9, 27.2,

and 25.6, respectively. All 4 regression lines are given in figure 2.9 Panel A, although they may be difficult to discern due to overlap of the GM line with the major axis line and overlap of the Bartlett's line with the ordinary least squares line. Inclusive of the three more acceptable methods for this type of regression, one (Bartlett's) provides a slope that contains -1.0 within the 95% confidence interval indicating an equal distribution of biomass across the measured biovolumes whereas the other two (GM and major axis regression) provide a slope significantly different than -1.0 indicating that biomass is unequally distributed across the range of volumes measured. Given this information and the preceding results on the probability distributions for biovolume (Fig. 2.6 B) and C_v (Fig. 2.5 B) as well as the analysis of the individual probability distributions, we can conclude that the relationship between volume and C_v is both non-linear and unequally distributed. Panel A of Figure 2.9 shows the power relationship between biovolume and C_v that has been noted by others on a log-log plot. Due to the impractical values for C_v produced by the log-log model of biomass distribution in the tails of the fit, a different, more reasonable approach was investigated through a weighted non-linear regression of the data. Panel B of Figure 2.9 shows a line fit to the averaged biovolumes divided into 7 size classes. This weighted non-linear regression was performed on the untransformed variables. Both x and y error bars (standard deviations) are included due to pooling of the data for both axes. Pooling of the C_v data into various biovolume classes and basing the weight on the inverse variance of C_v produced a variety of possible fits. The ultimate equation ($C_v^{-1} = a + bVol^{1.5}$) was chosen based on the behavior of the residuals (Gaussian), the lowest fit standard error, and the largest F-

statistic (102.3) The use of different weights (based on the inverse variance of the biovolume or a combination of the inverse variance of the biovolume and C_v) did not affect the final choice of fits, since the same function arose from several different choices of weights and division of size classes. Confidence (95%) intervals for a were 0.0005 to 0.0022 and the confidence intervals for b were 0.017 to 0.043. Better fits were attainable, but the tails of these fit-lines produced results inconsistent with both theoretical and practical considerations.

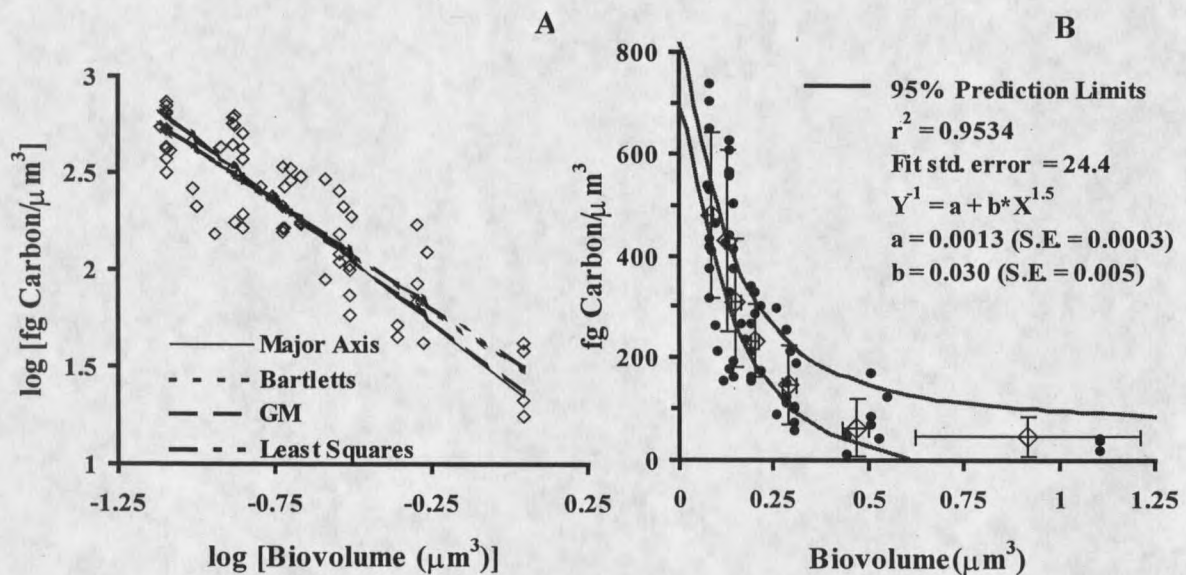


Figure 2.9 – Relationships between biovolume and C_v . Panel A shows the linear fit for the log-log transformed data, implying the power law relationship. Four regression lines are shown: GM, geometric mean method, Major axis method, Bartlett's 3-group method, and ordinary least squares. Panel B shows 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 ($1/\text{variance}$) data, $n = 68$.

Discussion

Probability Distributions

Vidondo, et al. (1997) noted the equivalence between the normalized biomass-size spectra (NB-SS) models used in the ecological analysis of biomass distributions and the Pareto distribution. They suggested that size distributions of organisms should be viewed from the standpoint of probability distributions. Analyzing this type of data within the context of probability distributions would allow ecologists to take advantage of progress in analyzing, identifying, and describing the underlying probability distribution based on individual measurements and not on the population (or size class) averages as is most common. Their prediction that a 3-parameter, type II Pareto distribution would fit biomass-size data better than the normalized biomass-size spectra models (i.e. power models) currently in use was borne out in our analysis of mixed microbial communities. The ability to analyze the biovolume data as individual measurements rather than community or size class averages allowed us to determine that biovolume did conform to a generalized Pareto distribution. As well, biovolume and C_v , as community averaged measures also conformed to the GPD while C_c fit a Weibull distribution, another type of extreme value distribution.

Using L-moment diagrams, rapid and easy determination of the probability distribution that best fit the biovolume data from both natural and carbon amended

communities was accomplished. L-moments are analogous to conventional moments but are estimated by linear combinations of the order statistics. As shown by Hosking (1990) and Vogel and Fennessey (1993) the ability to discriminate between alternate distributional hypotheses is made much easier with L-moments as compared to conventional moments. They can characterize a wider range of distributions and are more robust against outliers when estimated from samples. These diagrams have proven to be valid and often superior alternatives for summarizing probability distributions in environmental samples (Vogel and Fennessey, 1993; Vogel, et al., 1993; Hosking, 1990).

Measurements of cell volumes in a variety of habitats have led to estimates of a practical lower limit for bacterial cell volume around $0.005 \mu\text{m}^3$ which corresponds to a diameter (assuming a coccoid cell) of $0.105 \mu\text{m}$ (Psenner and Loferer, 1997). While there were occasional individual measurements within our datasets below this value in addition to reported values in the literature (Loferer-Kroßbacher, et al., 1998) the only theoretical distributions that provided for estimates smaller than this were from both the control and chlorinated humics reactors. Using the average biovolumes from all measurements, the lower limit to the biovolume GPD was $0.065 \mu\text{m}^3$ (ESD = $0.136 \mu\text{m}$) and the probability that an average biovolume measured in any of the experiments being below this value was effectively zero. Lower limits to the biomass distribution were around $7 \text{ fg C}/\mu\text{m}^3$, from the largest of the cells we measured from the carbohydrates experiments where average volumes ranged from 0.6 to $0.8 \mu\text{m}^3$.

Biovolume Conversion

While our measurements of the average C_v and C_c (42 ± 20 fg C/cell) were within the range of previously reported values, the obvious non-linear relationship implied that one should use a function for converting biovolume to biomass and not an individual value as is most often used. Since the parameters of the power function (a and b) have yet to be ascribed to any ecological function, the relationship is essentially empirical and the use of linear regression to fit a line to log-transformed data results in values of C_v that are inconsistent with current knowledge (Norland, et al., 1987; Bjornsen, 1986; Troussellier, et al., 1997). Using a biovolume from the low end of the bacterial size distribution can illustrate this problem of inconsistent values for carbon per unit volume. As an example, if we use a cell with a volume of $0.01 \mu\text{m}^3$, the conversion equation of Troussellier, et al. (1997) would result in a cell with 930 to 2353 fg C/ μm^3 , Kroers' (1994) 2721 fg C/ μm^3 , and this study 4954 to 5248 fg C/ μm^3 . The range of values given in the previous sentence result from using the high and low error values from the particular study. While Kroer was the only investigator to explicitly constrain his relationship between the bounds of 0.08 to $0.22 \mu\text{m}^3$, this non-linear relationship can be better fit on the untransformed variables by any of a number of non-linear regression programs such as TableCurve-2D (SPSS Inc.).

Based on literature values for the dry matter density of bacteria (1200 to 1600 fg/ μm^3) (Troussellier, et al., 1997; Norland, et al., 1987; Bjornsen, 1986) values above 500 fg C/ μm^3 would comprise 63 to 83% of the dry weight of a cell, if we assume that 50% of the dry weight is carbon, as is commonly done (Loferer-Kroßbacher, et al., 1998;

Troussellier, et al., 1997; Bjornsen, 1986; Nagata and Watanabe, 1990; Nagata, 1986). Using the power function models of Troussellier, et al. (1997) this value for C_v is attained at $0.05 \mu\text{m}^3$, while Kroers' (1994) function attains this value at $0.159 \mu\text{m}^3$, and our relationship produces this C_v value at approximately $0.09 \mu\text{m}^3$. If we do not assume any relationship between dry weight and carbon content the same value of $500 \text{ fg C}/\mu\text{m}^3$ would only account for 31 to 42% of the dry weight. Values significantly above $500 \text{ fg C}/\mu\text{m}^3$ have been measured before and at least one suggested conversion factor is above this value (Bratbak, 1985). Through use of a weighted non-linear regression (Figure 2.8) our data was fit to a non-linear function and this value for C_v is attained at $0.08 \mu\text{m}^3$. However, by constraining the equation at the upper and lower tails, reasonable estimates for C_v are attained throughout the size range. For example, the probability distribution in Figure 2.8 indicates that the probability of measuring a cell with a C_v greater than 879 is less than (or equal to) 0.001. By weighting the upper and lower tails of the function, values for C_v in agreement with literature and theoretical considerations are achieved. This function assigns approximately equal weight to the size class range of 0.005 to 0.01 (748 to $738 \text{ fgC}/\mu\text{m}^3$) but 1.7 times as much weight as for a biovolume of $0.1 \mu\text{m}^3$ ($440 \text{ fgC}/\mu\text{m}^3$) while C_v at unit volume is $32 \text{ fg C}/\mu\text{m}^3$.

Norland, et al. (1987) was perhaps the first to recognize the potential power relationship between volume and C_v in bacteria, although data from Bratbak (1985) also indicates this potential relationship. However, one must be aware that the linearity implied by a log-log plot of the data means that the log transformed variables follow a linear distribution but the data itself are related in a non-linear manner. More recently,

Trousellier, et al. (1997) found an allometric relationship in both marine and non-marine Gram negative bacteria in a starvation study. Their relationship, $C_v = 25.73 \cdot V^{-0.981}$, indicates a slope not significantly different from -1.0 on a log-log plot and implies a constant ratio relationship between volume and C_v . This implies that bacterial biomass declines as the reciprocal of bacterial size with the result that smaller cells have a higher C_v content than larger cells. This slope of -1.0 (on an abundance scale) also implies equal biomass distribution throughout the size classes studied (Harris, 1994).

The relationship found in this study produces similar results for C_v at unit volume (33.8, 34.9, 27.2, and 25.6 depending on regression method). However, depending on the regression method, the slope of the line cannot be unequivocally assigned a value of -1.0. The large amount of data on biovolume distribution from this study indicated that the relationship is significantly non-linear. More important in the context of biomass estimation, the C_v data also followed a Pareto distribution, indicating significant non-linearity in the relationship between volume and C_v and this has important implications for biomass estimations in aquatic environments.

The average C_v measured in this study was $268 \pm 194 \text{ fgC}/\mu\text{m}^3$ ($n = 68$). This value is in the range of the previously reported values by Simon and Azam (1989) (133 - 400 $\text{fgC}/\mu\text{m}^3$), Kroer (1994) (21 to 1610 $\text{fgC}/\mu\text{m}^3$), Fagerbakke, et al. (1996) (32 to 103 $\text{fgC}/\mu\text{m}^3$), Loferer-Kroßbacher, et al. (1998) (382 $\text{fgC}/\mu\text{m}^3$), Bratbak (1985) (560 $\text{fgC}/\mu\text{m}^3$), Norland et al. (1987) (50 to 100 $\text{fgC}/\mu\text{m}^3$), and Bjornsen (1986) (350 $\text{fgC}/\mu\text{m}^3$). A variety of reasons for this rather large range of values have been postulated including: (1) methodological differences, (2) trophic status, (3) differences in growth

phase, and (4) species differences. Nagata and Watanabe (1990) found no significant difference in conversion factors for native cells and nutrient-amended cultures and suggested a conversion factor of $120 \text{ fgC}/\mu\text{m}^3$. Bjornsen (1986) and Trouselier, et al. (1997) found significant differences between marine and non-marine bacteria with marine organisms having the higher conversion factor (410 and $218 \text{ fgC}/\mu\text{m}^3$, respectively). A significant factor for this variability could be the non-linear relationship between volume and C_v . Biovolume measurements ranged over approximately 4 orders of magnitude for the various studies reported above while the conversion factor had approximately 15 fold variability. Using the data (biovolume and C_v) supplied in these papers, there was a significant negative correlation ($p = 0.005$, $r^2 = -0.468$) between volume and C_v (log transformed variables). One reason that some investigators detect a potential relationship between volume and C_v while others do not is likely due to the measurement of a small range of average biovolumes. While Bjornsen (1986) did not specifically state if he checked for a relationship, data presented in his paper was used for a correlation analysis and no statistically significant relationship (correlation) between volume and C_v was indicated. This could be due to the smaller range of volumes measured in Bjornsen's (1986) studies (0.08 to $0.338 \mu\text{m}^3$) as compared to those measured by Norland (0.01 to $7 \mu\text{m}^3$) (1987). One can see from Figures 2.7 and 2.8 that if the range of data measured were not wide enough a linear relationship would be detected at almost any point in the curve. We noted in our data that the removal of either of the extreme tails significantly affected estimates of both the slope and the intercept.

Table 2.6 shows the effect that choice of conversion factor or equation has on estimation of the average carbon content and total carbon present in a sample from this study. The histogram of biovolumes presented in Figure 2.3, Panel E (carbohydrates, chlorinated, $n = 1097$, mean = $0.16 \mu\text{m}^3$, and median = $0.07 \mu\text{m}^3$) was used for generating the values in this Table. Shown in the Table are a high ($560 \text{ fgC}/\mu\text{m}^3$) and a low ($50 \text{ fgC}/\mu\text{m}^3$) conversion factor from Bratbak (1985) and Norland, et al. (1987), respectively. Conversion equations are from Fagerbakke, et al. (1996), Kroer (1994), Trousellier, et al. (1997), Norland (1993), and the weighted non-linear regression derived in this paper. Two methods were used to calculate the average C_v and the total carbon content of the 1097 cells for which the biovolume was measured. The first method used the average biovolume and the appropriate factor or equation to calculate the average C_v (i.e. mean Vol x conversion = mean C_v). Multiplying this by the total biovolume of the 1097 cells gives the total amount of bacterial carbon present. The second method used the conversion factor or equation for converting each individual volume and then calculated the average from this array of values. This average C_v value multiplied by the total biovolume provides the total carbon in the sample. The table points out the discrepancy between the two methods when a non-linear relationship between volume and C_v is present. Depending on the conversion factor or equation used, the final result for the total amount of carbon present in the sample has an 11 to 26 fold difference between the high and low estimates, depending on the method of calculation. These differences can be magnified to a larger degree depending on the distribution of the biovolumes within the sample measured.

Reference	Average C_v (fgC/ μm^3)		Total Carbon (pg C)	
	Method 1	Method 2	Method 1	Method 2
Fagerbakke ¹ , (1996)	58.1	53.4	10.4	9.5
Kroer ² , (1994)	477.2	1301	85.2	232.2
Trousellier ³ , (1997)	152.8	843.6	27.3	150.6
Norland ⁴ , (1993)	107.9	118.0	19.3	21.1
This study ⁵	304.1	463.8	54.3	82.8
Norland ⁶ , (1987)	50	50	8.9	8.9
Bratbak ⁷ , (1985)	560	560	99.9	99.9

Table 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.

It has been noted by a number of investigators that reconciling the NB-SS models with certain datasets has not resulted in very much success (Gaedke, 1993; Ahrens and Peters, 1991; Vidondo, et al., 1997; Gasol, et al., 1991). For bacteria, this has been suggested to be due to a variety of factors inherent in the estimation of bacterial biomass and activity measurements. For example, not all bacteria in a given ecosystem are actively respiring and top-down or a bottom-up control of bacterial biomass can occur in different ecosystems. Harris (1994) noted that values for the slope of the NB-SS in the range of -1.2 are often associated with oligotrophic systems where the food chains are dominated by small organisms whereas values in the range of -0.8 are found in eutrophic systems dominated by larger organisms. While these biomass-size spectra models are generally used across trophic groups, it has also been noted that the slope of the within functional groups spectra is steeper than the overall trend of the slope between functional

groups (Vidondo, et al., 1997; Gasol, et al., 1997; Thiebaut and Dickie, 1992; Chrost, 1991; Dickie, et al., 1987; Gasol, et al., 1991).

The use of scaling or conversion factors that assume a simple allometric distribution of biomass with abundance have proven ineffective for the prediction of bacterial biomass. This paper shows that one possible reason for this is that the distribution of bacterial biomass/abundance cannot be effectively described by a linear function composed of the shape and location parameters of the simple type 1 Pareto distribution. Sixty-eight of seventy-two datasets could be effectively modeled with a generalized Pareto distribution that is commonly used in hydrology and other extreme value distributions (Castillo and Hadi, 1997; Johnson, et al., 1987; Vogel, et al., 1993; Hosking and Wallis, 1987; Lewis and Chatwin, 1997; Vogel and Fennessey, 1993; Vidondo, et al., 1997). The datasets that did not fit the GPD could be fit quite well by either a lognormal or a generalized extreme value distribution as determined from their respective L-moment diagrams.

The use of L-moment diagrams (Fig. 2.5) were shown to be highly applicable to identifying the underlying distribution for biomass abundance, whether it was measured by biovolume, C_v , or C_c . These diagrams have proven highly effective in the identification of the underlying distribution in highly skewed datasets such as those for average daily streamflow, flood maxima, and the analysis of other extreme environmental events (Vogel and Fennessey, 1993; Hosking and Wallis, 1987; Castillo and Hadi, 1997; Hosking, 1990; Vidondo, et al., 1997). The identification of the GPD as an underlying distribution for bacterial biomass abundance and its potential use for identifying the

underlying parameters that effectively describe other distributions could, as stated by Vidondo, et al. (1997) allow for the analysis of biomass abundance across functional groups in food webs in a manner that is independent of the normalized biomass-size spectra models now in use.

Treatment Effects

Carbon Source and Concentration While biovolume was affected by all three treatments in this study, carbon source and chlorine had a larger effect than substrate concentration. As such, the parameters describing the biovolume distribution were not significantly affected by the carbon concentration. Individual datasets as shown in Figures 2.3, 2.4, and 2.5(A) as well as Table 2.2 and the combined datasets as shown in Figure 2.7 convey the highly skewed nature of the biovolume distribution. Average volume measurements were in the range of those measured by others in natural environments, although the carbohydrate-fed communities were in the upper range of reported values at the higher carbon concentration (Fagerbakke, et al., 1996; Bratbak, 1985; Kroer, 1994; Sommaruga, 1995). Carbohydrates as a carbon source produced cells with the highest measured volumes whether or not chlorination was a factor. Disregarding the carbohydrates, doubling the carbon loading rate produced a significant increase in the average cell biovolume (both mean and median) in the control reactors.

Chlorination Effects Chlorination was the most significant treatment affecting the cell volumes within any substrate group as well as across substrate types. Chlorination overrides the impact of substrate concentration and there is no significant

increase in either the mean or median cell volumes for any of the carbon sources when chlorine is present. Cell sizes were approximately 2 times smaller in the presence of chlorine for all carbon sources. This is especially apparent at the higher carbon concentration. In the presence of chlorine, far more cells are found at the left end of the distribution, but the population is still highly skewed to the right due to the presence of a small number of large cells. The histograms in Figure 2.3 show that chlorination selectively removes larger cells from the distribution. Table 2.2 shows that chlorine is especially effective at limiting cell size in the presence of carbohydrates and less so when the carbon source is humics or amino acids. This occurs despite the maintenance of a continual free chlorine residual at approximately 0.15 mg/L for all experiments. The histograms in Figure 2.3 were typical of the datasets collected in these experiments as well as datasets from other studies that show the frequency distribution of biovolume (Vidondo, et al., 1997; Norland, 1993; Norland, et al., 1987). The values next to the y-axis in each panel refer to the percentage of cells below an arbitrarily chosen biovolume of $0.15 \mu\text{m}^3$ (ESD = $0.236 \mu\text{m}$). It can be seen that the percentage of the population below $0.15 \mu\text{m}^3$ increases in the presence of chlorine, especially when the carbon source is carbohydrates. This implies one of two plausible explanations for the observed trends: (1) chlorination selectively removes larger cells or (2) chlorination prevents the formation of larger cells. Removal of larger cells could simply be due to the larger absolute surface area that can react with the chlorine, resulting in cell lysis and the loss of larger cells in the planktonic phase. Conversely, it has been noted by many researchers that bacterial cells confronting stress often undergo a size reduction (Siegele and Kolter, 1992;

Kjelleberg, 1993; Troussellier, et al., 1997). In either case, a significant decrease ($p = 0.015$) in the data variance occurs and the population becomes more homogeneous, although still highly skewed. The overall sample skewness and kurtosis were slightly reduced (not significantly) by chlorination compared to the control reactor when the carbon source was amino acids or humics. The opposite trend is seen with the carbohydrates, where both measures (skewness and kurtosis) were significantly increased upon chlorination ($p = 0.02$ and 0.009 respectively).

While biovolume was affected by all three treatments in this study, carbon source and chlorine had a larger effect than substrate concentration. As such, the parameters describing the biovolume distribution were not significantly affected by the carbon concentration. The probability plots generated from the averaged biovolume GPD parameters shown in Figure 2.7 provide a predictive ability with respect to measuring a given biovolume from a sample. It can be seen that for 90 to 95% of the population, chlorination reduces cell volumes by a factor of approximately 2. An interesting note for these distributions is that for the largest of every 100,000 cells from the humics or carbohydrate fed communities, the chlorinated communities will produce a larger cell than the control reactors. This arises due to the nature of the distribution as well as from averaging the values for the individual distributions. In the Methods section we mentioned that the skewness and kurtosis become infinite when the shape parameter, $\kappa \leq -1/3$ or $\kappa \leq -1/4$ respectively. For the humics and carbohydrates chlorinated reactor, κ is -0.39 and -0.32 respectively. Thus, the probability of measuring very large cells that cause

this highly skewed (or kurtotic) distribution will be larger than for the unchlorinated reactors.

Conclusions

- 1) Bacterial biomass-size distributions were modeled extremely well by a 3-parameter generalized Pareto probability distribution. Average biovolumes from all datasets were adequately described by two different versions of the Type II Pareto distribution.
- 2) While average biovolume increased significantly with a doubling of the carbon loading no significant changes in the parameters describing the probability distribution of volume were detected, with the exception of the location parameter for the humics fed control communities.
- 3) Chlorine overrides the impact of increasing substrate concentration. There was no no significant increase in either the mean or median cell volumes for any of the carbon sources when chlorine is present. This negative impact on cell size did not translate into major changes in the parameters describing the probability distribution of biovolume, with the exception of the shape parameter for the humics-fed communities.
- 4) While a power function could adequately fit the relationship between biovolume and C_v , the slope was significantly different from -1.0 , indicating that more biomass was present in the population of cells comprising the lower end of the spectrum. Significant non-linearity for the probability distributions, in conjunction

with the steeper slope mentioned above, indicate that a model other than a power function could describe this relationship with a better degree of certainty in the extreme tails of the size distribution.

- 5) Biomass is not independent of size and smaller cells have a higher C_v content than larger cells. This also implies that the utilization of a single conversion factor for the estimation of total biomass in bacterial communities comprising organisms of varying sizes produces results that are higher or lower than the true value. This over or underestimation is dependent on the shape, location, and scale parameters of the probability distribution describing the relationship between biomass/size and abundance.
- 6) L-moment diagrams were very effective for identifying the biovolume, C_c , and C_v probability distributions of these highly skewed datasets, while at the same time being easy to interpret.

Appendix I

Moments and their Meanings

The oldest and most widely understood method for fitting frequency distributions to observed data samples is known as the method of moments (MoM). This method estimates the parameters of a distribution by equating the sample moments with theoretical moments, providing a system of nonlinear equations that can often be easily solved. Traditionally, MoM is applied by equating the theoretical product moments

mean, $\mu = E[X]$

variance, $\sigma^2 = \text{Var} [X] = E[(X - \mu)^2]$

skewness, $\gamma = E[(X - \mu)^3]/\sigma^3$

kurtosis, $\kappa = E[(X - \mu)^4]/\sigma^4$

where E is the expected value, to the sample product moments

mean,

$$m = \frac{1}{n} \sum_{i=1}^n x_{(i)}$$

variance,

$$s^2 = \left[\frac{1}{n} \sum_{i=1}^n (x_i - m)^2 \right]$$

skewness,

$$G = \frac{1}{s^3} \left[\frac{1}{n} \sum_{i=1}^n (x_i - m)^3 \right]$$

and kurtosis,

$$k = \frac{1}{s^4} \left[\frac{1}{n} \sum_{i=1}^n (x_i - m)^4 \right]$$

where μ , σ^2 , γ , and κ represent the theoretical mean, variance, skewness, and kurtosis, respectively, and m , s^2 , G , and k represent the sample mean, variance, skewness, and kurtosis, respectively. Both small ($n \leq 100$) and large sample estimates of s^2 and G exhibit extensive bias and variance (Wallis, 1974). Monte Carlo experiments by Vogel and Fennessey (1993) have shown that both G and the coefficient of variation $C_v = s/m$ exhibit even more bias for highly skewed populations ($\gamma > 2$), with both large and small sample sizes ($n < 5000$). In experiments performed to understand the distributional properties of daily streamflow sequences, they found product moments to be of little value for discriminating among potential candidate distributions. They developed a simple alternative; the method of L-moments originally introduced by Hosking (1990). Hosking (1990) showed that the method of L-moments is an exact analogue to the method of moments. However, L-moment estimators are nearly unbiased for all sample sizes and all distributions. These L-moments are analogous to ordinary moments, in that their purpose is to summarize theoretical probability distributions and observed samples. They can be used for parameter estimation, hypothesis testing, and interval estimation. Several advantages of L-moments over conventional moments include:

- (1) L-moments are able to summarize and describe a wider range of distributions (ie. those expressible in inverse cumulative distribution function (CDF) form).
- (2) Sample estimates of L-moments are more robust to the presence of outliers and suffer less from the effects of sampling variability.
- (3) L-moments are less subject to estimation bias and are nearly unbiased for all underlying distributions.
- (4) L-moments often yield more accurate estimates of the parameters of a fitted distribution than do conventional moments.
- (5) Sample L-moments are able to take on the full range of values, as do the L-moments of the underlying population.

Since L-moments are linear combinations of the ranked observations, they are subject to less bias than ordinary product moments, which require squaring, cubing, or quadrupling of the variance, skewness, or kurtosis, respectively. L-moments and probability-weighted moments are defined in the next sections.

L-Moments

The first four L-moments are defined as:

$$\lambda_1 = E[X_{1:1}]$$

$$\lambda_2 = (1/2)E[X_{2:2} - X_{1:2}]$$

$$\lambda_3 = (1/3)E[X_{3:3} - 2X_{2:3} + X_{1:2}]$$

$$\lambda_4 = (1/4)E[X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4}]$$

where $X_{i:m}$ is the i^{th} smallest variable in a sample of size m and E stands for expectation (Wang, 1996; Vogel, et al., 1993; Hosking, 1990; Hosking and Wallis, 1987; Vogel and Fennessey, 1993). When the sample consists of many values, the expectations of λ_1 and λ_2 give measures of location and scale, and

$$\tau_3 = \lambda_3/\lambda_2$$

$$\tau_4 = \lambda_4/\lambda_2$$

give measures of skewness and kurtosis, respectively. The term

$$\tau_2 = \lambda_2/\lambda_1$$

can be defined as the L-coefficient of variation.

The estimation of L-moments has traditionally been carried out through linear combinations of the probability-weighted moments (PWM).

Probability Weighted Moments (PWMs)

L-moments can be defined in terms of PWMs by the relationships

$$\lambda_1 = \beta_0$$

$$\lambda_2 = 2\beta_1 - \beta_0$$

$$\lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0$$

$$\lambda_4 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0$$

where β_r are the PWMs defined as

$$\beta_r = E[X(F_x(x))^r]$$

where $F_x(x)$ is the cumulative distribution function of X . When $r = 0$, β_0 is the mean biovolume μ as defined above. Unbiased sample estimates of the PWMs can be computed for any distribution from

$$b_0 = m = \frac{1}{n} \sum_{i=1}^n x_i$$

$$b_1 = \sum_{i=1}^{n-1} \left[\frac{(n-i)}{n(n-1)} \right] x_{(i)}$$

$$b_2 = \sum_{i=1}^{n-2} \left[\frac{(n-i)(n-i-1)}{n(n-1)(n-2)} \right] x_{(i)}$$

$$b_3 = \sum_{i=1}^{n-3} \left[\frac{(n-i)(n-i-1)(n-i-2)}{n(n-1)(n-2)(n-3)} \right] x_{(i)}$$

where $x_{(i)}$ represents the ordered biovolumes, with $x_{(1)}$ being the largest observation and $x_{(n)}$ the smallest observation (Vogel, et al., 1993).

Direct Estimators for L-Moments

By following the definition of L-moments, a more logical way of estimating them can be obtained. As an example, consider the calculation of λ_2 for a sample of finite size n . For each combination of two values from the sample, the difference between the larger and smaller value can be found. Following the definition on the previously noted for λ_2 , the average for all possible combinations of values in the sample gives the sample λ_2 times two. λ_3 can be calculated in a similar manner by considering all possible combinations of three values from the sample and using the equation shown above. λ_4 , likewise, can be calculated in a similar manner by considering all possible combinations

of four values and applying the equation for λ_4 previously noted. Direct calculations for λ_r can be performed by using the following equations and have been shown to produce the same values as the PWM method (Wang, 1996).

$$\hat{\lambda}_1 = \frac{1}{n C_1} \sum_{i=1}^n x_{(i)}$$

$$\hat{\lambda}_2 = \frac{1}{2} \frac{1}{n C_2} \sum_{i=1}^n (i-1 C_1 - n-i C_1) x_{(i)}$$

$$\hat{\lambda}_3 = \frac{1}{3} \frac{1}{n C_3} \sum_{i=1}^n (i-1 C_2 - 2^{i-1} C_1 - n-i C_1 + n-i C_2) x_{(i)}$$

$$\hat{\lambda}_4 = \frac{1}{4} \frac{1}{n C_4} \sum_{i=1}^n (i-1 C_3 - 3^{i-1} C_2 - n-i C_1 + 3^{i-1} C_1 - n-i C_2 - n-i C_3) x_{(i)}$$

where $x_{(i)}$, $i = 1, 2, \dots, n$, are sample values ranked in ascending order and

$${}^m C_k = \binom{m}{k} = \frac{m!}{k!(m-k)!}$$

is the number of combinations of any k items from m items and is equal to zero when $k > m$. From this formula it can easily be seen that as the number of observations increase, the number of linear combinations can increase dramatically (see Table I, below)

Number of Items (k)	Number of Observations (m)			
	10	50	100	150
1	10	50	100	150
2	45	1,225	4,950	11,175
3	120	19,600	161,700	551,300
4	210	230,300	3,921,225	20,260,275

Table 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).

Generalized Pareto Distribution

The three-parameter generalized Pareto distribution has the probability density function (pdf)

$$f_X(x) = \alpha^{-1} [1 - \kappa(x - \beta)/\alpha]^{(1/\kappa - 1)} \quad \kappa \neq 0$$

and the cumulative distribution function (cdf)

$$F_X(x) = 1 - [1 - \kappa(x - \beta)/\alpha]^{1/\kappa} \quad \kappa \neq 0$$

For $\kappa < 0$ biovolume is bounded below so that $\beta \leq X \leq \infty$. For $\kappa > 0$, biovolume is bounded both above and below, so that $\beta \leq X \leq (\beta + \alpha/\kappa)$. If both $\beta = 0$ and $\kappa = 0$, the GPD reduces to the exponential distribution. When $\beta = 0$ and $\kappa = 1$ the GPD reduces to the uniform distribution.

The GPD distribution parameters can be related to conventional product moments, probability weighted moments, or L-moments. If they exist, the mean, variance, and skewness are,

$$\mu = \beta + \alpha/(1+\kappa)$$

$$\sigma^2 = \alpha^2/[(1+\kappa)^2(1+2\kappa)]$$

$$\gamma = 2(1+2\kappa)^{1/2}(1-\kappa)/(1+3\kappa)$$

So that the distribution has infinite variance when $\kappa \leq -1/2$, infinite skewness when $\kappa \leq -1/3$, and infinite kurtosis when $\kappa \leq -1/4$ (Hosking and Wallis, 1987; Hosking, 1990). Hosking (Hosking, 1990) showed that the L-moments λ_r , $r = 1, 2, \dots$, of any real-valued random variable X exist if and only if X has a finite mean. This implies that a distribution may be specified by its L-moments even if some of its conventional moments do not exist, as is the case for several of the datasets collected in this study. The parameters of the distribution in terms of L-moments are

$$\kappa = (1-\tau_3)/(1+\tau_3)$$

$$\alpha = \lambda_2(1+\kappa)(2+\kappa)$$

$$\beta = \lambda_1 - \alpha/(1+\kappa)$$

where λ_r , $r = 1, 2, 3$, are the first three L-moments.

For estimation of parameters, the cdf of the GPD can be inverted to obtain the quantile function

$$X_p = \beta + (\alpha/\kappa)[1-p^\kappa]$$

where p is the exceedance probability and X_p is that value of the biovolume which is exceeded with probability p . Substitution of sample estimates of λ_r into the above equations produces L-moment estimators of the quantile function. These L-moment estimators always produce feasible values for the estimated parameters.

Derivation of the relationship between L-moments and the GPD model parameters was performed by Hosking (Hosking, 1990) such that

$$\lambda_1 = \beta + \alpha/(1 + \kappa)$$

$$\lambda_2 = \alpha/[(1 + \kappa)(2 + \kappa)]$$

$$\tau_3 = (1 - \kappa)/(3 - \kappa)$$

$$\tau_4 = (1 - \kappa)(2 - \kappa)/[(3 + \kappa)(4 + \kappa)]$$

where τ_3 and τ_4 are defined previously. Hosking also derived an approximate relationship between L-kurtosis and L-skewness, which is useful for comparing the theoretical L-moment relationships with sample estimates

$$\tau_4 = 0.20196\tau_3 + 0.95924\tau_3^2 - 0.20096\tau_3^3 + 0.05061\tau_3^4$$

With the above definitions for τ_3 and τ_4 it becomes apparent that an L-moment diagram for the GPD depends only upon the important shape parameter κ . Estimates of $\alpha(\text{hat})$ and $\kappa(\text{hat})$ from the PWMs can provide for a method of determining the variances of these parameters. Namely, the variance-covariance matrix is defined by

$$n \text{ var} \begin{bmatrix} \hat{\alpha} \\ \hat{\kappa} \end{bmatrix} \sim \frac{1}{(1 + 2\kappa)(3 + 2\kappa)} \times \begin{bmatrix} \alpha^2(7 + 18\kappa + 11\kappa^2 + 2\kappa^3) & \alpha(2 + \kappa)(2 + 6\kappa + 7\kappa^2 + 2\kappa^3) \\ \alpha(2 + \kappa)(2 + 6\kappa + 7\kappa^2 + 2\kappa^3) & (1 + \kappa)(2 + \kappa)(1 + \kappa + 2\kappa^2) \end{bmatrix}$$

Thus, by placing the measured values of α and κ into the above matrix, estimates on the variance and covariance can be obtained (Hosking and Wallis, 1987). From the variances, the standard deviations for α and κ are easily computed, where $\text{S.D.} = (\text{var}/n)^{1/2}$.

Data Tables

n	[C]	rt	α	β	k	α^*	κ^*	x^1	x^2	med	λ_2^3	σ^2	t_3	t_4	τ_3	τ_4
226	1	1	0.063	0.001	-0.115	0.007	0.061	0.073	0.072	0.047	0.038	0.007	0.382	0.271	0.313	0.151
244	1	1	0.161	0.001	-0.532	*	*	0.288	0.344	0.142	0.234	-1.825	0.660	0.546	0.621	0.454
288	1	1	0.091	0.001	-0.339	0.009	0.067	0.139	0.139	0.077	0.083	0.059	0.514	0.353	0.503	0.321
503	1	1	0.083	0.005	-0.296	0.006	0.046	0.125	0.123	0.068	0.069	0.034	0.657	0.504	0.599	0.428
518	1	1	0.082	0.001	-0.159	0.006	0.038	0.089	0.089	0.062	0.053	0.014	0.556	0.410	0.479	0.297
190	1	1	0.063	0.001	-0.115	0.007	0.061	0.073	0.072	0.047	0.038	0.007	0.705	0.497	0.705	0.559
414	1	2	0.064	0.001	-0.376	0.005	0.063	0.104	0.103	0.052	0.063	0.042	0.459	0.345	0.387	0.211
497	1	2	0.088	0.001	-0.360	0.007	0.055	0.139	0.138	0.074	0.084	0.068	0.558	0.411	0.524	0.344
484	1	2	0.108	0.001	-0.366	0.008	0.056	0.172	0.171	0.083	0.104	0.108	0.544	0.400	0.515	0.334
304	1	2	0.102	0.001	-0.173	0.009	0.050	0.125	0.125	0.080	0.068	0.023	0.531	0.435	0.408	0.229
376	1	2	0.075	0.001	-0.240	0.006	0.049	0.100	0.099	0.062	0.056	0.019	0.480	0.333	0.415	0.236
456	1	2	0.074	0.001	-0.309	0.006	0.045	0.109	0.108	0.057	0.063	0.030	0.523	0.373	0.449	0.268
988	2	1	0.155	-0.021	-0.605	*	*	0.327	0.370	0.090	0.281	-0.728	0.551	0.405	0.486	0.304
161	2	1	0.213	0	-0.318	0.028	0.086	0.313	0.313	0.178	0.186	0.269	0.575	0.439	0.519	0.338
202	2	1	0.234	0.001	-0.069	0.025	0.058	0.252	0.252	0.175	0.130	0.073	0.506	0.365	0.492	0.310
586	2	2	0.099	0.001	-0.515	*	*	0.204	0.204	0.779	0.137	-1.352	0.423	0.279	0.365	0.192
452	2	2	0.085	0.001	-0.499	0.041	0.588	0.171	0.170	0.071	0.113	11.90	0.661	0.493	0.610	0.440
500	2	2	0.029	-0.004	-0.654	*	*	0.070	0.078	0.017	0.062	-0.022	0.670	0.466	0.670	0.514

Table 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

n	[C]	rt	α	β	k	α^*	κ^*	x^1	x^2	med	λ_2^3	σ^2	t^3-s	t^4-s	τ^3-t	τ^4-t
222	1	1	0.482	-0.077	-0.301	0.053	0.070	0.580	0.613	0.252	0.406	1.196	0.482	0.199	0.482	0.300
188	1	1	0.386	-0.053	-0.153	0.044	0.063	0.394	0.403	0.180	0.247	0.299	0.405	0.131	0.405	0.227
205	1	1	0.224	0.0004	-0.132	0.024	0.060	0.258	0.258	0.165	0.138	0.091	0.395	0.212	0.395	0.218
296	1	1	0.390	0.050	-0.094	0.035	0.049	0.376	0.381	0.191	0.226	0.229	0.377	0.115	0.377	0.202
303	1	1	0.830	-0.077	-0.09	0.073	0.048	0.827	0.835	0.509	0.478	1.016	0.375	0.16	0.375	0.200
637	1	1	1.018	-0.043	-0.045	0.062	0.033	1.021	1.023	0.695	0.545	1.247	0.354	0.187	0.354	0.183
270	1	2	0.166	-0.012	-0.114	0.016	0.051	0.173	0.174	0.083	0.099	0.045	0.386	0.132	0.386	0.210
467	1	2	0.087	-0.001	-0.300	0.006	0.039	0.123	0.124	0.059	0.073	0.039	0.481	0.266	0.481	0.299
170	1	2	0.107	0.002	-0.386	0.014	0.102	0.177	0.175	0.087	0.108	0.132	0.547	0.386	0.53	0.350
190	1	2	0.105	0.002	-0.543	*	*	0.233	0.231	0.092	0.157	-0.605	0.653	0.502	0.628	0.462
438	1	2	0.467	-0.039	-0.295	0.036	0.049	0.607	0.623	0.308	0.388	1.070	0.479	0.266	0.479	0.297
201	1	2	0.229	-0.015	-0.479	0.046	0.212	0.411	0.425	0.157	0.289	4.592	0.587	0.393	0.587	0.413
391	1	2	0.196	-0.012	-0.441	0.020	0.092	0.329	0.338	0.128	0.225	1.038	0.563	0.345	0.563	0.386
324	2	1	0.391	-0.035	-0.093	0.033	0.046	0.393	0.397	0.235	0.226	0.228	0.376	0.178	0.376	0.201
360	2	1	0.117	-0.001	-0.402	0.011	0.075	0.194	0.195	0.099	0.123	0.197	0.540	0.367	0.540	0.360
371	2	1	0.262	-0.022	-0.270	0.022	0.051	0.330	0.338	0.195	0.208	0.280	0.465	0.270	0.465	0.283
1862	2	1	0.250	0.001	-0.258	0.009	0.022	0.338	0.338	0.189	0.194	0.235	0.513	0.374	0.459	0.277
658	2	1	1.671	-0.030	-0.157	0.101	0.034	1.946	1.951	1.226	1.075	5.714	0.407	0.233	0.407	0.228
629	2	1	1.647	-0.133	-0.109	0.101	0.034	1.699	1.715	1.054	0.978	4.373	0.384	0.178	0.384	0.208
364	2	2	0.145	-0.008	-0.416	0.014	0.081	0.236	0.241	0.117	0.157	0.370	0.548	0.389	0.548	0.370
399	2	2	0.078	-0.005	-0.449	0.008	0.098	0.133	0.136	0.058	0.091	0.195	0.568	0.380	0.568	0.392
377	2	2	0.195	-0.006	-0.065	0.015	0.043	0.202	0.202	0.143	0.108	0.050	0.363	0.201	0.363	0.190
1362	2	2	0.341	-0.019	-0.128	0.014	0.023	0.370	0.373	0.226	0.209	0.206	0.393	0.211	0.393	0.216
237	2	2	0.650	-0.030	-0.076	0.065	0.054	0.670	0.673	0.455	0.365	0.583	0.368	0.188	0.368	0.195
408	2	2	0.251	-0.027	-0.486	0.041	0.180	0.436	0.461	0.154	0.322	8.335	0.591	0.363	0.591	0.418
220	2	2	0.554	-0.006	-0.278	0.060	0.068	0.758	0.761	0.455	0.445	1.324	0.47	0.307	0.469	0.287

Table 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.

n	[C]	rt	α	β	k	α^*	κ^*	x^1	x^2	med	λ_2^3	σ^2	t^3-s	t^4-s	τ^3-t	τ^4-t
240	1	1	0.198	0.002	0.078	0.020	0.053	0.186	0.186	0.131	0.089	0.029	0.372	0.208	0.299	0.141
615	1	1	0.101	0.001	-0.211	0.006	0.037	0.128	0.128	0.070	0.071	0.028	0.470	0.281	0.434	0.253
287	1	1	0.166	0.009	-0.330	0.016	0.066	0.234	0.238	0.118	0.148	0.180	0.498	0.297	0.498	0.316
973	1	1	0.280	0.004	-0.018	0.014	0.026	0.289	0.286	0.187	0.144	0.084	0.457	0.320	0.341	0.173
489	1	1	0.241	0.033	0.081	0.017	0.037	0.254	0.256	0.187	0.107	0.043	0.445	0.320	0.298	0.140
500	1	1	0.242	0.012	-0.131	0.017	0.038	0.292	0.290	0.171	0.149	0.105	0.601	0.501	0.394	0.217
316	1	2	0.129	0.002	0.002	0.011	0.046	0.130	0.137	0.099	0.064	0.016	0.333	0.227	0.333	0.166
512	1	2	0.112	0.001	-0.092	0.008	0.037	0.124	0.133	0.077	0.064	0.019	0.403	0.214	0.375	0.201
423	1	2	0.110	0.002	-0.095	0.008	0.041	0.124	0.130	0.099	0.064	0.018	0.427	0.236	0.377	0.202
1019	1	2	0.141	0.001	-0.016	0.007	0.026	0.145	0.145	0.094	0.072	0.021	0.482	0.363	0.340	0.172
491	1	2	0.130	0.023	0.015	0.009	0.037	0.151	0.152	0.109	0.064	0.016	0.436	0.319	0.327	0.161
494	1	2	0.123	0.025	-0.033	0.008	0.037	0.153	0.152	0.113	0.065	0.017	0.475	0.376	0.348	0.178
647	2	1	0.280	0.004	-0.003	0.017	0.032	0.284	0.285	0.225	0.141	0.079	0.350	0.201	0.334	0.168
640	2	1	0.165	0.007	-0.458	0.015	0.085	0.318	0.312	0.136	0.197	1.111	0.582	0.379	0.574	0.398
241	2	1	0.303	0.007	0.051	0.030	0.053	0.295	0.295	0.263	0.141	0.076	0.375	0.279	0.311	0.150
1021	2	1	0.476	0.001	-0.059	0.023	0.026	0.507	0.461	0.305	0.260	0.29	0.479	0.346	0.360	0.188
495	2	1	0.211	0.043	0.004	0.014	0.037	0.253	0.253	0.176	0.105	0.044	0.460	0.327	0.332	0.165
500	2	1	0.305	0.015	-0.216	0.021	0.041	0.408	0.404	0.232	0.218	0.267	0.630	0.520	0.437	0.256
456	2	2	0.084	0.004	-0.245	0.006	0.050	0.116	0.115	0.066	0.063	0.024	0.542	0.400	0.452	0.270
229	2	2	0.090	0.003	0.057	0.009	0.054	0.087	0.088	0.057	0.041	0.006	0.480	0.34	0.308	0.148
174	2	2	0.065	0.003	-0.230	0.008	0.071	0.089	0.088	0.048	0.048	0.013	0.444	0.214	0.444	0.263
1013	2	2	0.187	0.008	-0.117	0.009	0.027	0.220	0.219	0.133	0.112	0.058	0.499	0.348	0.388	0.211
501	2	2	0.169	0.015	0.060	0.011	0.036	0.173	0.174	0.124	0.077	0.023	0.461	0.341	0.307	0.147
501	2	2	0.210	0.046	-0.235	0.015	0.042	0.334	0.327	0.196	0.156	0.143	0.568	0.409	0.447	0.265

Table 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.

References Cited

- Ahrens, M.A., and Peters, R.H. (1991). Patterns and limitations in limnoplankton size spectra. *Can. J. Fish. Aquat. Sci.*, 48:1967-1978.
- Aroner, R. (1994). WQHYDRO, Water quality/ hydrology graphics/ analysis system.
- Bjornsen, P.K. (1986). Automatic determination of bacterioplankton biomass by image analysis. *Appl. Environ. Microbiol.*, 51:1199-1204.
- Borgmann, U. (1987). Models on the slope of, and biomass flow up, the biomass size spectrum. *Can. J. Fish. Aquat. Sci.*, 44:136-140.
- Boudreau, P.R., Dickie, L.M., and Kerr, S.R. (1991). Body-size spectra of production and biomass as system-level indicators of ecological dynamics. *J. Theor. Biol.*, 152:329-339.
- Bratbak, G. (1985). Bacterial biovolume and biomass estimations. *Appl. Environ. Microbiol.*, 49:1488-1493.
- Castillo, E., and Hadi, A.S. (1997). Fitting the generalized Pareto distribution to data. *J. Amer. Stat. Assoc.*, 92:1609-1620.
- Chrost, R.J. (1991). *Microbial Enzymes in Aquatic Environments*. New York: Springer-Verlag.
- Dickie, L.M., Kerr, S.R., and Boudreau, P.R. (1987). Size-dependant processes underlying regularities in ecosystem structure. *Ecol. Monog.*, 57:233-250.
- Emerson, D., Worden, R.M., and Breznak, J.A. (1994). A diffusion gradient chamber for studying microbial behavior and separating microorganisms. *Appl. Environ. Microbiol.*, 60:1269-1278.
- Fagerbakke, K.M., Heldal, M., and Norland, S. (1996). Content of carbon, nitrogen, oxygen, sulfur and phosphorous in native aquatic and cultured bacteria. *Aquat. Microb. Ecol.*, 10:15-27.
- Felip, M., Pace, M.L., and Cole, J.J. (1996). Regulation of planktonic bacterial growth rates: The effects of temperature and resources. *Microb. Ecol.*, 31:15-28.

- Fukuda, R., Ogawa, H., Nagata, T., and Koike, I. (1998). Direct determination of carbon and nitrogen contents of natural bacterial assemblages in marine environments. *Appl. Environ. Microbiol.*, 64:3352-3358.
- Gaedke, U. (1993). Ecosystem analysis based on biomass size distributions: A case study of a plankton community in a large lake. *Limnol. Oceanogr.*, 38:112-127.
- Gasol, J.M., del Giorgio, P.A., and Duarte, C.M. (1997). Biomass distribution in marine planktonic communities. *Limnol. Oceanogr.*, 42:1353-1363.
- Gasol, J.M., Guerrero, R., and Pedros-Alio, C. (1991). Seasonal variations in size structure and procaryotic dominance in sulfurous Lake Ciso. *Limnol. Oceanogr.*, 36:860-872.
- Grossmann, S. (1994). Bacterial activity in sea ice and open water of the Weddell Sea, Antarctica: A microautoradiographic study. *Microb. Ecol.*, 28:1-18.
- Harris, G.P. (1994). Pattern, process and prediction in aquatic ecology. A limnological view of some general ecological problems. *Freshwat. Biol.*, 32:143-160.
- Hosking, J.R.M. (1990). L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *J. R. Statist. Soc. B*, B:105-124.
- Hosking, J.R.M., and Wallis, J.R. (1987). Parameter and Quantile estimation for the generalized pareto distribution. *Technometrics*, 29:339-349.
- Johnson, N.L., Kotz, S., & Balakrishnan, N. (1987). In: N.L. Johnson, S. Kotz, & N. Balakrishnan. (Eds.) *Pareto Distributions*. 2 ed, New York, USA: Wiley Interscience Publication. pp. 573-626. Wiley Series in Probability and Mathematical Statistics.
- Karner, M., and Fuhrman, J.A. (1997). Determination of active marine bacterioplankton: a comparison of universal 16S rRNA probes, autoradiography, and nucleoid staining. *Appl. Environ. Microbiol.*, 63:1208-1213.
- Kjelleberg, S. (1993). *Starvation in Bacteria*. New York: Plenum Press.
- Kroer, N. (1994). Relationships between biovolume and carbon and nitrogen content of bacterioplankton. *FEMS Microbiol. Ecol.*, 13:217-224.
- LeChevallier, M.W., Shaw, N.E., Kaplan, L.A., and Bott, T.L. (1993). Development of a rapid assimilable organic carbon method for water. *Appl. Environ. Microbiol.*, 59:1526-1531.
- Lee, S., and Fuhrman, J.A. (1987). Relationships between biovolume and biomass of naturally derived marine bacterioplankton. *Appl. Environ. Microbiol.*, 53:1298-1303.

- Lewis, D.M., and Chatwin, P.C. (1997). A three-parameter PDF for the concentration of an atmospheric pollutant. *J. Appl. Meteor.*, 36:1064-1075.
- Loferer-Kroßbacher, M., Klima, J., and Psenner, R. (1998). Determination of bacterial cell dry mass by transmission electron microscopy and densitometric image analysis. *Appl. Environ. Microbiol.*, 64:688-694.
- Meyer, J.L. (1994). The microbial loop in flowing waters. *Microb. Ecol.*, 28:195-199.
- Nagata, T. (1986). Carbon and nitrogen content of natural planktonic bacteria. *Appl. Environ. Microbiol.*, 52:28-32.
- Nagata, T., and Watanabe, Y. (1990). Carbon- and nitrogen-to-volume ratios of bacterioplankton grown under different nutritional conditions. *Appl. Environ. Microbiol.*, 56:1303-1309.
- Norland, S. (1993). The relationship between biomass and volume of bacteria. In P.F. Kemp, B.F. Sherr, E.B. Sherr, & J.J. Cole (Eds.), *Handbook of Methods in Aquatic Microbial Ecology*. (pp. 303-307). Boca Raton: Lewis Publishers.
- Norland, S., Heldal, M., and Tumyr, O. (1987). On the relation between dry matter and volume of bacteria. *Microb. Ecol.*, 13:95-101.
- Pernthaler, J., Sattler, B., Simek, K., Schwarzenbacher, A., and Psenner, R. (1996). Top-down effects on the size-biomass distribution of a freshwater bacterioplankton community. *Aquat. Microb. Ecol.*, 10:255-263.
- Psenner, R. (1990). From image analysis to chemical analysis of bacteria: A long-term study? *Limnol. Oceanogr.*, 35:234-237.
- Psenner, R., and Loferer, M. (1997). Nannobacteria: Size limits and evidence. *Science*, 272:1776-1777.
- Psenner, R., and Sommaruga, R. (1992). Are rapid changes in bacterial biomass caused by shifts from top-down or bottom-up control? *Limnol. Oceanogr.*, 37:1092-1100.
- Robertson, B.R., Button, D.K., and Koch, A.L. (1998). Determination of the biomasses of small bacteria at low concentrations in a mixture of species with forward light scatter measurements by flow cytometry. *Appl. Environ. Microbiol.*, 64:3900-3909.
- Sheldon, R.W., Prakash, A., and Sutcliffe, W.H. (1972). The size distribution of particles in the ocean. *Limnol. Oceanogr.*, 17:327-340.
- Siegele, D.A., and Kolter, R. (1992). Life after log. *J. Bacteriol.*, 174(2):345-348.

- Simon, M., and Azam, F. (1989). Protein content and protein synthesis rates of marine planktonic bacteria. *Mar. Ecol. Prog. Ser.*, 51:201-213
- Sokal, R.R., & Rohlf, F.J. (1995). *Biometry, The Principles and Practice of Statistics in Biological Research*. New York: W.H. Freeman and Company.
- Sommaruga, R. (1995). Microbial and classical food webs: A visit to a hypertrophic lake. *FEMS Microbiol. Ecol.*, 17:257-270.
- Thiebaux, M.L., and Dickie, L.M. (1992). Models of aquatic biomass size spectra and the common structure of their solutions. *J. Theor. Biol.*, 159:147-161
- Troussellier, M., Bouvy, M., Courties, C., and Dupuy, C. (1997). Variation of carbon content among bacterial species under starvation condition. *Aquat Microb Ecol*, 13:113-119.
- Tuomi, P., Fagerbakke, K.M., Bratbak, G., and Høldal, M. (1995). Nutritional enrichment of a microbial community: The effects on activity, elemental composition, community structure and virus production. *FEMS Microbiol. Ecol.*, 16:123-134.
- Van Der Kooij, D., Visser, A., and Hijnen, W.A.M. (1982). Determining the concentration of easily assimilable organic carbon in drinking water. *J. AWWA*, 74 :540-545.
- Van Der Wende, E., Characklis, W.G., and Smith, D.B. (1989). Biofilms and bacterial drinking water quality. *Wat. Res.*, 23:1313-1322.
- Verity, P.G., Robertson, C.Y., Tronzo, C.R., Andrews, M.G., Nelson, J.R., and Sieracki, M.E. (1992). Relationships between cell volume and the carbon and nitrogen content of marine photosynthetic nanoplankton. *Limnol. Oceanogr.*, 37:1434-1446.
- Vidondo, B., Prairie, Y.T., Blanco, J.M., and Duarte, C.M. (1997). Some aspects of the analysis of size spectra in aquatic ecology. *Limnol. Oceanogr.*, 42:184-192.
- Vogel, R.M., and Fennessey, N.M. (1993). L moment diagrams should replace product moment diagrams. *Water Resour. Res.*, 29:1745-1752.
- Vogel, R.M., McMahon, T.A., and Chiew, F.H.S. (1993). Floodflow frequency model selection in Australia. *J. Hydrol.*, 146:421-449.
- Wallis, J.R. (1974). Just a moment! *Water Resour. Res.*, 10:211-219.
- Wang, Q.J. (1996). Direct sample estimators of L moments. *Water Resour. Res.*, 32:3617-3619.

Zelterman, D. (1992). A statistical distribution with an unbounded hazard function and its application to a theory from demography. *Biometrics*, 48:807-818.

CHAPTER 3

RELATIONSHIPS BETWEEN TOTAL AND ACTIVE CELLS IN HETEROTROPHIC BACTERIAL POPULATIONS SUBJECTED TO CONTROLLED CARBON SOURCES, CARBON CONCENTRATIONS, AND CHLORINATION

Introduction

Bacterial activity measurements and the enumeration of the active fraction of a bacterial community are important objectives in the field of microbiology. Enumeration of the metabolically active component of microbial communities is required for the accurate estimation of productivity, biomass turnover and nutrient cycling. A number of methods have been developed for activity assays, including viable plate counts, microautoradiography, direct viable counts, and fluorescent stains (Nybroe, 1995; Poulsen, et al., 1993; Porter, et al., 1996; Kogure, et al., 1979; Peele and Colwell, 1981; Karner and Fuhrman, 1997; Meyer-Reil, 1978). However, no single technique has proven adequate for all ecosystems or organisms.

Severe underestimation of the active fraction of heterotrophic bacteria in natural samples occurs when attempts are made to culture the organisms (Peele and Colwell, 1981; Byrd, et al., 1991; Barcina, et al., 1997; Nystrom, 1998; McDougald, et al., 1998; Bloomfield, et al., 1998). The presence of viable but nonculturable bacteria (VBNC) or sub-lethal injury due to exposure to stressors exacerbates the difficulty of enumerating

active bacteria. Attempts to utilize culturable plate counts as a surrogate for the enumerating active bacteria have met with limited success (McDougald, et al., 1998; Bloomfield, et al., 1998; Amann, et al., 1995).

Direct viable counts (DVC) by the method of Kogure, et al. (1979) have been used by a number of investigators to measure the response of planktonic and biofilm bacteria in nutrient supplemented samples (Peele and Colwell, 1981; Yu, et al., 1993; Smith, et al., 1994; Singh, et al., 1990; Byrd, et al., 1991). Long incubation periods, resistance to cell division inhibitors, selective substrate responsiveness, and the difficulty in determining cell elongation in microbial communities containing diverse morphologies has hampered its use in natural ecosystems (Boon, 1992; Peele and Colwell, 1981; Rodriguez, et al., 1992).

Fluorescent stains can target basic metabolic processes common to all heterotrophic bacteria. Some of these targets include membrane potential, membrane permeability, respiration and intracellular esterases (Porter, et al., 1996; Lopez-Amoros, et al., 1995; Diaper and Edwards, 1998; Davey and Kell, 1996; Porter, et al., 1996; Deere, et al., 1995). Rhodamine 123 (Rh123) detects the presence of a membrane potential whereas the oxonols detect the absence of a membrane potential. A number of workers have indicated that the oxonols [e.g. bis-(1,3)-dibutylbarbituric acid) trimethine oxonol, DiBAC₄(3)] are more appropriate than Rh123 for the determination of viability in both Gram negative and Gram positive bacteria. This is mainly due to the need for treatment of Gram negative bacteria with Tris-HCl and EDTA (i.e. TE buffer) in order to permeabilize the outer membrane and allow entry of Rh123 (Deere, et al., 1995; Lopez-

Amoros, et al., 1995; Lopez-Amoros, et al., 1997; Ratinaud and Revidon, 1995). Other investigators have shown good correlations between culturability and viability when using either stain on a variety of organisms. Oxonols and Rh123 are inherently fluorescent and dye accumulation is dependent on the magnitude of the membrane potential. Therefore, a distribution of fluorescence intensity is observed in the cells. Other fluorescent stains such as carboxyfluorescein diacetate (CFDA) require the passive entry of a non-fluorescent substrate into the cell, cleavage by non-specific esterases and the production of a fluorescent analog of fluorescein within the cell (Diaper and Edwards, 1998; Davey and Kell, 1996; Porter, et al., 1996).

Chemchrome B is also a fluorogenic ester that has been shown superior to fluorescein derivatives for the staining of a variety of bacteria (Gram positive and negative). In a study of *Klebsiella pneumoniae* survival in sterile lake water, staining with Chemchrome B was found to correlate well with DVC and total cell counts whereas culturability decreased dramatically over an 80 day period (Diaper and Edwards, 1998). For these dyes, fluorescent cells indicate the presence of both an intact membrane and active enzymatic activity.

5-cyano-2,3-ditolyl tetrazolium chloride (CTC) is a non-fluorescent tetrazolium salt that is reduced by various dehydrogenases within a cells respiratory chain and the accumulation of a fluorescent deposit within the cells indicates active respiratory functionality. Schaule, et al. (1993) used CTC to quantify respiring planktonic and sessile bacteria in drinking water. Their estimate of 1-10% actively respiring organisms agrees well with other studies that have utilized CTC in natural waters (Rodriguez, et al., 1992;

Ullrich, et al., 1996; Karner and Fuhrman, 1997). Reduction of CTC requires active respiration in the target cells and it has been shown that this dye is detrimental to cellular integrity of the, possibly due to the accumulation of precipitated CTC crystals (Ullrich, et al., 1996). These rather low estimates for activity are in contrast to microautoradiographic methods that generally provide estimates for active cells in the range of 10 to 80% of the total population (Grossmann, 1994; Meyer-Reil, 1978; Karner and Fuhrman, 1997; Schuppler, et al., 1995; Karner and Fuhrman, 1997; Pernthaler, et al., 1996; Tabor and Neihof, 1985; Ramsay, 1974). Live/Dead BacLight (Molecular Probes, Eugene, OR) is a dual stain that contains propidium iodide (PI) and Syto-9. The live fraction of cells are stained by the membrane permeable Syto-9 while PI is impermeable to a cell with an energized membrane.

Microautoradiography (MAR), first introduced by Brock and Brock (1968a; 1968b) for studies in microbial ecology more than 30 years ago may still be the most applicable for enumerating the active microorganisms in aquatic samples. Problems such as substrate selectivity and standardization still exist, yet the method has proven to be sensitive for organisms active in the uptake of specific compounds (Knoechel and Kalff, 1976; Meyer-Reil, 1978; Paerl, 1991; Tabor and Neihof, 1985; Karner and Fuhrman, 1997; Ramsay, 1974; Nilsson and Sundback, 1996; Andreasen and Nielsen, 1997; Grossmann, 1994; Tuomi, et al., 1995; Pernthaler, et al., 1996; Fuentes, et al., 1998). These studies have investigated active organisms in samples such as activated sludge, sediments, sea ice, seawater, and freshwater.

Studies of bacterial activity in natural ecosystems can be confounded by changing variables such as temperature, pH, or carbon and energy source. These variables can change between sampling periods, resulting in the inability to precisely determine causative effects. In this study, we have used a semi-controlled experimental system to generate heterotrophic bacterial biofilm communities in a continuous flow reactor. The bacterial source is a colonized biologically activated carbon (BAC) filter that had been in place for 3 years prior to the initiation of these experiments. Organisms from the BAC filter effluent colonize the reactors. Carbon supplementation was initiated in order to study the community response to substrates typically found in aquatic environments. In separate experiments, the supplements included amino acids, carbohydrates, humics, and a mixture of the 3 substrates. Bacterial communities grown on amino acids were found to have the highest number of active cells as well as the highest fraction of active cells. Substrates were supplied at three different concentrations and chlorination of one community provided a third level of treatment that could be compared to a control community. At the two higher carbon levels during pseudo-steady state growth, activity measurements were made by MAR and a suite of fluorescent stains. Carbon concentration had a relatively small impact on the fraction of active cells and the total number of active cells. Chlorination was found to severely reduce the ability of the cells to grow on agar plates without causing a major decrease in the active fraction of cells.

This method has not been applied to the heterotrophic microorganisms present in drinking water and as far as we know, no MAR study as yet has estimated the active fraction of a bacterial community under conditions of controlled carbon limitation,

differential substrate supply and chlorination stress. Comparison between MAR and fluorescent stains for detection of the active fraction allowed us to determine that a combination of any two stains that did not include CTC was effective for enumerating the total population as determined from direct DAPI (4',6-diamidino-2-phenylindole) staining. Both total cell counts and plate counts were used as methods for normalizing and comparing the activity measurements across treatment levels.

Materials and Methods

Experimental System

See Chapter 2 for an overview of the experimental system, media, and reactor operation. For each carbon substrate (amino acids, carbohydrates, humics) utilized in this study one of the replicate experiments was used to investigate the activity of the effluent bacteria from the annular reactors. One additional experiment was performed for which the substrate was a mixture of the 3 different carbon sources. In this case the carbon source comprised a 2:1:1 equimolar (as carbon) mixture of humics + amino acids + carbohydrates respectively. Samples were collected from the reactor effluents prior to increasing the carbon concentration from 1000 ppb to 2000 ppb as well as at the end of the experiments when the reactors were receiving 2000 ppb C.

Sampling Protocol

Microautoradiography was the main activity measurement used, however several fluorescent stains that assess different physiological activities were also tested. Approximately 100-200 ml of reactor effluent was collected in fired (500°C for at least 4 hours) Erlenmeyer flasks. These were dechlorinated with 20 μ L of a filter sterilized solution of sodium thiosulphate where appropriate (chlorinated reactor only) and homogenized with a Tekmar tissue homogenizer for one minute at the maximum setting. Five mL subsamples were aliquoted into 7 mL acid cleaned and fired (as above) vials in triplicate. One 10 mL sample was collected for total cell counts and spread plating on R2A agar (Reasoner and Geldreich, 1985) in triplicate. R2A plates were incubated for 7 days at room temperature (approximately 22°C) prior to enumeration.

Fluorescent Activity Stains

Five hundred μ L subsamples were aliquoted from the aforementioned 100 mL effluent sample into a 1.5 ml microcentrifuge tube containing 500 μ L of the respective stain. See Table 3.1 for particulars of the stains and protocols used.

Stain	Final Concentration	Incubation Time	Stock Solution	Working Solution
CTC	5 mM	4 – 6 hours	fresh (2x)	water (pH 6.8)
DiBAC ₄ (3)	20 μ M	30 minutes	1 mg/mL in ethanol (-20°C)	water (pH 6.8)
Rh123	0.013 mM	30 minutes	0.5 mg/mL in PBS (pH 7.5)	TE (pH 7.5)
BacLight	*	30 minutes	*	*

Table 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.

Microautoradiography (MAR)

MAR was performed according to the technique (MARG-E) devised by Tabor and Niehof (1985). Briefly, microautoradiographic emulsion (Kodak, NTB-2) was coated onto acid-cleaned (conc. H₂SO₄ + NoChromix, (Godax Laboratories, Inc.)) microscope slides in a darkroom. Preliminary experiments were performed in order to determine the amount of radioactivity added, time of incubation with the isotope, and the length of time for the exposure of the sample to the emulsion. Three concentrations of radioactivity (0.25, 1.0, and 5.0 μ Ci/mL), two incubation periods (2 and 4 hours), and two exposure times (3 and 5 days) were tested. These preliminary experiments also indicated that 4% formalin was inadequate for the killed controls. These killed controls contained cells that were approximately ten times higher in dpm/mL (scintillation counts) compared to blank controls, and in addition showed increased background counts on the autoradiography emulsions (compared to blanks). Based on these results, 5% ice-cold

trichloroacetic acid (TCA) was used to kill the cells, prevent uptake, and avoid abiotic adsorption. This did not interfere with the detection of cells as indicated by the various controls employed, including killed samples, samples incubated without added isotope, and separate subsamples from the original sample fixed, stained, and counted separately.

Based on the above results the MAR samples were incubated for 4 hours in the presence of ^3H -leucine (5 ml total volume, $1\ \mu\text{Ci/mL} \approx 7\ \text{nM}$) and uptake was stopped by adding ice-cold TCA at a final concentration of 5%. Samples were filtered on black polycarbonate $0.2\ \mu\text{m}$ filters (Poretics), rinsed with 15 mL ($3 \times 5\ \text{mL}$) ice-cold TCA (5%) followed by 8 rinses with 5 mL phosphate buffered saline (PBS, pH 7.2) containing 0.13 g/L nonradioactive carrier leucine. Filters were air dried and applied to emulsion dipped slides (in the presence of 15W red darkroom light, Kodak), placed on a cold metal plate to solidify the emulsion and incubated for 3 days in the dark with desiccant at 9°C prior to developing in Kodak D-19 developer. Development was as per the manufacturer. The developed slides were stained in DAPI for 30 minutes, rinsed in Nanopure water for 3 minutes, dipped in 1% glycerin to facilitate filter removal from the emulsion, and allowed to dry prior to gently peeling the filter off of the emulsion.

Aliquots were also removed at various time intervals and counted by liquid scintillation (Packard, Tri-Carb 1900CA, Liquid Scintillation Analyzer) in order to follow the turnover of leucine in the samples taken for MAR. These samples were collected and processed as described above (without DAPI staining) but $0.2\ \mu\text{m}$ pore size white polycarbonate filters (Nucleopore) were used in place of the black-stained polycarbonate filters.

Epifluorescence microscopy was used to quantitate DAPI stained cells and transmitted light microscopy for exposed silver grains with an Olympus BH-2 microscope. Samples were observed at a magnification of 1000X to 1500X depending on the number of cells per field of view with a 1.3 N.A. Nikon PlanFluor objective. A charge-coupled device (Optronics, 470T) was used to record images of random fields of view for all specimens under UV and transmitted light. Superposition of the images allowed for correlating direct counts of cells with the exposed silver grains, indicating active incorporation of the radioactive leucine. Active cells were sorted in two subgroups: (1) low activity cells with one or two exposed silver grains within 1 μm of the cell and (2) high activity cells with 3 or more silver grains within 1 μm of the cell.

Statistical Analyses

Sampling was performed in triplicate for all experiments and reported values are the mean and standard deviation for active cells as determined by MAR unless otherwise indicated. Triplicate subsamples from each sample were taken for scintillation counting at time intervals such that 3 to 4 samples were taken over the course of the four hour incubation period. All samples were analyzed by ANOVA with the statistical program Minitab (ver 12.1, Minitab Inc.) at an α -level of 0.05. Factors in the ANOVA were the treatments (see Chapter 1, Figure 1.1), which were broken down into three levels: (1) substrate type, (2) carbon concentration, and (3) reactor. The "substrates" used for the ANOVA were (1) amino acids, (2) carbohydrates, and (3) mixture. The "carbon concentrations" were (1) 1000-ppb carbon and (2) 2000 pp2 carbon as the specific

substrate type. The base level treatment was "reactor" and could be split into two treatments: (1) unchlorinated (control) and (2) chlorinated (test). With this experimental protocol, the carbon concentration could be nested within the substrates and the reactors could be nested within the substrates and carbon concentrations. Given that some of the measured parameters were balanced while others were unbalanced due to the different n-values for different techniques, a general linear model (GLM) ANOVA was used to analyze all data. All variables were checked to ensure they conformed to the inherent assumptions for an ANOVA and, if not, appropriate transformations were performed prior to analysis. Reported values are the back-transformed means and 95% confidence intervals. Turnover rates were analyzed by analysis of covariance (ANCOVA) and the slopes were compared with time as the covariate.

Activity Determination

One common problem with microautoradiography is the determination of active cells when a background level of exposed silver grains will always occur due to abiotic causes. Some investigators will only score cells as active when they are associated with multiple exposed silver grains, more than likely erring on the conservative side and thus underestimating the active proportion of the population. Others will score all cells as active if they are associated with an exposed silver grain. Depending on background exposed silver grain density and the area determined to be associated with a cell this could produce a very liberal estimate of activity with no basis in quantitative observations. Thus, we looked for a statistical basis for scoring active cells. In our case

cells were scored as active if they were associated with 1 or 2 (low activity) or ≥ 3 (high activity) developed silver grains within 1 μm distance from a cell. One can treat the two images (UV for DAPI stained cells and transmitted light for silver grains) obtained from a sample as independent, randomly (normally) distributed samples in the same field of view. In the case of low activity cells, where the problem of deciding if they are indeed active is the greatest, they (i.e. the two images) can be treated as a binomial sample. In this case, one can define the density of background grain formation as

$$\lambda = \text{density} = (\# \text{ of grains})/\mu\text{m}^2$$

based on the killed sample observations. The area associated with a cell can be defined as

$$A = \text{area associated with a cell} = \pi (r)^2$$

where r is the radius of the defined area, which is 1 μm .

N = # of cells with 0, 1, or 2 grains associated.

X = # of cells with 1 or 2 grains associated (random variable)

As a binomial sample the expected value for X is

$$E(X) = \frac{N \bullet e^{-\lambda A} \left[\lambda A + \frac{(\lambda A)^2}{2} \right]}{e^{-\lambda A} \left[1 + \lambda A + \frac{(\lambda A)^2}{2} \right]}$$

and the realized (observed) value for X is X_R , which is the value enumerated from the samples. From this, the best estimate of the number of cells (low activity) with uptake and incorporation becomes

$$X_R - E(X)$$

Thus, the silver grains formed in the killed samples act as a background control. This correction would only apply in cases where: (1) a very heavy background grain count was observed in killed samples, (2) the area associated with a cell was excessively large, or (3) the number of cells per field of view was high. Since our background never exceeded 7.5 grains per $1000 \mu\text{m}^2$ and averaged $4.4 (\pm 2.1)$ grains per $1000 \mu\text{m}^2$ the correction was deemed negligible. Therefore, as long as the area associated with a cell is not too large, then $E(X)$ is sufficiently small (rarely exceeded 0.5 per field of view in our samples) and can be neglected. In this study the value of $X_R - E(X)$ was $<4\%$ of the low activity cells ($<1\text{-}2\%$ of total active cells) in the sample and therefore the correction for mistakenly identifying a cell as active when it was not would be negligible. This method of analysis has the potential of adding a level of standardization to the technique of MAR. It would also add a statistical/quantitative basis to MAR that is currently lacking, removing some of the subjectivity and resulting in a higher level of confidence in the results.

Results

Measurements of potential microbial activity in aquatic habitats are normally limited to the planktonic fraction due to the difficulty in measuring surface associated activity. The annular reactors in these experiments were sampled during steady-state

reactor operation (2.1 hour detention time) and the bacteria in the effluent consist of two populations. One population results from the organisms that are in the dilution water from the GAC/BAC filter (reactor influent). The second population of bacteria result from biofilm growth and detachment in the experimental system. In these experiments we studied the recently (i.e. within 2.1 hours) detached planktonic fraction of a bacterial biofilm community in a model biofilm reactor. Activity was measured by viability (plate counts) and MAR while potential activity was estimated using five different fluorescent stains. Total cell counts were used as a basis for comparing the results of the different methods, in addition to controls for the MAR results. The turnover rates, total cells/mL, dpm/cell, cfu/mL, and active cells/mL were all $\log_{10}$ transformed and ANOVAs were performed to determine the response of the cells to chlorination, substrate type, and concentration. All measurements were significantly different for the chlorinated vs. control reactors ($p = 0.05$, < 0.0005 , < 0.0005 , < 0.0005 , and 0.019 respectively). Appendix II (Tables 3.4 and 3.5) contains a summary of all results from the control and chlorinated treatments respectively.

Figure 3.1 shows the results for dpm/cell and total cells/mL. The results indicated that all factors (substrate, carbon level, and reactor) were significantly different ($p < 0.0005$) when a fully nested, GLM ANOVA was carried out for these parameters. Based on these results the dataset was broken down into individual substrates and reanalyzed according to carbon levels and reactors.

This analysis indicated that for the substrate mixture, total cells/mL were significantly different for the two carbon feed levels ($p < 0.0005$) for both the chlorinated

and control reactors (see Figure 3.1). The mixture was the only substrate to produce results that were not significantly different between the two reactors at 2000 ppb C ($p = 0.089$) with respect to total cell counts. Increasing the carbon loading produced no change in total cell counts for either reactor when the substrate was amino acids whereas the carbohydrate chlorinated reactor had significantly less total cells/mL at 2000 ppb C ($p < 0.0005$). No significant change was detected in cell counts for the carbohydrate control reactor at the two carbon feed levels ($p = 0.069$).

The carbohydrate control reactor shows the only increase in the dpm/cell with increased carbon loading ($p = 0.016$) whereas there was no significant change for the chlorinated reactor ($p = 0.41$). Chlorination in the presence of amino acids produced the only significant increase in the dpm/cell ($p < 0.0005$) when the carbon loading was increased. In the unchlorinated reactors dpm/cell were highest for the amino acids and similar for the other substrates, except the humics (at 2000 ppb C) which were the lowest by an order of magnitude. Figure 3.2 shows the results of an ANOVA comparing the control and chlorinated reactors. Chlorination was the most significant treatment for turnover rates, total cells, active cells and culturable cells and

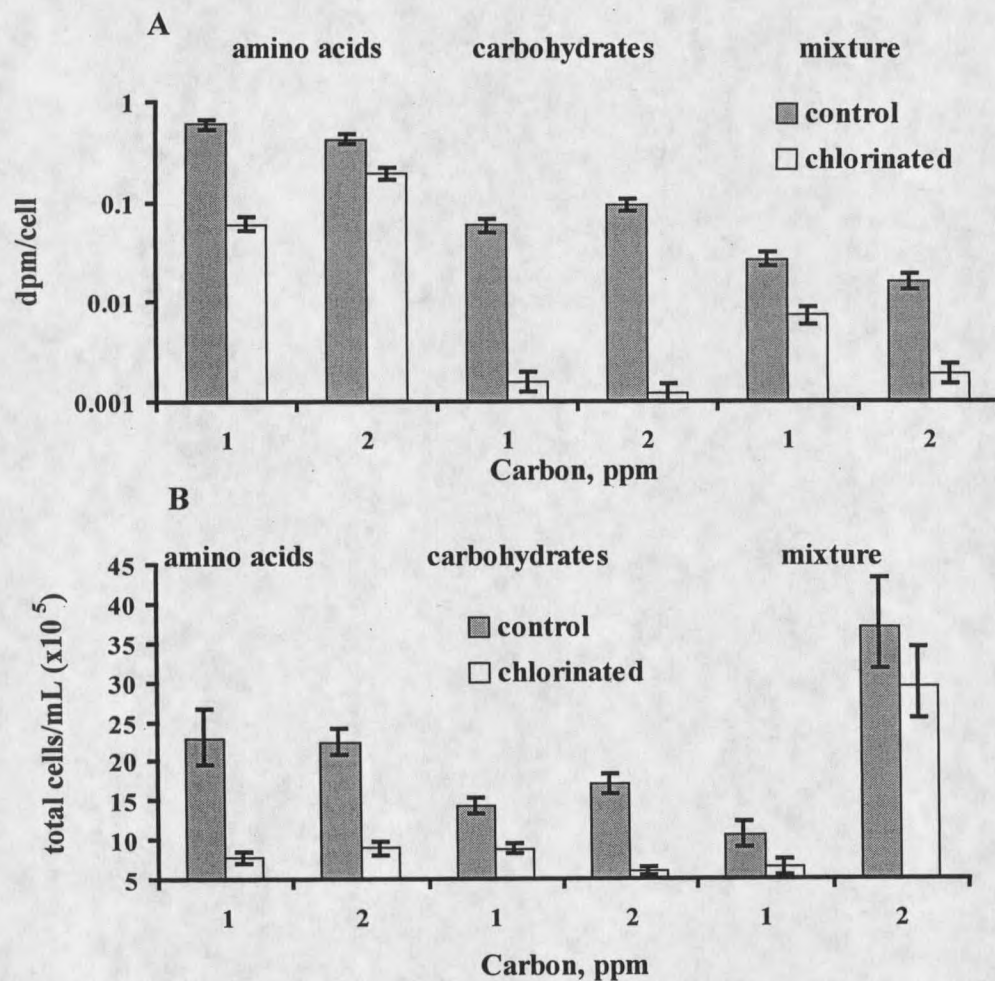


Figure 3.1 – Panel A, radioactivity counts per cell and Panel B, total cells per mL according to reactor, substrate, and carbon concentration. Error bars are 95% confidence intervals and plotted values are the back-transformed geometric means.

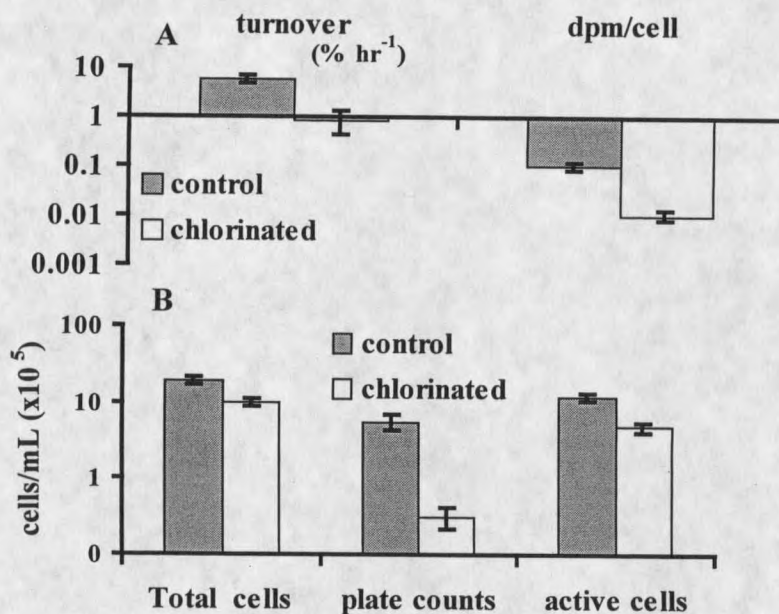


Figure 3.2 – Panel A shows the turnover rates and dpm/cell for the chlorinated and control reactors. Turnover rates and dpm/cell were significantly different ($p = 0.006$ and < 0.0005 , respectively) between control and chlorine. Panel B shows the total cells, plate counts and active cells for the chlorinated 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.

³[H]-Leucine uptake was linear for all samples over the labeling period, indicating that significant changes in the composition of the community in the subsamples did not occur over the incubation period. Figure 3.3 shows an example of the turnover curves for cells from the effluent of the carbohydrates experiment. Panel A shows the maximum leucine uptake normalized to the total cells in the sample (i.e. dpm/cell) and the turnover rate (%hr⁻¹) for the control reactor. Panel A shows the results for cells from the effluent of the chlorinated reactor and Panel B shows results for the control reactor. Table 3.2 contains the turnover rates and cfu/ml according to the reactors, substrates, and carbon

levels used in the experiments. Two separate MAR samples were taken from the amino acids control reactor at 1000 ppb C while 3 independent samples were taken at 2000 ppb C. The turnover rates were 0.163 and 0.129 hr^{-1} at 1000 ppb C and an ANCOVA indicated no difference in the slopes ($p = 0.286$). At 2000 ppb C the turnover rates were 0.251, 0.258, and 0.229 hr^{-1} and an ANCOVA indicated no difference in the slopes ($p = 0.528$). Therefore, Table 3.2 shows the average of these individual measurements. The average turnover rate for the control reactor was 14.5 ($\pm 0.7\%$, S.D.) and 25 ($\pm 1.0\%$) at 1000 and 2000 ppb C respectively. Both reactors showed an increase in the turnover rate upon increased carbon loading for all carbon sources. This increase was significant between the two carbon levels for the amino acids only ($p < 0.0005$ for both reactors). This is in contrast to the carbohydrates and mixture reactors where the turnover rates within reactors between carbon levels did not change significantly. While substrate concentration had a modest affect on the turnover rates, chlorination decreased the rates by 1 to 3 orders of magnitude depending on the substrate. The native influent cells arising from the BAC column and used for colonization of the reactors were measured on 2 separate occasions and turnover rates were low ($0.09 \pm 0.01\%$) as were dpm/cell (0.044 ± 0.02). These values, in addition to the active fraction determined by MAR ($46 \pm 6.8\%$) were comparable to the chlorinated carbohydrates, humics reactors, and mixture reactors.

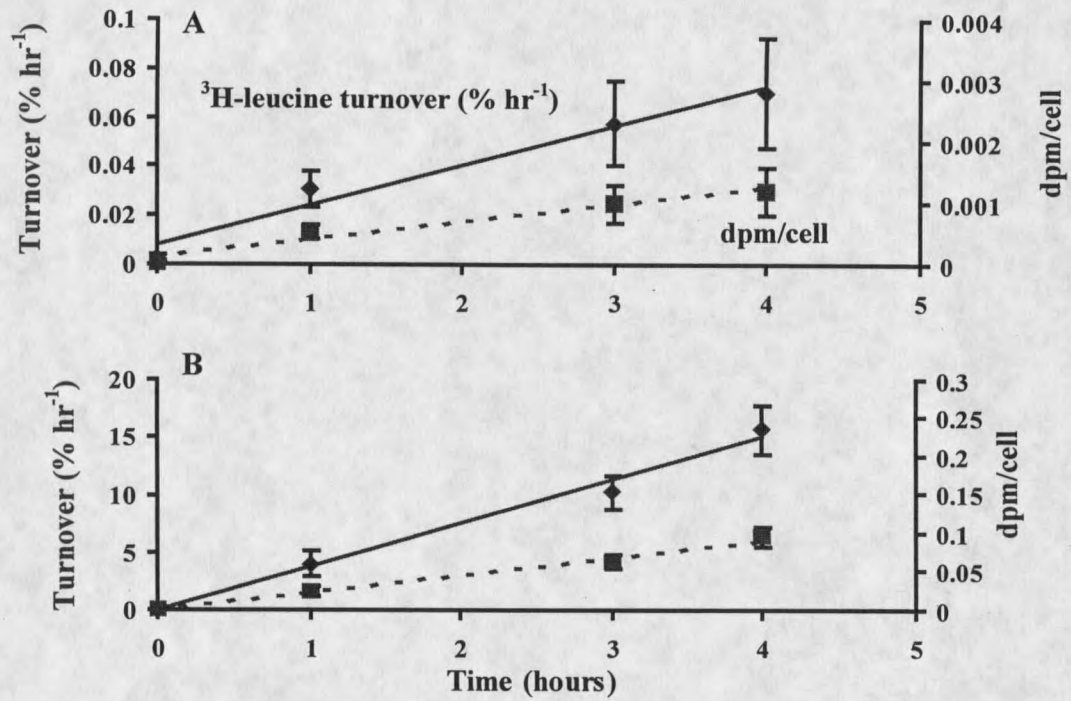


Figure 3.3 – An example of the linear uptake of tritiated leucine in cells from the effluent of the carbohydrates-fed reactors at 2000 ppb carbon. Solid lines and diamonds refer to the percentage turnover (hr⁻¹) of the labeled leucine while the dashed line and squares refer to the normalized (dpm/cell) amount of label incorporated. Error bars are standard deviations. Panel A is for the chlorinated reactor while Panel B is for the control reactor.

	1000 ppb Carbon				2000 ppb Carbon			
	Control		Chlorinated		Control		Chlorinated	
	rate ¹	cfu ²	rate ¹	cfu ²	rate ¹	cfu ²	rate ¹	cfu ²
aa ³	14.4 (0.7, 0.94)	8.7 (2.3)	1.58 (0.02, 0.999)	0.9 (0.09)	25.0 (1, 0.93)	8.1 (2.0)	5.5 (0.2, 0.98)	8.2 (0.6)
cho ⁴	1.4 (0.1, 0.90)	3.75 (0.9)	0.025 (0.004, 0.77)	0.06 (0.01)	3.8 (0.2, 0.96)	1.82 (0.09)	0.017 (0.002, 0.85)	0.09 (0.02)
mix ⁵	1.19 (0.03, 0.996)	2.0 (0.4)	0.12 (0.01, 0.94)	0.12 (0.01)	1.7 (0.1, 0.95)	7.5 (1.1)	0.17 (0.02, 0.85)	0.5 (0.1)
hu ⁶	nd	nd	nd	nd	1.7 (0.02, 0.998)	3.6 (0.05)	0.006 (0.001, 0.64)	0.02 (0.004)

Table 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.

Active cells/mL were not significantly different for the carbohydrates between carbon levels, and the distribution of cells between high and low activity did not change when the carbon loading was increased (see Figure 3.4, Panels A and B). The amino acids showed a significant difference in active cells between carbon levels for both reactors ($p < 0.002$ for both) while the numbers of high activity cells decreased significantly ($p = 0.016$) for the control reactor at the higher carbon feed level. Activity enumeration by MAR in the mixture reactor was significantly different for the reactors at the different carbon feed levels ($p < 0.008$ for both) and the distribution of both the high and low activity cells changed significantly in both reactors ($p < 0.001$ for both reactors and activity distribution). As a fraction of the total population, active cells were highest in the amino acids experiment and lowest in the carbohydrates as shown in Figure 3.5.

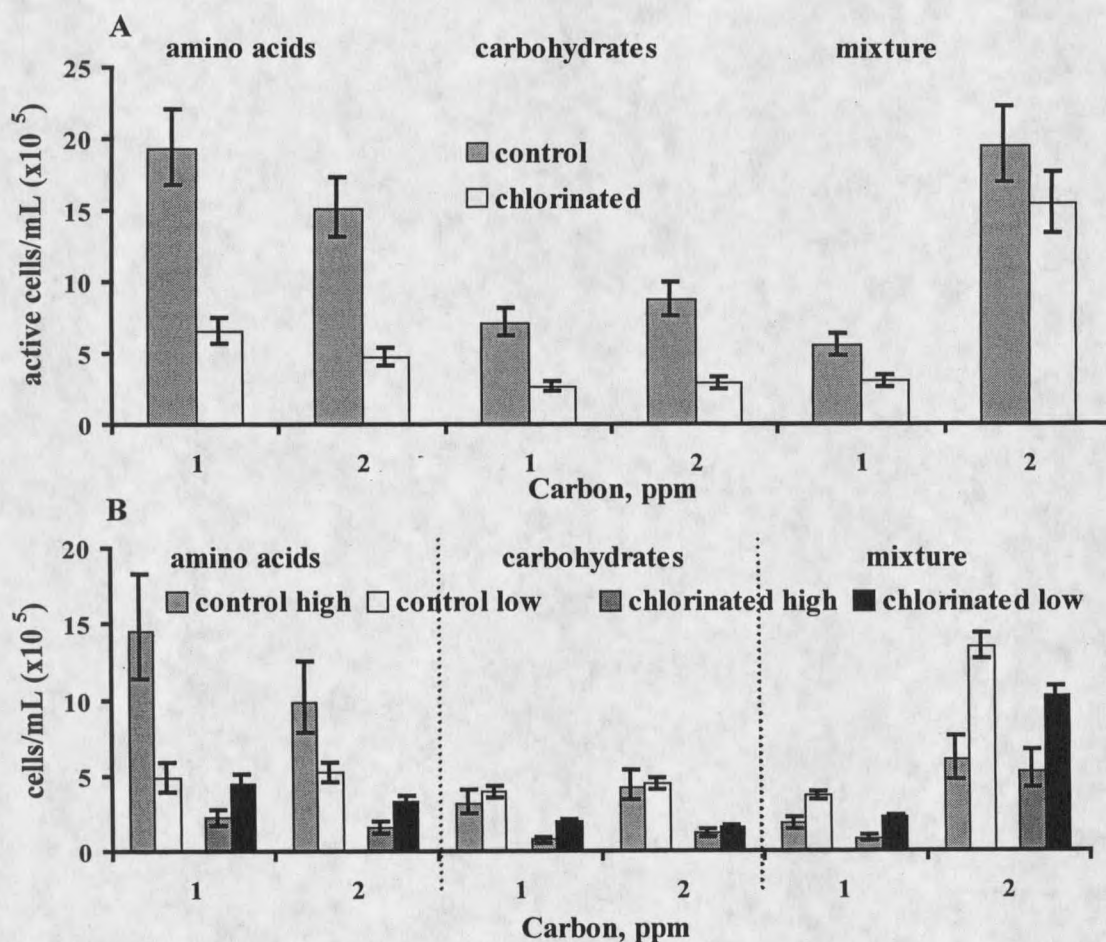


Figure 3.4- Panel A – active cells/mL in reactor effluents. Panel B shows the distribution of activity in the high and low activity classes. Results are back-transformed geometric means resulting from an ANOVA. Error bars are 95% confidence intervals.

Data analysis after normalization to total cell counts showed that the high and low activity fractions, had different distributions for all substrates (see Figure 3.5). The amino acids-fed control reactor was the only treatment that produced more cells in the highly

active fraction. The high activity fraction included $72 \pm 2\%$ more of the population than the low activity fraction for the amino acid grown communities. This can be compared to the other two substrates where activity was approximately evenly divided between high and low fractions (44 and 51% for the carbohydrates and mixture respectively). The active fraction of cells for the amino acids fed communities decreased by 20% with the increase in the carbon feed concentration whereas the carbohydrates increased by 10% and the mixture showed an insignificant increase of 3%. The fraction of active cells in both categories did not change significantly for the substrate mixture within any of the treatments (carbon level and chlorination). Overall, under chlorination, the fraction of active cells scored as low activity were always higher than the fraction scored as highly active. This trend suggests that chlorination depresses bacterial activity as measured by tritiated leucine uptake and is supported by the turnover rates (Table 3.2 and Figure 3.1) and the results from the amount of radioactivity incorporated per cell (i.e. dpm/cell, Appendix II, Tables 3.4 and 3.5)

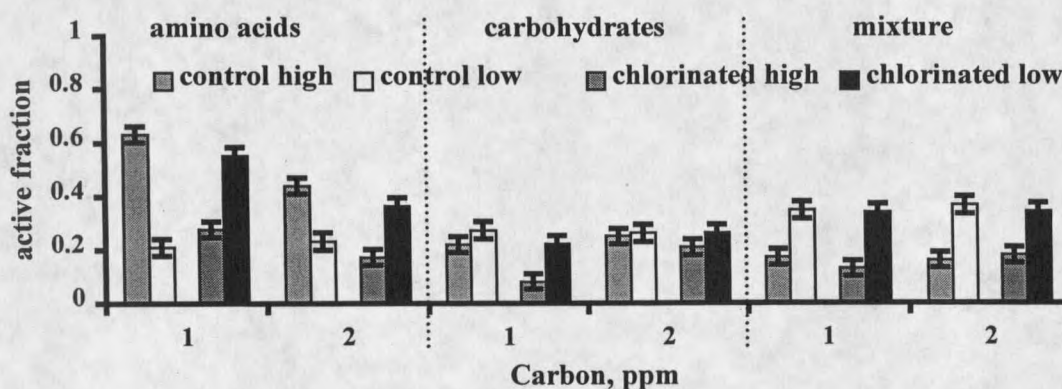


Figure 3.5 – Activity distribution normalized to total cell counts and separated into high and low fractions according to treatments.

The ANOVA results presented above were supported by a correlation analysis, the results for which are reported in Table 3.3. For this analysis, active cells/mL in the high and low fraction were summed for the total number of active cells/mL. No correlations were detected between carbon concentration and any of the measured parameters. Significant correlations were detected between substrate type (and reactors) and the measured parameters. These correlations as well as the ANOVAs indicated that all of the measured parameters are lower for the chlorinated reactor.

	substrate	reactor	turnover (% $\cdot$ hr $^{-1}$)	Total cells/mL	dpm/cell	cfu/mL
turnover (%, hr $^{-1}$)	-0.62 (0.003)	-0.583 (0.006)	ns			
Total cells/mL	-0.121 (0.6)	-0.514 (0.017)	0.588 (0.005)			
dpm/cell	-0.685 (0.001)	-0.471 (0.031)	0.91 (<0.0005)	0.419 (0.059)		
cfu/mL	ns	-0.752 (<0.0005)	0.752 (<0.0005)	0.711 (<0.0005)	0.561 (0.008)	
active cells/mL	ns	-0.592 (0.005)	0.634 (0.002)	0.959 (<0.0005)	0.344 (0.077)	0.784 (<0.0005)

Table 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.

In order to compare the various fluorescent stains with MAR, all of the active cell counts were normalized to total cell counts determined by DAPI staining. With this normalization a comparative analysis between the stains that indicate activity and MAR could be performed in order to assess the potential of the different stains with respect to MAR. Due to the sensitivity of radioisotopic methods, MAR was assumed to be the most

sensitive indicator of active cells as compared to fluorescence. Since the cells must incorporate the $^3\text{[H]}$ -leucine into protein during the incubation period, this demonstrates active cells and not potentially active cells.

Since two of the fluorescent stains used in this study enumerate the inactive cells, conversion to active cells could be made by subtracting these counts from the total cell counts. This allowed us to compare all of the stains with respect to their ability to enumerate the active population of cells in the samples. Conversion of the cell counts for the "dead" activity stain DiBAC₄(3) and the BacLight "dead" activity stain PI into active cells by subtracting the "dead" counts from the total counts allows for a comparison between all of the activity stains used in these experiments. All data for the various stains normalized as described above are shown in Figure 3.6. A significant difference was detected between the stains ($p = 0.003$) and this difference was entirely due to CTC underestimating the number of active cells as compared to all of the other stains. Removal of CTC from the analysis indicated that the other methods produced equivalent results ($p = 0.589$). Therefore, enumeration of active bacterial cells was equivalent between MAR and the fluorescent stains. Utilization of these fluorescent stains is much simpler and faster and the results are obtained quickly compared to the more difficult and tedious technique of MAR.

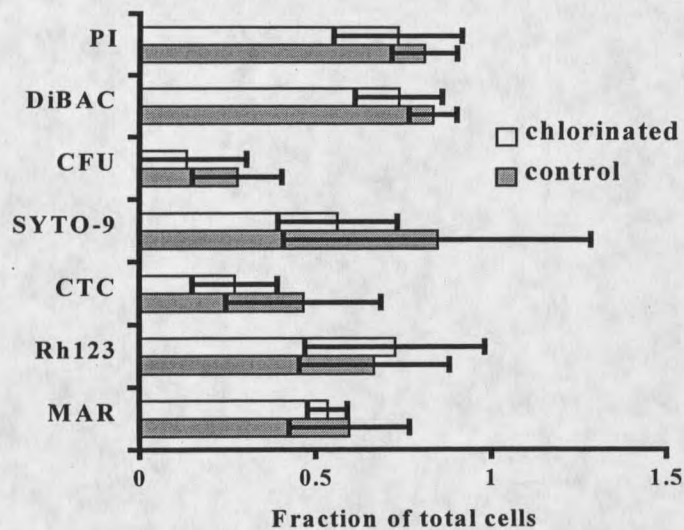


Figure 3.6 – Activity measurements by the various methods, normalized to the total cell counts as determined by DAPI staining. PI and DiBAC₄(3) counts were converted to active cells by subtraction from the total cells prior to normalization. Error bars are standard deviations.

Various combinations of the activity detection methods can be summed to get an estimate of the best combination of “live + dead” detection methods used in these experiments with respect to the total cell counts (see Figure 3.7). Based on these results, any combination of fluorescent stains, that does not include CTC, produces similar values. The four combinations including – BacLight (PI + SYTO-9), Rh123 + PI, SYTO-9 + DiBAC₄(3), and Rh123 + DiBAC₄(3)- accounted for 94, 93, 92, and 91% of the total cell counts respectively. This is in contrast to CTC + DiBAC₄(3) and CTC + PI which only accounted for 58 and 60% of the total population enumerated by DAPI direct counts, respectively.

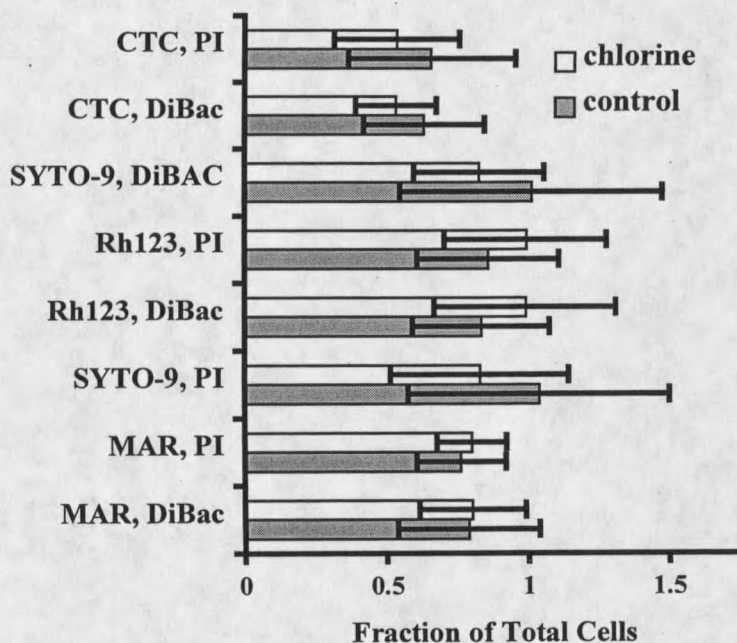


Figure 3.7 – Activity stains as a function of the various combinations of live plus dead components of the individual methods. Error bars are standard deviations.

Discussion

Statistical Models

The overall treatment that had the greatest effect on the measurements of total, culturable, and active cells/mL in addition to the leucine turnover rates and dpm/cell was chlorination, followed by substrate type and carbon concentration respectively. This can be seen readily in Figures 3.1 to 3.4. The individual variance components resulting from the ANOVAs can provide more information on treatment effects. The treatment effect of chlorination accounted for 93, 57, and 49%, of the total variance for the total cells, dpm/cell, and active cells respectively. The unexplained error components were 1.5, 0.8,

and 1.9%, for each of the above variables respectively. This indicates good reproducibility between subsamples for MAR. With this analysis, a model can be compiled with 3 nested factors (see Appendix II). The model was inadequate to explain the culturable cell counts results since the carbon levels within substrates were not significantly different. In this case, a simpler design with reactor nested within substrates was sufficient to model the data and the reactors accounted for 62% of the total variation with the unexplained error component comprising 22% of the total. The remaining variance (16%) can be attributed to the substrates indicating that substrates had a comparatively small impact on the viability of the bacteria with respect to culturability.

Activity Measurements

Activity measurements can be made by a variety of methods that target different cellular physiologies. Some of these targets include membrane potential (Rh123, oxonols), respiration (CTC and INT), isotope incorporation (labeled leucine, thymidine glucose, etc.), ATP, and intracellular esterases. While the use of MAR suffers from problems such as standardization and difficulty, it allows for the measurement of turnover rates and differentiation of subpopulations rather easily. Through a statistical analysis of the MAR results, as described in the Materials and Methods section, I think that less subjectivity in the interpretation and reporting of MAR results can be achieved. This could allow for a more appropriate basis of comparison between studies using this technique across different environmental conditions. In this study, MAR was the main technique, however, dyes that target membrane potential [Rh123 and DiBAC₄(3)], permeability (BacLight), and respiration (CTC) were also used. No nutrient

supplementation was used for any of the assays (except the leucine for MAR) since we were interested in the activity of the organisms in the reactor.

Aquatic environments have been reported to have concentrations of amino acids in the range of 0.01 to 5 μM (Volk, et al., 1997; Nilsson and Sundback, 1996; Jorgensen, et al., 1998). These values are contributions from the dissolved free amino acids (DFAA), dissolved combined amino acids (DCAA) and humic bound amino acids (HBAA). Thus in this study the addition of labeled leucine at a final concentration of 7 nM can be considered a trace supplement and would not likely affect a change in community composition or activity over the relatively short incubation period of 4 hours. As indicated by the linear uptake of the ^3H -leucine over the incubation period, there was no change in the manner in which isotope was taken up over the course of the incubation period. The organisms in the effluent of the reactor can be considered recently detached surface associated bacteria although there is a constant influx of bacteria from the BAC column that constitutes the dilution water. These influent organisms generally comprised 20% (range 4-44%) of the effluent cells as total cell counts. The high value of 44% arises due to the effect of chlorine on the carbohydrate fed community. Chlorination had a much larger impact on the number of cells in the effluent of the carbohydrates grown bacteria than for the other substrates (see Figure 3.1).

Data for the active fraction of cells and the dpm/cell agree quite well with other published information for aquatic ecosystems where ^3H -leucine (or amino acid mixtures) was used as the radioactive tracer (Karner and Fuhrman, 1997; Tabor and Neihof, 1985; Kirchman, et al., 1985). Although turnover rates and dpm/cell were significantly less for

the carbohydrates, mixture, and humics than for the amino acids, the fraction of the population that was active did not change as dramatically when compared to dpm/cell and turnover rates (Figures 3.1 panel A, 3.4 panel A and Table 3.2). This is also apparent for the dilution water from the BAC filter where low uptake rates and dpm/cell were measured, yet approximately 50% of the cells were active by MAR. For the humics reactors at 2000 ppb carbon, the difference in both dpm/cell (0.1 vs. 0.001 dpm/cell, control vs. chlorinated) and turnover rate was 2 to 3 orders of magnitude, the largest difference detected, indicating that chlorination had a highly significant impact. There was no significant correlation ($r = 0.344$, $p = 0.077$) between the fraction of MAR active cells and the dpm/cell, indicating that the methodology employed was sensitive to the lowest per-cell activity levels in the samples (Grossmann, 1994). This implies that the lowest counts of active bacteria (carbohydrate fed populations) did not result from a substantial number of organisms that incorporated isotope in amounts below the threshold for autoradiographic detection. This is also supported by the results from the other activity stains, excepting CTC, which consistently underestimated the active fraction. Presumably, this is due to the deleterious effects that CTC has on cellular respiration (Ullrich, et al., 1996; Rodriguez, et al., 1992; Smith, et al., 1994).

Karner and Fuhrman (1997) and Ullrich, et al. (1996) compared MAR with CTC activity counts in marine and freshwater samples. Their conclusion that CTC regularly underestimates potential activity was confirmed in our study. While Karner and Fuhrmann (1997) measured CTC reduction from 0 to 4.6 % of the MAR fraction in marine waters, Ullrich, et al. (1996) determined that 2.6 to 9.6 times as many cells were

detected by MAR as compared to CTC in freshwater. Both papers, as well as several others indicate that the fraction of the population detectable by CTC reduction is on the order of 1 to 30%, although higher values have been measured (Kalmbach, et al., 1997; Schaule, et al., 1993; Rodriguez, et al., 1992; Choi, et al., 1996). In our experiments, 47% of the control communities and 27% of the total cells in the chlorinated communities were actively reducing CTC. All carbon sources consistently produced a significantly higher number of active cells in the control vs. chlorinated reactors. The control amino acids were the only substrate that produced a significantly higher number of active cells in the high activity subpopulation.

In contrast to CTC active cells, MAR activity counts are generally in the range of 10 to 60%, with exceptions at either end of the spectrum. These exceptions can have different meanings, including: (1) genuinely low or high activity, (2) underestimation due to only scoring cells with multiple silver grains, (3) cells that do not incorporate the particular isotope used, and (4) populations that disproportionately incorporate the isotopic tracer used. In an examination of bacterial communities in sea ice, Grossman (1994) measured activities as low as 3 to 5%. Karner and Fuhrman (1997) measured activities from 33 to 76% by MAR in marine waters. Ullrich, et al. (1996) detected activities from approximately 35 to 80% (estimated from Fig. 1 in (1996)), in fresh and brackish waters. Meyer-Reil (1978) measured activities by MAR in the Kiel Fjord, Germany at 4.5 to 56.2% which compares well with the measurements of Ullrich, et al. (1996), in the same Fjord (38 to 78%, estimated from Fig. 1b in (Ullrich, et al., 1996)) with the exception of the one low value by Meyer-Reil. In samples from Chesapeake

Bay, Tabor and Niehof (1985) determined that 50% of the cells were active in amino acid uptake by MAR.

Interestingly, the percentage of the total cells detected with any combination of the putative "live" and "dead" stains was fairly consistent when normalized to the total cell counts. Stains that are intended to include only the active fraction include Rh123 and SYTO-9 (BacLight active component) whereas the inactive fraction could be enumerated with PI (BacLight inactive component) or DiBAC₄(3). As such, four combinations can be generated to estimate the total population of cells. These combinations are BacLight (PI + SYTO-9), Rh123 + PI, SYTO-9 + DiBAC₄(3), and Rh123 + DiBAC₄(3). The various combinations, when divided by the total cell counts accounted for 94, 93, 92, and 91% of the total cell counts respectively. This is in contrast to CTC + DiBAC₄(3) and CTC + PI which only accounted for 58 and 60% of the total cell counts, respectively.

Plate counts were significantly higher in the control reactor and averaged approximately 30% of total cell counts whereas chlorination reduced culturability to 10% of total cell counts. These percentages are generally higher than values reported for other aquatic environments which range from approximately 0.01-10%. This could be due to the fact that R2A agar was developed for drinking water, a seven day incubation at room temperature, or a combination of these factors (Reasoner and Geldreich, 1985). The culturability of the organisms entering the reactors was $20 \pm 12\%$ which was not significantly different than culturability in the control effluent for all substrates. Conversely, substrate type had a major impact on the culturability of the organisms in the chlorinated reactors. After chlorination, culturability drops to 3% when the amino acids

are not included in the analysis. From Table 3.2, it can be seen that chlorination did not have as large an effect on the culturability of the organisms grown with amino acids as compared to the other substrates.

Control vs. chlorinated culturable cells for the amino acids were 55 and 59% of the active fraction, respectively. For the carbohydrates, culturable cells were 25 and 3% of the active fraction, while for the mixture 70 and 5% of the fraction was active, respectively. MAR activity measurements indicated that approximately 35 times as many cells were active in the chlorinated reactor effluents than grew on the R2A plates, and for the control reactor this decreases to about 2 times. For the amino acids, active cells were approximately 3 times and 2 times the culturable cell counts for the chlorinated and control reactors respectively. For the mixture experiment, active cells vs. culturable cells were 19 times and 1.5 times for the chlorinated and control reactors respectively. The carbohydrates produced the second highest difference in activity and culturability between the chlorinated (58 times) and control treatments (4 times). Data for the humics was only available at the higher carbon concentration but indicated that chlorination affected culturability to the largest degree. Active cells in the chlorinated reactor were 122 times greater than culturable cells while the control reactor produced results similar to the other substrates at 1.5 times as many active cells compared to culturable cells.

All of the data indicate that while chlorination can severely inhibit growth on agar plates, the carbon source utilized influences culturability. While activity measurements (by MAR or fluorescent stains) are significantly affected by chlorination, a much larger

fraction of the community is enumerated by activity as compared to plate counts. In either case substrate concentration played a relatively minor and insignificant role.

Conclusions

- 1) Chlorination had the most significant impact on the activity measurements.
- 2) Substrate type had a larger impact on all measured parameters (turnover rates, dpm/cell, active cells/mL, total cells/mL, and cfu/mL) than substrate concentration. This is especially apparent when evaluating turnover rates (Table 3.2 and Table 3.3) and dpm/cell (Figure 3.1, A), parameters that cannot be measured by the use of a fluorescent activity stain.
- 3) Amino acids produced the most active populations. Proportional activities (i.e. active cells normalized to total cell counts) in the chlorinated amino acids reactor exceeded or equaled the activities in the control reactors for all other substrates tested.
- 4) Neither chlorination nor substrate concentration affected the proportion of active cells when the substrate was a mixture of amino acids, carbohydrates, and humics (average $51 \pm 7\%$, $n = 12$).
- 5) Any combination of live/dead fluorescent stains that does not include CTC produces cell counts similar to the total cell counts determined by DAPI staining of bacterial communities in this study.

- 6) Culturable cells were effective in estimating the active population under certain conditions (e.g. control amino acids and mixture grown communities) but severely underestimated the active component in the presence of chlorine.

References Cited

- Amann, R., Ludwig, W., and Schleifer, K.-H. (1995). Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol. Rev.*, 59:143-169.
- Andreasen, K., and Nielsen, P.H. (1997). Application of microautoradiography to the study of substrate uptake by filamentous microorganisms in activated sludge. *Appl. Environ. Microbiol.*, 63:3662-3668.
- Barcina, I., Lebaron, P., and Vives-Rego, J. (1997). Survival of allochthonous bacteria in aquatic systems: a biological approach. *FEMS Microbiol. Ecol.*, 23:1-9.
- Bloomfield, S.F., Stewart, G.S.A.B., Dodd, C.E.R., Booth, I.R., and Power, E.G.M. (1998). The Viable but nonculturable phenomenon explained? *Microbiol.*, 144:1-2.
- Boon, P.I. (1992). Antibiotic resistance of aquatic bacteria and its implications for limnological research. *Austral. J. Freshwat. Res.*, 43:847-859.
- Brock, M.L., and Brock, T.D. (1968a). The application of micro-autoradiographic techniques to ecological studies. *Internat. Assoc. Theor. Appl. Limnol.*, 15:1-29.
- Brock, T.D., and Brock, M.L. (1968b). Autoradiography as a tool in microbial ecology. *Nature*, 209:734-736.
- Byrd, J.J., Huai-Shu, X., and Colwell, R.R. (1991). Viable but nonculturable bacteria in drinking water. *Appl. Environ. Microbiol.*, 57:875-878.
- Choi, J.W., Sherr, E.B., and Sherr, B.F. (1996). Relation between presence-absence of a visible nucleoid and metabolic activity in bacterioplankton cells. *Limnol. Oceanogr.*, 41:1161-1168.
- Davey, H.M., and Kell, D.B. (1996). Flow cytometry and cell sorting of heterogeneous microbial populations: the importance of single-cell analyses. *Microbiol. Rev.*, 60:641-696.
- Deere, D., Porter, J., Edwards, C., and Pickup, R. (1995). Evaluation of the suitability of bis-(1,3-dibutylbarbituric acid) trimethine oxonol, (diBA-C₄(3)-), for the flow cytometric assessment of bacterial viability. *FEMS Microbiol. Lett.*, 130:165-170.
- Diaper, J.P., and Edwards, C. (1998). The use of fluorogenic esters to detect viable bacteria by flow cytometry. *J. Appl. Bact.*, 77:221-228.

Fuentes, F.A., Santo Domingo, J.W., and Hazen, T.C. (1998). Survival of *Candida albicans* and *Pseudomonas aeruginosa* in oil polluted tropical coastal waters. *Wat. Res.*, 32:2154-2170.

Grossmann, S. (1994). Bacterial activity in sea ice and open water of the Weddell Sea, Antarctica: A microautoradiographic study. *Microb. Ecol.*, 28:1-18.

Jorgensen, N.O.G., Tranvik, L., Edling, H., Graneli, W., and Lindell, M. (1998). Effects of sunlight on occurrence and bacterial turnover of specific carbon and nitrogen compounds in lake water. *FEMS Microbiol. Ecol.*, 25:217-227.

Kalmbach, S., Manz, W., and Szewzyk, U. (1997). Dynamics of biofilm formation in drinking water: Phylogenetic affiliation and metabolic potential of single cells assessed by formazan reduction and in situ hybridization. *FEMS Microbiol. Ecol.*, 22:265-279.

Karner, M., and Fuhrman, J.A. (1997). Determination of active marine bacterioplankton: a comparison of universal 16S rRNA probes, autoradiography, and nucleoid staining. *Appl. Environ. Microbiol.*, 63:1208-1213.

Kirchman, D., K'nees, E., and Hodson, r. (1985). Leucine incorporation and its potential as a measure of protein synthesis by bacteria in natural aquatic systems. *Appl. Environ. Microbiol.*, 49:599-607.

Knoechel, R., and Kalff, J. (1976). The applicability of grain density autoradiography to the quantitative determination of algal species production: A critique. *Limnol. Oceanogr.*, 21:583-590.

Kogure, K., Simidu, U., and Taga, N. (1979). A tentative direct microscopic method for counting living marine bacteria. *Can. J. Microb.*, 25: 415-420.

Lopez-Amoros, R., Castel, S., Comas-Riu, J., and Vives-Rego, J. (1997). Assessment of *E. coli* and *Salmonella* viability and starvation by confocal laser microscopy and flow cytometry using rhodamine 123, diBAC₄(3), propidium iodide, and CTC. *Cyt.*, 29:298-305.

Lopez-Amoros, R., Comas, J., and Vives-Rego, J. (1995). Flow cytometric assessment of *Escherichia coli* and *Salmonella typhimurium* starvation-survival in seawater using rhodamine 123, propidium iodide, and oxonol. *Appl. Environ. Microbiol.*, 61:2521-2526.

McDougald, D., Rice, S.A., Weichart, D., and Kjelleberg, S. (1998). Nonculturability: adaptation or debilitation? *FEMS Microbiol. Ecol.*, 25:1-9.

Meyer-Reil, L.-A. (1978). Autoradiography and epifluorescence microscopy combined for the determination of number and spectrum of actively metabolizing bacteria in natural waters. *Appl. Environ. Microbiol.*, 36:506-512.

- Nilsson, C., and Sundback, K. (1996). Amino acid uptake in natural microphytobenthic assemblages studied by microautoradiography. *Hydrobiologia*, 332:119-129.
- Nybroe, O. (1995). Assessment of metabolic activity of single bacterial cells - new developments in microcolony and dehydrogenase assays. *FEMS Microbiol. Ecol.*, 17:77-84.
- Nystrom, T. (1998). To be or not to be: The ultimate decision of the growth-arrested bacterial cell. *FEMS Microbiol. Rev.*, 21:283-290.
- Paerl, H.W. (1991). Ecophysiological and trophic implications of light-stimulated amino acid utilization in marine picoplankton. *Appl. Environ. Microbiol.*, 57:473-479.
- Peele, E.R., and Colwell, R.R. (1981). Application of a direct microscopic method for the enumeration of substrate-responsive marine bacteria. *Can. J. Microb.*, 27:1071-1075.
- Pernthaler, J., Sattler, B., Simek, K., Schwarzenbacher, A., and Psenner, R. (1996). Top-down effects on the size-biomass distribution of a freshwater bacterioplankton community. *Aquat. Microb. Ecol.*, 10:255-263.
- Porter, J., Deere, D., Pickup, R., and Edwards, C. (1996). Fluorescent probes and flow cytometry: New insights into environmental bacteriology. *Cyt.*, 23:91-96.
- Poulsen, L.K., Ballard, G., and Stahl, D.A. (1993). Use of rRNA fluorescence in situ hybridization for measuring the activity of single cells in young and established biofilms. *Appl. Environ. Microbiol.*, 59:1354-1360.
- Ramsay, A.J. (1974). The use of autoradiography to determine the proportion of bacteria metabolizing in an aquatic habitat. *J. Gen. Microbiol.*, 80:363-373.
- Ratinaud, J.-H., and Revidon, S. (1995). A flow cytometric method to assess functional state of the *Listeria* membrane. *J. Microbiol. Met.*, 25:71-77.
- Reasoner, D.J., and Geldreich, E.E. (1985). A new medium for the enumeration and subculture of bacteria from potable water. *Appl. Environ. Microbiol.*, 49:1-7.
- Rodriguez, G.G., Phipps, D., Ishiguro, K., and Ridgway, H.F. (1992). Use of a fluorescent redox probe for direct visualization of actively respiring bacteria. *Appl. Environ. Microbiol.*, 58:1801-1808.
- Schaule, G., Flemming, H.-C., and Ridgway, H.F. (1993). Use of 5-cyano-2,3-ditolyl tetrazolium chloride for quantifying planktonic and sessile respiring bacteria in drinking water. *Appl. Environ. Microbiol.*, 59:3850-3857.

- Schuppler, M., Mertens, F., Schon, G., and Gobel, U.B. (1995). Molecular characterization of nocardioform actinomycetes in activated sludge by 16S rRNA analysis. *Microbiol.*, 141:513-521.
- Singh, A., Yu, F.-P., and McFeters, G.A. (1990). Rapid detection of chlorine-induced bacterial injury by the direct viable count method using image analysis. *Appl. Environ. Microbiol.*, 56(2):389-394.
- Smith, J.J., Howington, J.P., and McFeters, G.A. (1994). Survival, physiological response, and recovery of enteric bacteria exposed to a polar marine environment. *Appl. Environ. Microbiol.*, 60(8):2977-2984.
- Tabor, P.S., and Neihof, R.A. (1985). Improved microautoradiographic method to determine individual microorganisms active in substrate uptake in natural waters. *Appl. Environ. Microbiol.*, 44:945-953.
- Tuomi, P., Fagerbakke, K.M., Bratbak, G., and Heldal, M. (1995). Nutritional enrichment of a microbial community: The effects on activity, elemental composition, community structure and virus production. *FEMS Microbiol. Ecol.*, 16:123-134.
- Ullrich, S., Karrasch, B., Hoppe, H.G., Jeskulke, K., and Mehrens, M. (1996). Toxic effects on bacterial metabolism of the redox dye 5-cyano-2,3-ditolyl tetrazolium chloride. *Appl. Environ. Microbiol.*, 62:4587-4593.
- Volk, C.J., Volk, C.B., and Kaplan, L.A. (1997). Chemical composition of biodegradable dissolved organic matter in streamwater. *Limnol. Oceanogr.*, 42:39-44.
- Yu, F.P., Pyle, B.H., and McFeters, G.A. (1993). A direct viable count method for the enumeration of attached bacteria and assessment of biofilm disinfection. *J. Microbiol. Met.*, 17:167-180.

Appendix II

Statistical Model for Active and Total Cells/mL

Carbon concentration is nested in substrate type and reactors are nested within substrates and concentration such that the final statistical model is

$$Y_{(ijkl)} = \mu + \alpha_i + \beta_{j(i)} + \kappa_{k(ij)} + \varepsilon_{l(ijk)}.$$

Each measurement (average) for any of the above parameters is denoted by $Y_{(ijkl)}$, where i , the number of the treatment (substrate, 1 is amino acids, 2 is carbohydrates, and 3 is mixture) runs from 1 to a ($a = 3$). The grand mean μ , is the overall mean for all substrates, concentrations and reactors. j , the number of treatments in the first nested variable (carbon concentration, 1 is 1ppm and 2 is 2 ppm) runs from j to b ($b = 2$). k , the number of the treatment in the second nested variable (chlorinated or control) runs from k to c ($c = 2$). l is the observation number which runs from l to n ($n = 3$) from the k^{th} experimental unit subjected to the j^{th} carbon concentration receiving the i^{th} substrate. For example, consider the active cells: $Y_{(ijkl)}$ is the mean number of active cells/mL from the l (3) observations from the treatment level k at the carbon concentration j that was receiving the carbon substrate i . Practically, if we consider the amino acids ($i = 1$) control reactor ($k = 1$) receiving carbon at a concentration of 1000 ppb ($j = 1$) we get the following results. The grand mean (in $\log_{10}$ transformed units) μ is 5.862 (± 0.0009), α_i is 0.129 (± 0.013), $\beta_{j(i)}$ is 0.061 (± 0.016), $\kappa_{k(ij)}$ is 0.233 (± 0.023) and $\varepsilon_{l(ijk)}$ is 0.00312. Addition of all these values gives us the least squares mean of 6.285 (± 0.032) which fits the observed value of 6.285 (± 0.01) very well.

Summarized Data

Substrate	Substrate Concentration									
	1000 ppb					2000 ppb				
	turnover ¹ %, hr ⁻¹	dpm/cell	Total cells/mL (x10 ⁵)	Active ² (cells/mL)	cfu/mL (x10 ⁵)	turnover ¹ %, hr ⁻¹	dpm/cell	Total cells/mL (x10 ⁵)	Active ² (cells/mL)	cfu/mL (x10 ⁵)
amino acids	16.4 (0.8, 0.97)	0.54	19.7	14.4 (2.2)	6.3	25.1 (2.3, 0.91)	0.45	22.6	14.9 (7.0)	6.5
	12.9 (0.5, 0.99)	0.68	19.7	14.0 (0.7)	6.3	25.8 (1.4, 0.97)	0.49	22.6	8.9 (18.0)	6.5
	nd	nd	nd	nd	nd	22.9 (1.7, 0.92)	0.36	57.9	40.5 (4.9)	31.
carbohydrates	1.4 (0.1, 0.90)	0.006	14.4	7.7 (5.2)	1.14	3.8 (0.2, 0.96)	0.0095	17.1	9.33 (5.0)	3.3
humics	nd	nd	nd	nd	nd	1.7 (0.02, 0.998)	0.1	8.9	5.2 (8.0)	3.6
mixture	1.19 (0.03, 0.996)	0.03	10.6	5.6 (8.0)	3.73	1.7 (0.1, 0.95)	0.016	37.3	19.7 (5.0)	14.1
	0.9 (0.05, 0.98)	0.03	9.1	6.7 (6.8)	2.70	4.3 (0.3, 0.94)	0.065	22.8	16.6 (5.2)	7.3

Table 3.4 – Control reactor results for MAR experiments. 1 – the numbers in brackets 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. cfu/mL is colony forming units on R2A agar.

	Substrate Concentration									
	1000 ppb					2000 ppb				
Substrate	turnover ¹ (%, hr ⁻¹)	dpm/cell	Total cells/mL (x10 ⁵)	Active ² (cells/mL)	cfu/mL (x10 ⁵)	turnover ¹ (%, hr ⁻¹)	dpm/cell	Total cells/mL (x10 ⁵)	Active ² (cells/mL)	cfu/mL (x10 ⁵)
amino acids	1.58 (0.02, 0.999)	0.061	8.0	5.3 (6.4)	1.1	5.5 (0.2, 0.98)	0.19	9.1	4.73 (4.4)	4.6
carbohydrates	0.025 (0.004, 0.77)	0.0017	8.9	2.7 (0.3)	0.13	0.017 (0.002, 0.85)	0.0012	6.03	2.8 (2.2)	0.03
humics	nd	nd	nd	nd	nd	0.006 (0.001, 0.64)	0.0006	5.4	2.4 (7.4)	0.02
mixture	0.12 (0.01, 0.94)	0.0074	6.5	3.0 (3.4)	0.15	0.17 (0.02, 0.85)	0.0019	29.6	15.7 (8.4)	0.87
	3.5 (0.4, 0.93)	0.165	8.6	4.3 (2.3)	0.90	1.27 (0.07, 0.97)	0.018	24.1	12.4 (2.3)	0.44

Table 3.5 – Chlorinated reactor results for MAR experiments. 1 – the numbers in brackets 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. cfu/mL is colony forming units on R2A agar.

CHAPTER 4

EFFECTS OF CARBON SOURCE, CARBON CONCENTRATION, AND CHLORINATION ON GROWTH RELATED PARAMETERS OF HETEROTROPHIC BIOFILM BACTERIA

Introduction

Microbial utilization of dissolved organic carbon (DOC) in natural freshwater ecosystems has mainly been studied in terms of planktonic bacteria (Amon and Benner, 1996; Chrzanowski, et al., 1995; Coveney and Wetzel, 1992; Currie, 1990; Gremm and Kaplan, 1997). Traditionally, low molecular weight organics such as amino acids and carbohydrates have been estimated to be responsible for a significant fraction of net carbon production in planktonic bacterial communities from a variety of lakes and rivers (Jorgensen, et al., 1998; Jorgensen and Jensen, 1994; Hanisch, et al., 1996). Conversely, in recent years, high utilization of DOC material other than the simple compounds mentioned above has been estimated to comprise a large fraction of net bacterial carbon production in freshwater and marine ecosystems (Amon and Benner, 1994; Volk, et al., 1997; Amon and Benner, 1996; Tranvik, 1993; Tranvik, 1988). Amon and Benner (Amon and Benner, 1994; Amon and Benner, 1996) and more recently Volk, et al. (Volk, et al., 1997) have demonstrated that the high molecular weight fraction of dissolved organic matter (DOM) is less recalcitrant than previously thought.

While most studies have mainly focused on the planktonic component of the food web, a number of investigations have implicated surface-associated communities as being responsible for significant carbon turnover rates (Volk, et al., 1997; Bryers, 1987; Paerl and Pinckney, 1996; Kaplan and Newbold, 1993; van Loosdrecht, et al., 1990; Wolfaardt, et al., 1994). These bacterial communities have been shown to contribute disproportionately to nutrient turnover rates as compared to their planktonic counterparts. Nitrogen, carbon, and sulphur cycles can be compressed into linear zones comprising millimeters in microbial mats, trickling filters, and sediments as compared to meters and tens of meters in lakes and oceans. Kaplan and Newbold (Kaplan and Newbold, 1993) (and references therein) estimated that benthic populations in streams and rivers are significant sites for DOC utilization and are the primary mediators of DOC diagenesis. This pool of natural organic matter (NOM) contains the reduced organic carbon required for the production of energy and biomass during heterotrophic bacterial metabolism. Utilization of NOM can result in the growth of bacteria in both natural and engineered aquatic environments.

An area of applied microbiology that is significantly impacted by the presence of utilizable carbon is drinking water. Excess organic material in drinking water is associated with bacterial growth, increased chlorine demand in finished waters, and the formation of disinfection by-products in distribution systems (LeChevallier, et al., 1987; LeChevallier, et al., 1991; Putnam and Graham, 1993; McFeters, 1990; Sobsey, et al., 1993; LeChevallier, et al., 1992; Glaze, et al., 1993). While the measurement of specific utilizable components of NOM has not been the focus of drinking water research,

measurement of the total utilizable fraction has encompassed a significant amount of research in the last 10 years (van der Kooij, et al., 1982; van der Kooij, 1992; Frias, et al., 1994; Lucena, et al., 1990; Ribas, et al., 1995; Kaplan and Bott, 1990; Huck, 1990). A number of methods have been applied to measure the fraction of NOM present in aquatic systems, which can be divided into two basic groups: (1) assimilable organic carbon (AOC) and (2) biodegradable organic carbon (BDOC).

AOC is generally measured by the method of van der Kooij (van der Kooij, et al., 1982; van der Kooij, 1992), whereas BDOC is generally measured by the method of Servais, et al. (Servais, et al., 1987). While there are differences in the application, utilization, and interpretation of the two methods, ultimately, the major difference between the two is the parameter measured. AOC measures biomass/cell number changes over time utilizing a known inoculum and represents the potential of the water source to support bacterial growth (van der Kooij, et al., 1982; Shirey and Bissonnette, 1991; Bradford, et al., 1994; Kaplan and Bott, 1990). BDOC measures a change in the TOC/DOC content of the water over time utilizing an indigenous inoculum. Biomass/cell number (AOC) measurements are generally used when the concern is bacterial growth whereas BDOC is used when the concern is a reduction in the chlorine demand or formation of disinfection by-products (DBPs). Both methods can be time consuming and laborious, with a significant amount of time passing before the results are known. Lucena, et al. (Lucena, et al., 1990) introduced a BDOC method that utilizes surface attached, indigenous bacterial communities continually supplied with the water of interest and provides results within 2 to 3 hours (Kaplan and Newbold, 1995; Frias, et al., 1994;

Volk, et al., 1997; Kott, et al., 1997; Ribas, et al., 1995). While there are problems with this approach, such as the potentially extended period of adaptation required for colonization of the inert support in the columns, it has been used for estimation of kinetic constants (eg. K_m and V_{max}), growth rates, and BDOC component identification. Volk, et al. (1997) used this method as an aid in determining the contribution of different components of the BDOC responsible for a decrease in DOC across these reactors. None of these studies has investigated the effects of chlorine on the communities present in the columns, drinking water, or distribution systems.

In this study, we utilized a biofilm reactor system (annular reactors) to study community response to three of the commonly accepted primary DOC components in aquatic environments. Three different organic substrates comprising amino acids, carbohydrates, and humics were separately tested for their ability to support heterotrophic bacterial communities under varying trophic conditions. AOC-“free” drinking water was generated by use of granular activated carbon/biologically activated carbon (GAC/BAC) filtration and used as dilution water for the reactors. This dilution water entered the reactors at a maximum carbon loading rate of 1.0 mg C/L_oday, whereas the reactors themselves had carbon loading rates of approximately 6, 12, and 24 mg C/L_oday. These values were comparable to the mean BDOC loading rate entering (i.e. finished drinking water) the GAC/BAC filters, which was 18 mg C/L_oday. Chlorinated and control reactors were operated in parallel to examine the effects of oxidative stress and nutrient limitation on growth parameters. Operationally, the annular reactors could be viewed as biological filters similar to BDOC columns, but they are much easier to model as continuously

stirred tank reactors (CSTRs). This allowed us to measure specific growth rates (SGRs), observed yields/growth efficiency, bacterial carbon production (BCP), and biomass/cell density at different substrate loading rates. BDOC columns and a carbon unsupplemented (negative) control annular reactor were used to measure the BDOC content of the AOC-“free” water.

Materials and Methods

See Chapter 2 for the specifics of experimental setup and general protocols of the experimental procedures.

Experimental Overview

A total of nine experiments were run, with amino acids ($n = 3$), carbohydrates ($n = 3$), humics ($n = 2$), and a mixture ($n = 1$) of the three substrates. From these experiments, 5 individual samples ($n = 5 \times 3 = 15$) were used for substrate evaluation with respect to total direct counts (TDC) and colony forming units (cfu) in the effluent and biofilm at 3 different carbon loads: (1) 500 ppb, (2) 1000 ppb, and (3) 2000 ppb. The same samples were used to determine SGRs, BCP, and observed yields. The reactors were designated as control (no chlorination), or test (chlorination at 1000 and 2000 ppb C) on a random basis. Since both reactors are receiving the same treatment at the 500 ppb C concentration, evaluation of bacterial density in the effluent and biofilm was used to assess the reproducibility of the annular reactors.

The substrate mixture experiment was designed to complement the other experiments, but with a few changes in protocol. Carbon feed was initiated at a concentration of 250 ppb (equimolar as carbon, 2:1:1 mixture of humics, amino acids, and carbohydrates respectively) and chlorination was started at this stage.

An additional annular reactor was operated as a negative control without supplemental carbon feed. Buffer was supplied to this reactor (as nitrate and phosphate, 612 and 591 $\mu\text{g/L}$ respectively) and sampling was performed periodically over the course of twelve months.

Since the effluent from the GAC/BAC columns was used for colonization of the annular reactors as well as a supply of dilution water, background carbon in the drinking water could have been a confounding factor in substrate evaluations. Therefore, tap water influent and effluent of the GAC/BAC columns was sampled for DOC removal in order to ascertain the amount of carbon removed across these biological filters.

An additional control used BDOC columns and were operated as described by Ribas, et al. (1995). Briefly, SIRAN glass bead (Schott Glassworks) filled chromatography columns (Kontes) were operated (in the dark) for the measurement of the BDOC present in the AOC "free" water. These columns (in series) were maintained on the influent water (i.e. GAC/BAC effluent) used for the annular reactors for 3 months prior to initiation of sampling (sampling period of 261 days). Prior to this the columns had been operating separately (i.e. in parallel) for 1 year on City of Bozeman tap water. Hydraulic detention time for the BDOC columns was set to be equivalent to the annular

reactors (2.1 hours) while the hydraulic detention time in the GAC/BAC columns was 15 minutes.

TOC Measurements

All TOC measurements were made on the nonpurgeable organic carbon (NPOC) fraction of the water.

Carbon Balance Across Reactors. NPOC was determined as described above for all input parameters (AOC-“free” dilution water, buffer and carbon supplement) and reactor output. The following designations are utilized:

- 1) AOC_f is the filtered AOC dilution water (DOC)
- 2) AOC_{nf} is the unfiltered AOC dilution water (TOC)
- 3) B_{nf} is the unfiltered buffer solution (TOC)
- 4) C_f is the filtered carbon solution (DOC)
- 5) C_{nf} is the unfiltered carbon solution (TOC)
- 6) $C\text{-out}_f$ is the filtered carbon in the reactor effluent (DOC)
- 7) $C\text{-out}_{nf}$ is the unfiltered carbon in the reactor effluent (TOC)

The total amount of soluble carbon entering the reactors was determined from the relationship

$$\text{Carbon in} = (4.5/5.0) * (AOC_f) + (0.25/5.0) * (B_{nf}) + (0.25/5.0) * (C_{nf}) \quad (1)$$

where the numbers refer to the flow rates (mL/min) for each solution. Similarly, the total amount of soluble carbon in the reactor effluent is

Carbon out = C_{out_f}

Both the buffer and the AOC dilution water (AOC_f) were assumed to be unavailable for bacterial utilization and are summed to give the non-utilizable fraction as shown below:

$$\text{Non-utilizable} = (0.25/5.0) * (B_{\text{nf}}) + (4.5/5.0) * (AOC_f) \quad (2)$$

C_{out_f} represents the total concentration of DOC that was not removed across the annular reactor. If the non-utilizable fraction was wholly represented in both the influent and effluent carbon measurements, it could be subtracted from the total amount of dissolved carbon in the reactor effluent (i.e. C_{out_f}). This difference represents the amount of the supplemental dissolved organic carbon substrate feed that was not removed across the reactors. Subtracting this difference from the supplemental substrate feed concentration in the influent gave us the amount of the supplemental feed carbon that was removed across the reactors. The supplemental substrate feed concentration was determined directly from the feed carboy. The difference between the unfiltered (TOC) and filtered (DOC) solutions gave us the amount of particulate organic carbon (POC). This POC arises from the bacterial cells in the water and was used to determine an average value for carbon/cell when divided by the appropriate total cell count for that sample. One to three carbon determinations were made at the pseudo-steady state stage of duplicate experiments and average carbon substrate removal amounts were computed from these samples. Similar measurements were used to determine the amount of carbon removed across the BDOC columns, GAC/BAC filter, and negative control reactor.

Carbon Source Utilization

Biolog GN microplates (Hayward, CA) are 96 well microtitre plates. Each well contains a redox indicator (tetrazolium violet) and a single dehydrated carbon substrate as well as all other essential nutrients. A Bio-Rad (Richmond, CA) Model 450 Microplate Reader equipped with a 570 nm interference filter was used for all measurements using the dual wavelength mode. At the biofilm sampling stages (1000 and 2000 ppb C only) for duplicate experiments a 0.15 mL inoculum from each reactor (scraped and homogenized biofilm) was introduced into each well in triplicate plates. Inoculum density was approximately 5×10^6 cells/mL as determined by direct microscopic count of 4',6-diamidino-2-phenylindole (DAPI) stained cells prior to inoculation. Time series readings of the Biolog plates were taken every 24 hours for 4 to 5 days (until the response in the wells had leveled off) and the baseline response in the blank well was subtracted from all other absorbance readings. This baseline reading is for a negative control well which contains the redox dye (+ nutrients), but no carbon substrate. For our application, the plates contained the 4 amino acids and 3 carbohydrates chosen as model substrates for this study. One carbohydrate substrate used in our experiments (D - arabinose) was represented in the Biolog plates as the L-epimer (L-arabinose). The response to these substrates were compared to evaluate the potential effect of priming the communities grown in the reactors for the particular set of substrates that they had been utilizing. Namely, would growth on amino acids affect the rate of carbohydrate utilization in the Biolog plates? Conversely, would growth on carbohydrates affect amino acid utilization in the Biolog plates?

Influent and Effluent Cell Sampling

Routine cell sampling of the reactors included samples of the dilution water (i.e. reactor influent from the GAC/BAC filter) and effluent from each reactor for total and culturable cells. Samples for total cell counts were fixed using formalin (0.7% final concentration) for 24 hours at 4°C. Fixed samples were filtered onto 0.2 µm pore size black polycarbonate filters (Poretics), stained on the filter using DAPI (0.1 µg/ml) for 30 minutes and rinsed with 5 ml of filter sterilized Nanopure water. Culturable cell counts were determined by spread plating dilutions in triplicate on R2A agar medium and incubating at room temperature (22 - 23°C) for 7 days. Chlorine concentrations were determined using the N,N-diethyl-p-phenylenediamine (DPD) colorimetric method and a digital colorimeter (LaMotte). Both total free residual and combined residuals were determined in the effluent as well as chlorine dosage levels applied to the reactors.

Specific Growth Rates, Yield, and Bacterial Carbon Production

In order to determine SGRs, observed yields, and BCP the reactors were modeled as CSTRs such that mass balances across the reactors can be solved for the particular parameters, assuming complete mixing and uniform conditions with the reactor. The mass balance for planktonic cells gives the following equation:

$$V \frac{dX_1}{dt} = F(X_0 - X_1) + \mu X_1 V + r_d X_b A \quad (1)$$

where

V = liquid volume (cm^3)

X_1 = planktonic cell concentration in reactor liquid (mass cm^{-3})

X_0 = planktonic cell concentration entering the reactor (mass cm⁻³)

X_b = biofilm cell density (mass cm⁻²)

F = volumetric flow rate through reactor (cm³ hr⁻¹)

μ = specific growth rate of planktonic cells (hr⁻¹)

r_d = specific biofilm detachment rate (hr⁻¹)

A = surface area for biofilm accumulation in reactor (cm²)

A mass balance for the biofilm cells gives the following equation:

$$A \frac{dX_b}{dt} = \mu_b X_b A - r_d X_b A \quad (2)$$

where,

μ_b = specific growth rate of biofilm cells (hr⁻¹)

At steady state all derivatives are assumed to equal zero such that equations (1) and (2) simplify to:

$$\frac{F}{V}(X_1 - X_0) = \mu X_1 + r_d X_b \frac{A}{V} \quad (3)$$

$$\mu_b = r_d \quad (4)$$

If the hydraulic residence time is short enough to preclude significant planktonic growth, the term μX_1 can be neglected such that equations (3) and (4) can be solved, thus obtaining the kinetic parameters for the biofilm cells. The term r_d can be solved using the following equation:

$$\frac{F}{V}(X_1 - X_0) = r_d X_b \frac{A}{V} \quad (5)$$

Equation (4) can then be used to solve for the specific growth rate of the biofilm cells (μ_b). Assuming steady state conditions and minimal planktonic growth an observed yield (Y_{obs}) can be calculated for the reactors using the following equation:

$$Y_{obs} = \frac{(X_1 - X_0)}{(S_0 - S_1)}$$

where,

S_0 = substrate concentration in the reactor influent and S_1 = substrate concentration in the reactor effluent.

Bacterial carbon production (BCP, $\mu\text{g Cell carbon}/\text{cm}^2 \cdot \text{hr}$) can be calculated using the biovolume and carbon/ μm^3 information obtained from Chapter 2 (weighted nonlinear regression equation) in conjunction with the physical parameters of the reactors and operating protocols. Assuming that the difference between influent and effluent total cell counts at steady state is due to biofilm growth in the reactor, a simple subtraction ($X_1 - X_0$) gives the number of bacteria produced from the biofilm and conversion to biomass provides BCP.

Biofilm Sampling

The biofilm in the reactors was sampled by removal of a polycarbonate slide (area of approximately 30 sq. cm) and aseptically scraping the biofilm from the slide into a sterile buffer solution. The biofilm/buffer solution was then homogenized for 1 min using

a tissue homogenizer (Tekmar) at the maximum setting. This suspension was then analyzed for total and culturable cell counts as described in previous paragraphs. Evaluation of the homogenization procedure was performed by direct microscopic examination. Cell clustering was not a factor and the homogenization procedure was effective in dispersing the cells in the biofilms.

Statistical Analyses

A general linear model (GLM) analysis of variance (ANOVA) was used to analyze all data with nested treatments designated as chlorination/control, substrate type, and carbon feed level (Sokal and Rohlf, 1995; Minitab, 1998; Dunn and Clark, 1974) (see Chapter 1, Figure 1.1). In addition, Tukey's pairwise comparisons were used when one factor had more than 2 levels and data analysis was within one treatment factor. A one sample t-test was used where specific means were being compared. Acceptance or rejection of the null hypothesis was based on an α -level of 0.05 in all cases. The statistical program Minitab (1998) was used for all analyses. Model II regressions were carried out by the method of geometric means as detailed in Sokal and Rohlf (1995)

Results

GAC/BAC Filter and BDOC Column Performance

The BDOC columns were operated as separate parallel units on unsupplemented tap water for over one year. Carbon removal across the two columns was generally lower than for the GAC/BAC filter. There was no significant difference between the two BDOC columns, the first column removed $9.2 \pm 4.8 \%$, ($n = 11$, $\pm$ S.D.) while the second column removed $10.0 \pm 4.7 \%$ ($n = 10$) of the influent native DOC. Both were significantly different from the GAC/BAC filter where carbon removal averaged $18 \pm 4.7 \%$ ($n = 19$) of the influent native DOC. In terms of absolute carbon removal the BDOC columns averaged $93 \pm 51 \mu\text{g C/L}$ while the GAC/BAC filter removed an average of $157 \pm 86 \mu\text{g C/L}$ over the same sampling period. These differences are magnified when one considers that the hydraulic residence time in the GAC/BAC filter was 15 minutes compared to 2.1 hours for the BDOC columns. For all samples (773 days) the GAC/BAC filter removed an average of $190 \pm 178 \mu\text{g C/L}$ ($17 \pm 10\%$). When the BDOC columns were run in series and receiving AOC "free" water the percentage removal of utilizable carbon was $2.4 \pm 5.4\%$ and equated to $32 \pm 38 \mu\text{g C/L}$. This compares to the negative control reactor operated with the AOC "free" water where supplementation with nitrate and phosphate produced no net (average) removal across the reactors of $-3.3 \pm 7.4\%$ or $-11 \pm 38 \mu\text{g C/L}$ ($n = 7$). Neither of these values was significantly different than zero ($p = 0.28$ and 0.67 ,

respectively) by a one-sample t-test where the null hypothesis mean was zero (i.e., $H_0 = 0$). The GAC/BAC filter released bacteria at a steady rate with the total cell counts producing less variability than the culturable cells. There did tend to be a seasonal component with total cells decreasing in the order March-May > June-August > September-November > December-February. Over the course of the experiments, cell numbers in the influent (i.e. GAC/BAC effluent) to the annular reactors averaged 3.89×10^5 total cells/ml (95% C.I., $3.64 \times 10^5 - 4.15 \times 10^5$, $n = 247$) and 5.66×10^4 cfu/ml ($4.19 \times 10^4 - 6.53 \times 10^4$, $n = 260$).

Control Stage (500 ppb Carbon Feed Level) Annular Reactors

Effluent. Frequent sampling was performed at this stage of the study in order to evaluate the reproducibility of the carbon supplemented annular reactors within any one experiment. Figure 4.1 shows the results from biofilm and effluent sampling at the control stage of the experiments. A GLM ANOVA was used to analyze the effluent data from the amino acids ($n = 15$), carbohydrates ($n = 15$), and humics ($n = 10$). The mixture experiment was not included in this analysis due to the different protocols used (initial feed concentration was 250 ppb C) and the lack of true replication. However, the means and standard deviations ($n = 5$) for the unchlorinated mixture control reactor are shown for comparison. Determination of steady state at this stage (500 ppb C) was made by the equivalency of total and culturable cells with respect to time in the reactor effluent. The GLM ANOVA ($n = 5$), effluent averages from the low carbon feed level of each experiment (3 experiments gives $n = 15$, or 2 experiments gives $n = 10$) and subsequent

analysis showed that the individual reactors were equivalent within substrates for both total cells ($p = 0.581$) and culturable cells ($p = 0.178$). Based on these results, the within substrate data was pooled and both the total cell counts (tdc) and culturable cell counts (cfu) were significantly different between substrates ($p < 0.0005$ for both cfu and tdc, $n = 30, 30$ and 20). Tukey's pairwise comparisons indicated that the substrates supported a significantly different number of total and culturable cells ($p < 0.0005$ for both), but culturable cells were equivalent for the amino acids and carbohydrates. The percentage variance in the data for the unexplained error components was 1.5% for total cells and 3.2% for culturable cells and the between experiments variance component was 10.7% for total cells and 8.8% for culturable cells. Most of the data variance occurred at the level of substrate. As a measure of the culturability within individual substrates we can look at the percentage of culturable cells with respect to total cells (i.e. $\text{cfu}/\text{tdc} \times 100$) in the reactor effluents. This ratio and the correlation between culturable and total cells are reported in Table 4.1 for paired samples from the control and test stage of the experiments. All three supplements produced a significant increase in the ratio of cfu/tdc ($p < 0.0005$) over that measured in the influent.

Biofilm. Biofilm sampling was performed once for each of the amino acids and carbohydrates experiments ($n = 3$). The humic experiment was performed twice and two samples were obtained from one of the experiments giving a pseudo-replication of $n = 3$. The mixture experiment was performed once and the results are based on 3 separate samples for the total cells ($n = 3$) and culturable cells ($n = 3$). Biofilm sampling at this stage also indicated that the reactors within any one substrate were equivalent for total

and culturable cells ($p = 0.088$ and 0.868 , respectively) prior to increasing the carbon loading and initiation of chlorination.

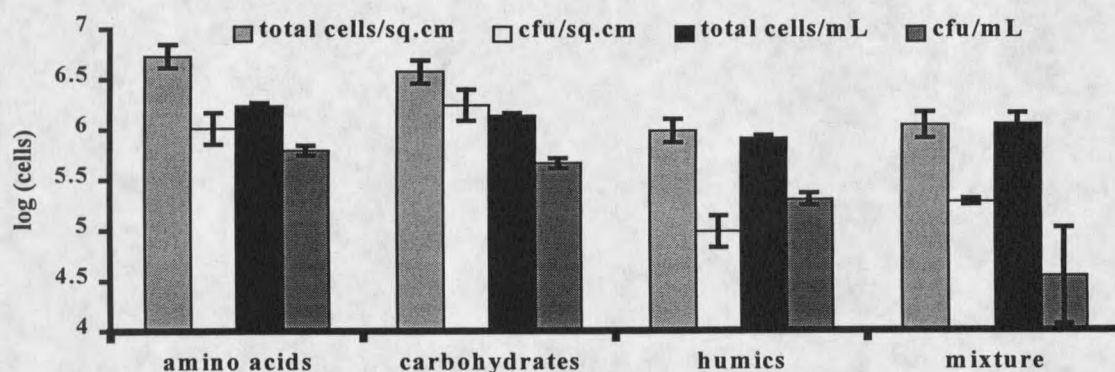


Figure 4.1 – Cell counts in the effluent and biofilm from the control stage of the experiments separated by substrate type. Total cells/mL are significantly different for all substrates while the amino acids and carbohydrates are equivalent for culturable cells/mL (Tukey's pairwise comparisons). $n =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

Test Stage (1000 and 2000 ppb Carbon Feed Levels) Annular Reactors

Figure 4.2 Panels A and B show the total cells and cfu per square centimeter in the biofilms for the control and chlorinated treatments at the carbon loading levels of 1000 and 2000 ppb. Reported values are means and standard deviations resulting from a GLM ANOVA, with carbon levels nested within substrates, performed on the separate reactors. Response to the substrates in the control reactor was compared at the three carbon levels (Figures 4.1 and 4.2). In the presence of chlorine, comparison between substrates was analyzed at the mid and high carbon feed levels due to the added treatment of chlorination.

Substrate	Measure	500 ppb C	Control	Chlorine	
			Test Level	1000 ppb C	2000 ppb C
amino acids	r^2	0.593 (71)	0.763 (100)	0.667 (72)	0.424 ¹ (46)
	cfu/tdc(%)	49 (26)	42 (21)	23 (16)	21 (15)
carbohydrates	r^2	0.511 (47)	0.675 (60)	0.684 (42)	0.508 ² (16)
	cfu/tdc(%)	41 (22)	42 (23)	8 (12)	5 (14)
humics	r^2	0.424 ¹ (63)	0.366 (90)	0.565 (51)	0.365 ³ (48)
	cfu/tdc(%)	35 (31)	36 (16)	8 (14)	1.5 (2.5)

Table 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. 500ppb 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.

Biofilm. All substrates supported higher biofilm cell numbers with increasing carbon feed levels in the control reactor although the only significant increases in the total ($p = 0.001$) and culturable cells ($p < 0.0005$) were from the humics and mixture experiments. The carbohydrates and amino acids biofilm cell numbers (total and culturable cells) did not change significantly between the three carbon loading levels ($p = 0.093$, $p = 0.403$, and $p = 0.502$ and $p = 0.003$ respectively). Total and culturable cells in the biofilm were significantly impacted by chlorination in the presence of carbohydrates ($p = 0.001$ and $p < 0.0005$, respectively) and humics ($p = 0.001$ and $p < 0.0005$, respectively) resulting in decreased cell numbers compared to the low carbon feed level. Increased carbon loading (1000 ppb to 2000 ppb) had a relatively minor affect on the total cell density in the biofilm such that cell numbers stabilized in the presence of

chlorine for humics and carbohydrates whereas the mixture and amino acids showed increased population densities. Chlorination did not affect the culturable biofilm cell density in the presence of amino acids. Comparing substrates within the control reactors indicated that total cell counts in the biofilm were equivalent ($p = 0.246$) while the culturable cells were significantly different ($p = 0.005$). This difference is due to the humics, which supported a significantly lower amount of culturable cells. Culturable cells in the biofilm were significantly different ($p < 0.0005$) for all substrates once chlorination was initiated. While the mixture experiment was performed only once, it was the only carbon source that produced significantly higher total cell densities at each carbon level ($p < 0.0005$). Similar to the amino acids, chlorination had no affect on the number of total cells ($p = 0.080$) when carbon concentrations were increased (i.e. 1000 ppb to 2000 ppb) in the presence of the mixture. The negative control reactor had 2.45×10^6 total cells/cm² (95% C.I. $1.97 \times 10^6 - 3.01 \times 10^6$) and 4.36×10^4 cfu/cm² ($1.87 \times 10^4 - 1.02 \times 10^5$) in the biofilm determined on two separate sampling dates. This can be compared to the maximum cell density in the carbon supplemented reactors for the mixture experiment at 2000 ppb C [1.06×10^8 total cells/cm² ($4.36 \times 10^7 - 2.57 \times 10^8$)] and the amino acids experiments at 1000 ppb C [9.12×10^6 cfu/cm² ($6.8 \times 10^6 - 1.22 \times 10^7$)].

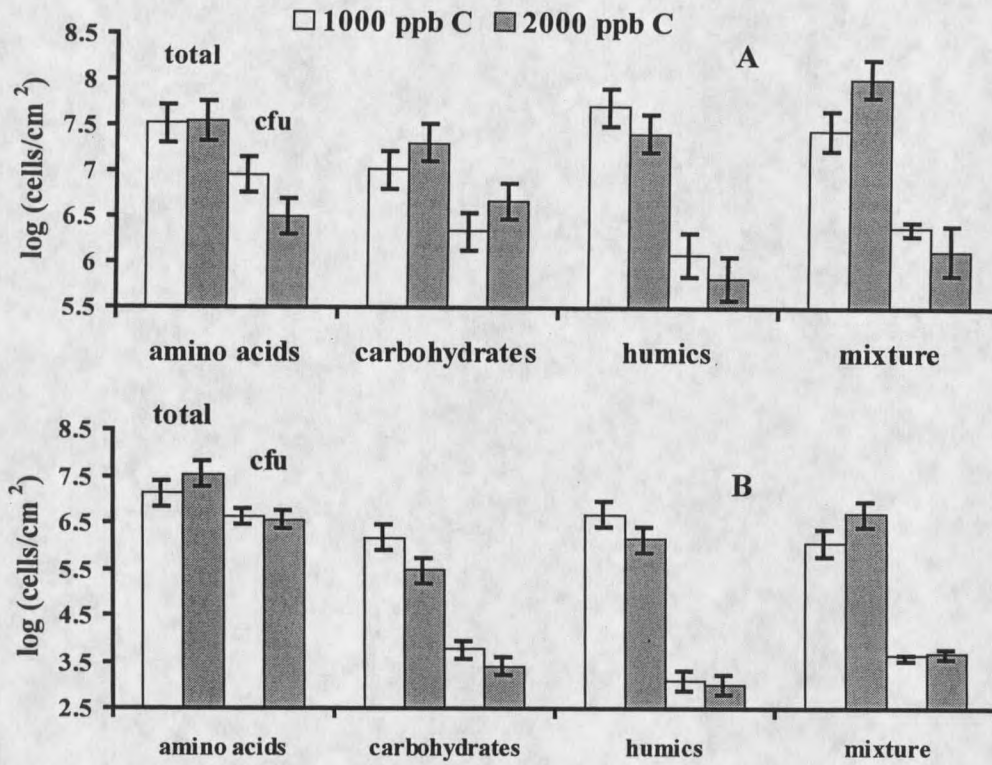


Figure 4.2 – Total and culturable cells per square cm in the biofilm for the four substrates and two higher carbon feed levels. Reported values are means and standard deviations resulting from a GLM ANOVA on the separate reactors. Panel A is the control reactor and Panel B in the chlorinated reactor. Error bars are standard deviations. See text for further details.

Effluent. Effluent cell measurements were required for the determination of growth rates, yields, and BCP as shown in the Materials and Methods section in addition to the biofilm cell densities shown above. See Figure 4.3 panels A and B for effluent culturable and total cells from the replicated experiments and Table 4.1 for percent culturable cells and correlations between total cells and culturable cells at this stage of the study. There was no significant difference in percent culturable cells between the two

higher carbon feed levels in the control reactors. While culturability tended to decrease as chlorination time proceeded, this decrease was not significant. The substrates and carbon levels within substrates were significantly different for total and culturable cells ($p < 0.002$ for all four) in both the control and chlorinated reactor effluents. Total and culturable cells in the control reactor effluents generally increased over the course of the experiment, however, significant increases were only noted for the 1000 to 2000 ppb C levels. These increases were only significant for the amino acids and carbohydrates ($p < 0.002$) whereas the humics did not support a significantly different number of effluent cells at any of the carbon feed levels. Increased carbon loading in the amino acids chlorinated reactor (from 1000 to 2000 ppb C) did not have a significant effect on the total or culturable cells in the effluent, whereas the carbohydrate cell counts decreased upon increased carbon loading in the presence of chlorine. The humics were the only substrate that produced significantly higher cell counts (2-sample t-test, $p = 0.022$ and $p = 0.0001$ for total and culturable cells, respectively) in the presence of chlorine when the carbon feed was increased between the mid and high levels. As an indication that the responses we were measuring were due to the added carbon supply and not utilization of organic carbon in the AOC "free" water from the GAC/BAC columns, we compared the influent and effluent cell numbers in the negative control reactor. There was no significant difference in either the total or culturable effluent cell counts in the negative control over the course of 250 days when compared to the influent cell numbers ($p = 0.223$ and 0.237 respectively, $n = 16$). Effluent total cells were 2.90×10^5 cells/ml (95%

C.I., $2.18 \times 10^5 - 3.86 \times 10^5$) while culturable cells were 1.03×10^5 cfu/ml ($6.00 \times 10^4 - 1.79 \times 10^5$).

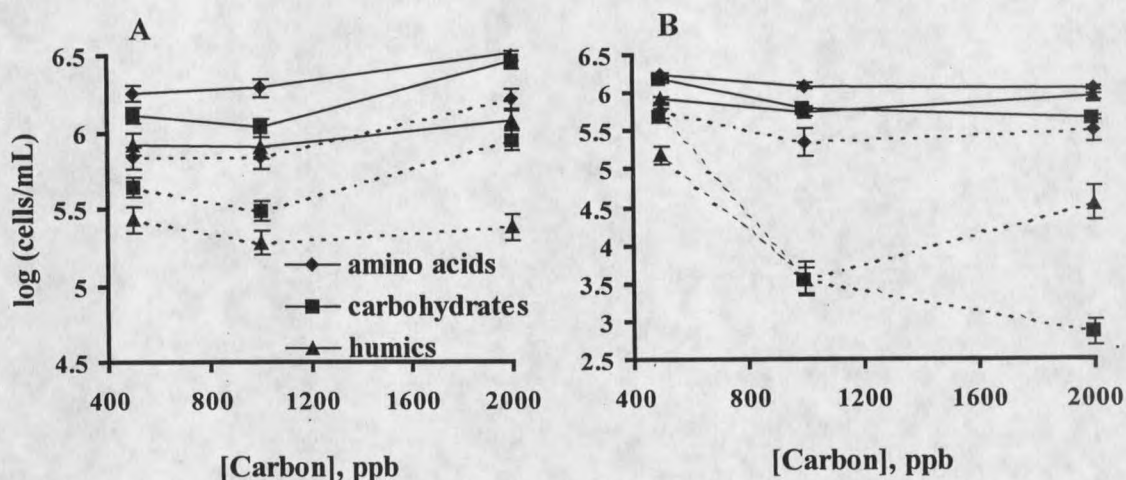


Figure 4.3 – Total and culturable cells per mL in reactor effluents. Panel A is control reactor effluent and Panel B is the chlorinated reactor effluent. Solid lines are total cells/mL and dashed lines are cfu/mL.

Growth Rates, Yields, and Bacterial Carbon Production

Data collected during operation of the annular reactors was used to obtain the information necessary to determine specific growth rates (SGR), substrate flux, observed yield, and bacterial carbon production (BCP) for the biofilm organisms. The data collected included influent, effluent and biofilm cell numbers (tdc and cfu), cell volumes, influent and effluent DOC concentrations (as NPOC).

The overall response for the data from all substrate groups (combined data) was first analyzed to determine general trends. Line plots for SGR, BCP, and observed yields for the control and chlorinated biofilms broken down into substrate and carbon levels are

presented in Figures 4.4 and 4.5. P-values indicate the results from a one-way ANOVA for the carbon levels, with the control reactors compared at 500, 1000, and 2000 ppb C while the chlorinated reactors were compared at 1000 and 2000 ppb C. There were significant differences for all three parameters between substrate groups ($p < 0.0005$) and carbon levels within substrate groups ($p < 0.0005$). The only exceptions were the chlorinated SGRs where the carbon levels within the substrates were not significantly different ($p = 0.409$). A one-way ANOVA indicated there was a significant difference between reactor biofilm growth rates within substrate groups ($p < 0.0005$) and that there was no significant difference between amino acids and carbohydrates. However, there was a significant difference between SGRs on the humics and the other two substrates (Tukey's pairwise comparisons). In general, SGRs for the humics reactors were less than those from the carbohydrates and amino acids reactors. SGRs in the presence of chlorine were significantly different ($p = 0.006$) and greater than those in the absence of chlorine. There was a significant difference between carbon levels ($p < 0.0005$), with the low level significantly different from the mid and high levels (Tukey's pairwise comparisons) but no significant difference between the mid and high level. Reactor biofilm SGRs at the 500 ppb C level were greater than both the 1000 and 2000 levels. The 500 level was always prior to the commencement of chlorination. There was a significant negative correlation between SGR and the biomass areal density in the biofilm for both the control and chlorinated communities ($r^2 = -0.637$, $p = 0.001$ and $r^2 = -0.570$, $p = 0.021$, respectively). No correlations were detected between SGR and influent or effluent substrate concentration.

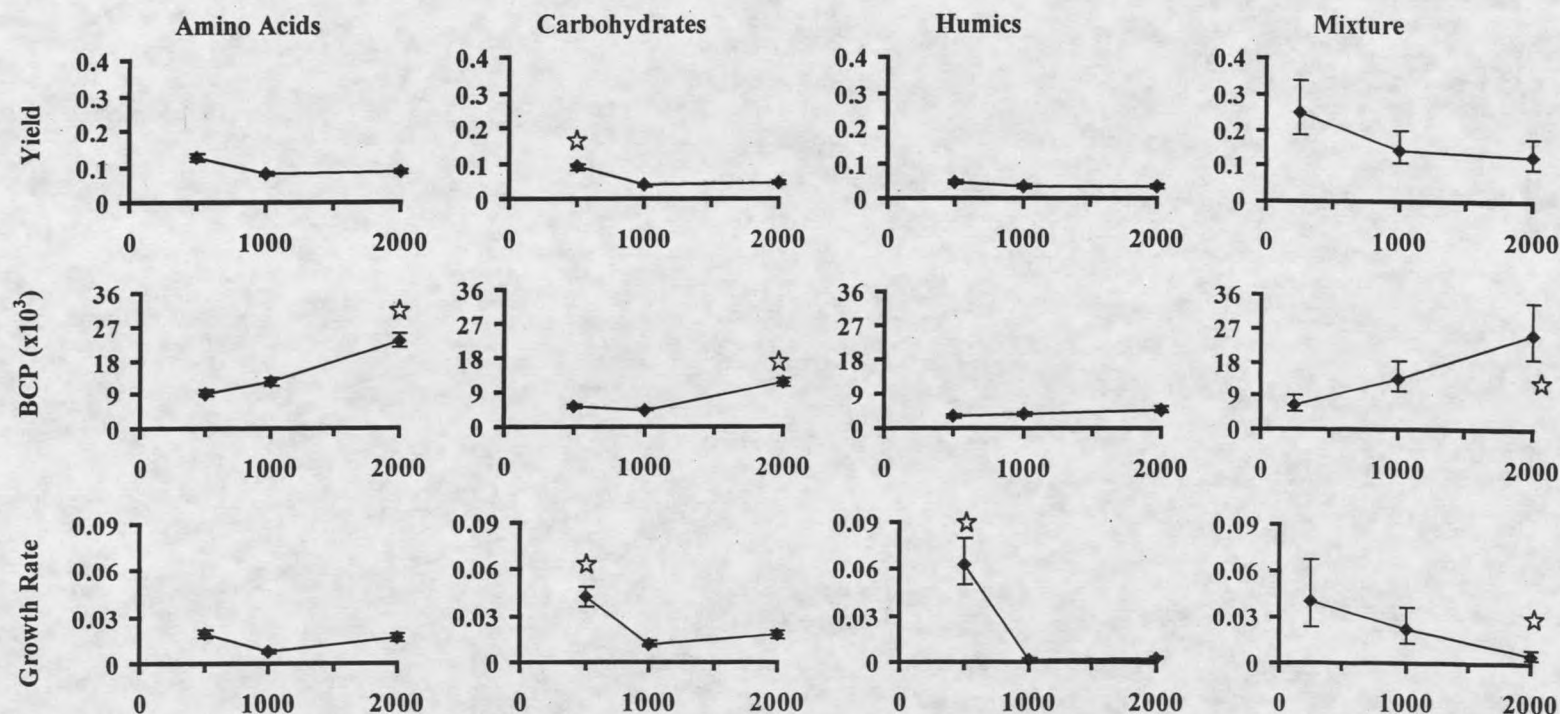


Figure 4.4 – Line plots showing observed yield ($\mu\text{g cell C}/\mu\text{g C}$), bacterial carbon production (BCP, $\mu\text{g C}/\text{cm}^2\cdot\text{hr} \times 10^3$), and specific growth rate (SGR, hr^{-1}) vs. nominal carbon feed concentration with respect to the separate substrates for the 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.

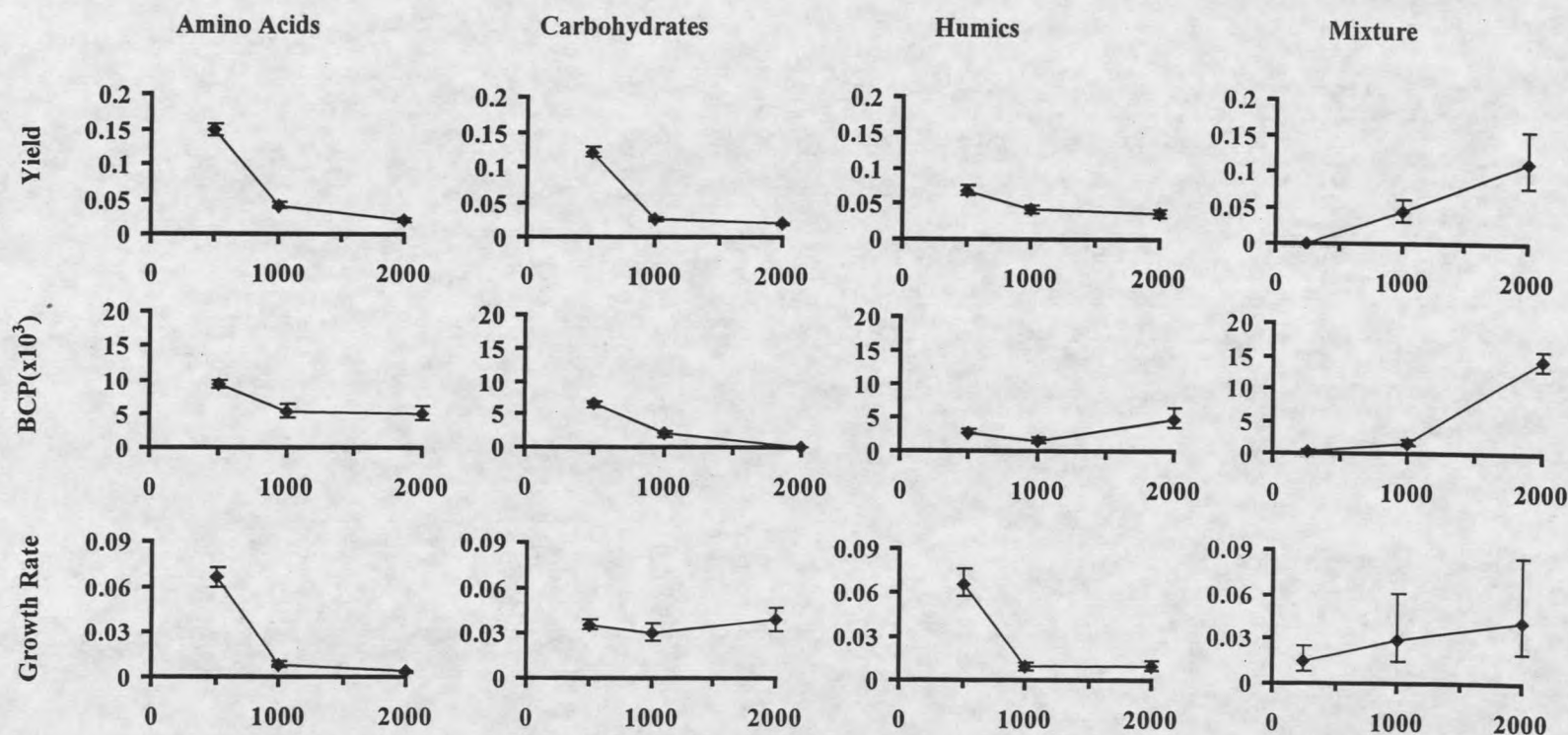


Figure 4.5 - Line plots showing observed yield ($\mu\text{g cell C}/\mu\text{g C}$), bacterial carbon production (BCP, $\mu\text{g cell C}/\text{cm}^2\cdot\text{hr} \times 10^3$), and specific growth rate (SGR, hr^{-1}) vs. nominal carbon feed concentration with respect to the separate substrates for 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.

Yield and BCP were significantly and negatively impacted by chlorination, with the exception of the carbon mixture experiment. Yields ranged from a high of 0.11 μg cell carbon/ μg substrate carbon in the mixture reactors at the high carbon feed level to a low of 0.019 in the carbohydrate communities. Biomass yield on humics (0.069, 0.042, and 0.036 at 500, 1000, and 2000 ppb C, respectively), amino acids (yield = 0.147, 0.041, 0.020), and carbohydrates (0.121, 0.025, and 0.019) all decreased with increasing carbon concentrations. The mixture was the only substrate which showed an increase in cell yield (0.0, 0.044, and 0.111) with increasing carbon concentration.

In the control biofilms, yield also tended to decrease with increasing carbon concentration. A high of 0.25 (μg cell carbon/ μg substrate carbon) was measured for the mixture biofilm at the low feed level (250 ppb C) and a low of 0.034 was measured for the humics biofilm at the 1000 ppb C feed level. While the cell yield on humics was fairly steady between the three carbon feed levels (0.045, 0.034, and 0.035 at 500, 1000, and 2000 ppb C, respectively) the amino acids (yield = 0.124, 0.080, 0.084), carbohydrates (0.089, 0.038, and 0.046), and mixture (0.250, 0.142, and 0.126) all showed a significant decrease in yield at the higher substrate feed levels compared to the low carbon feed level.

BCP tended to increase with increasing carbon concentration ($r^2 = 0.475$, $p = 0.019$) in the control reactors. In terms of $\mu\text{g C}/\text{cm}^2\cdot\text{hr}$, BCP for the amino acids (0.009, 0.012, and 0.023 at 500, 1000, and 2000 ppb C, respectively), carbohydrates (0.005, 0.004, and 0.012), humics (0.003, 0.004, and 0.005), and mixture (0.007, 0.014, and 0.025) all showed this increasing trend. In contrast, there was no correlation between

carbon concentration and BCP for the chlorinated reactors. The mixture (BCP = 0.0004, 0.0016, and 0.014) and the humics (0.0028, 0.0016, and 0.0048) showed an increase in production rates while the amino acids (0.009, 0.005, 0.005) and the carbohydrates (0.007, 0.002, and 3.79×10^{-5}) showed a decrease in production rates following chlorination. Biofilm biomass was positively correlated ($r^2 = 0.619$, $p = 0.011$) with BCP in the chlorinated reactors whereas no correlation was detected for the control reactors.

All three parameters measured in the negative control reactor were significantly lower than the supplemented reactors. BCP was 8.4×10^{-4} $\mu\text{g cell carbon}/\text{cm}^2 \cdot \text{hr}$ ($\pm 6.4 \times 10^{-4}$), SGR was 0.0034 hr^{-1} (± 0.0032), and the observed yield was $0.003 \mu\text{g cell carbon}/\mu\text{g carbon}$ (± 0.003).

Carbon Source Utilization by Biolog Plates

There are a number of methods used to analyze the data arising from the measurement of single substrate oxidation as occurs in BIOLOG microplates. Originally developed for the identification of isolated organisms based on the substrate utilization profile, they have been adopted for differentiating communities through the use of multivariate statistical techniques such as principal component analysis (PCA) or cluster analysis (Klingler, et al., 1992; Hitzl, et al., 1997; Konopka, et al., 1998; Garland and Mills, 1991; Ellis, et al., 1995; Haack, et al., 1995; Vahjen, et al., 1995). We intended to use the microplates as a method of differentiating the biofilm populations from the test and control reactors at the biofilm sampling stages of 1000 and 2000 ppb C and between substrates. PCA did effectively separate the chlorinated from control communities, mid

level from high level, as well as the communities grown on the different substrates. Due to the large amount of data generated with these plates, the fact that PCA can detect differences in utilization is not surprising and generally the first two principal components from the analysis were effective in differentiating the communities (approximately 75 –85% of the variance could be explained in this manner). Addition of the third principal axis added another 5 – 10% to this value and resulted in a very good separation of the communities in three dimensions (data not shown).

Since we collected the data as a time series, we could compare the oxidation rate of the BIOLOG substrates among the different substrates used to generate the biofilm communities. We were interested to see if the substrates used to grow the biofilms predisposed the organisms to a preference for the equivalent substrate in the BIOLOG plates. The ability of the control communities grown on amino acids or carbohydrates to utilize the converse substrates was not affected by the history of the community as seen in Figure 4.6. The major effect was due to chlorination and this had a significant impact on the ability of the carbohydrate and humic grown communities to utilize the specific carbohydrates or amino acids used in this study. However, the oxidation rates were not significantly different for either set of BIOLOG substrates within the control or chlorinated communities. This effect was also extended to the majority of other substrates present in the BIOLOG plates. The dye reduction rates for the amino acids grown communities were not significantly different between the chlorinated and control for the substrates shown in Figure 4.6, whereas they were for the carbohydrate and humics grown communities. Dye reduction rates were in the range of the specific growth rates

measured in the reactors suggesting that this may be a rather simple method of estimating growth rates.

Given that the Biolog plates used for substrate utilization contained the substrates chosen for evaluation with the exception of humics and the L-epimer of arabinose, we chose to compare the dye reduction rates across the communities arising from the different substrates. Optical density in the individual wells was plotted against time and the rates were calculated from the linear portion of the resulting curve. The rate of dye reduction was assumed to be directly proportional to the rate of substrate oxidation. Comparison between the utilization of the amino acids and carbohydrates in the BIOLOG plates versus the substrate used to grow the biofilms is shown in Figure 4.6. Both the chlorinated and control reactor showed no difference in the reduction rate between amino acids or carbohydrates regardless of the substrates used for biofilm growth. The rate of dye reduction and the number of substrates utilized by the homogenized biofilms was negatively impacted by chlorination. Chlorination had a significant impact ($p < 0.0005$) on the ability of the homogenized biofilm to utilize the substrates present in the BIOLOG plates, whereas there was no significant difference in the control reactors ($p = 0.073$). However, chlorination had no effect on the ability of the amino acid grown cells to utilize either the amino acids or carbohydrates present in the BIOLOG plates.

The carbohydrate and amino acid grown control communities oxidized virtually every substrate present in the BIOLOG plates whereas the humic grown biofilms showed a reduced capacity to utilize many of the substrates. This response was measured as the maximum optical density recorded in the individual wells and represents the potential of the cells to utilize the individual substrates. The substrates in each plate can be divided into classes which include: (1) polymers, (2) sugar alcohols, (3) monosaccharides (both D and L forms), (4) di- and tri-saccharides, (5) carboxyl- and amine- substituted sugars, (6) fatty acids (short chain and substituted), (7) dipeptides and amino acids (2 D-amino acids as well), (8) nucleic acid bases, and (9) various other indicator substrates. These other substrates include phenyl ethylamine, 2-amino ethanol, 2,3-butanediol, glycerol, D,L- α -glycerol phosphate, glucose-1-phosphate, glucose-6-phosphate, and bromosuccinic acid. Based on maximal response over the course of the 4 – 5 day incubations the control reactors had substrate preferences in the order amino acids > carbohydrates > organic acids > polymers = fatty acids. The chlorinated communities followed the same trend except that polymers were definitely preferred over fatty acids. The poorest overall substrate was thymidine with only 2 (average optical density > 0.15) out of 17 sample communities showing an ability to utilize it as a single carbon source.

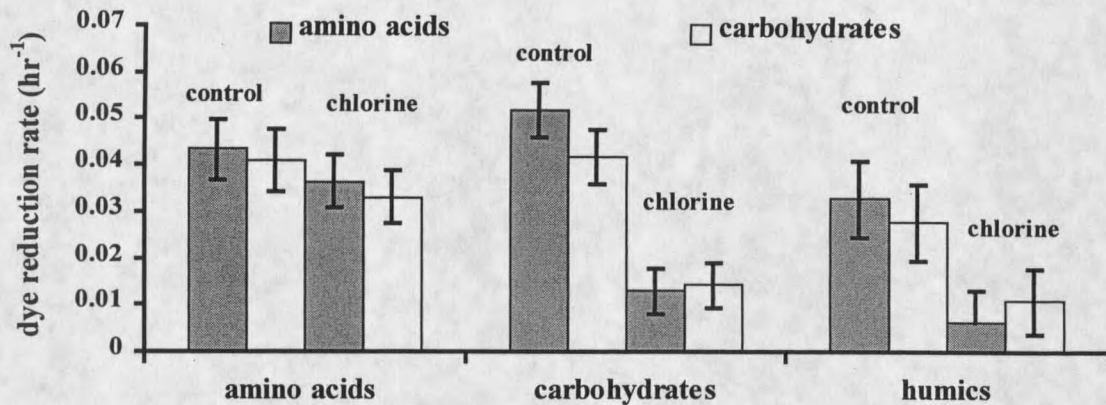


Figure 4.6 – Dye reduction rates averaged over the four amino acids and four carbohydrates used for biofilm development. Y-axis is the substrate source used for biofilm development and the bars are the response of those cells to either the amino acids or carbohydrates present in the BIOLOG-GN microplates. Error bars are standard deviations.

Comparative Analysis of Methods used for Growth Rate Estimations

Table 4.2 shows SGRs compared by method of calculation for the amino acids experiments. Growth rates could be calculated by carbon per unit volume, volume, carbon per cell, bacterial carbon production divided by total biofilm biomass, total cells and active cells. Carbon per unit volume was calculated using the equation from Chapter 2 (weighted non-linear regression) and carbon per cell was calculated from the average carbon/cell value of 42 fg measured in Chapter 2. Active cells were calculated using the active fraction of cells according to microautoradiography from Chapter 3. Additionally, a correction factor could be applied to the carbon per unit volume, volume, and carbon per cell calculations. Consider, for example, the volume measurements. The cells entering the reactor have a different average volume than those in the effluent. Assuming that the influent cells do not attach and are represented as a fraction of the effluent cells

the correction factor for this fraction of cells is different from that for cells arising from within the reactor (i.e. from biofilm detachment). Therefore, the number of influent cells was subtracted from the number of effluent cells, this value was multiplied by the average effluent cell volume (i.e. conversion factor) and added to the influent total cell volume. The active fraction of the cells in the influent and effluent of the reactor were used to correct the growth rates to active cells. An assumption for this calculation is that the active fraction of cells in the biofilm is equivalent to the active fraction of cells in the effluent. While individual methods produce slightly different results, the differences are not significant.

Reactor	Carbon	fgC/ μm^3		μm^3		fgC/cell		cells		BCP
		1	2	1	2	1	2	1	2	3
control	1000	0.024	0.021	0.028	0.030	0.025	0.027	0.026	0.029	0.026
	2000	0.019	0.017	0.023	0.024	0.021	0.022	0.021	0.022	0.021
chlorine	1000	0.021	0.018	0.014	0.016	0.017	0.019	0.017	0.021	0.017
	2000	0.008	0.008	0.010	0.010	0.009	0.012	0.009	0.010	0.009

Table 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 fgC/ μm^3 , μm^3 , and fgC/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 ($\mu\text{g C}/\text{cm}^2$).

Carbon Substrate Removal

Carbon removal across the reactors using filtered non-purgeable organic carbon (NPOC) data was used for determination of the substrate removal rates. When carbon removal flux ($\mu\text{g C used}/\text{cm}^2\cdot\text{hr}$) was plotted against the influent loading rate ($\mu\text{g C}/\text{cm}^2\cdot\text{hr}$) a linear relationship was seen in the data where the slope represents the

fraction of carbon removed across the reactors. Table 4.3 shows the fraction of substrate removed, regression coefficients and 95% confidence intervals for the slopes. Table 4.4 shows the specific removal rate normalized to biofilm biomass. All intercepts from the substrate flux regression were negative (range -0.069 to $-0.0024 \mu\text{g C/cm}^2\cdot\text{hr}$) except for the carbohydrate control reactor ($y = 0.0081 \mu\text{g C/cm}^2\cdot\text{hr}$), indicating that in most cases there was a positive x - intercept or a minimum substrate loading rate which would support a steady state biofilm community. Since this was a Model II regression, the method of geometric means was used for the regression analysis. The linear relationship noted above did not always occur for the normalized carbon removal rate plotted in the same manner. Using a GLM ANOVA, there was a significant difference in normalized carbon removal rates between substrate groups ($p = 0.005$) and between carbon levels within substrate groups ($p < 0.0005$) for the control reactor. In the presence of chlorine, specific removal rates were not significantly different between substrates ($p = 0.123$) but were significantly different between carbon levels within substrates ($p < 0.0005$). Removal rates were calculated for the chlorinated reactor after chlorination had been initiated.

Substrate	Reactor	Slope (n_1, n_2)	r^2	95% C.I.
amino acids	control	0.98 (2, 15)	0.992	0.89 - 1.07
	chlorinated	1.02 (2, 15)	0.981	0.91 - 1.14
carbohydrates	control	0.90 (2, 11)	0.980	0.80 - 1.0
	chlorinated	0.08 (2, 11)	0.932	0.008 - 0.14
humics	control	0.78 (2, 10)	0.952	0.63 - 0.92
	chlorinated	0.58 (2, 10)	0.818	0.37 - 0.79
mixture	control	0.64 (1, 6)	0.991	0.54 - 0.74
	chlorinated	0.50 (1, 6)	0.993	0.43 - 0.57
control	control	0.05 (1, 7)	0.904	0.03 - 0.07

Table 4.3 – Linear regression data for carbon flux ($\mu\text{g Carbon used/cm}^2 \cdot \text{hr}$) as a function of influent substrate loading rate ($\mu\text{g Carbon/cm}^2 \cdot \text{hr}$). 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.

Substrate	Reactor	Mean Norm. Carbon Removal Rate (hr^{-1})	95% C.I. for Mean
Amino Acids	control	0.080	0.075 - 0.083
	chlorinated	0.234	0.204 - 0.269
Carbohydrates	control	0.129	0.121 - 0.137
	chlorinated	0.243	0.198 - 0.299
Humics	control	0.118	0.110 - 0.127
	chlorinated	0.200	0.165 - 0.243
Mixture	control	0.078	0.042 - 0.146
	chlorinated	0.538	0.341 - 0.849
Control	control	0.048	0.029 - 0.081

Table 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.

Discussion

In these experiments, we have focused on the role of biofilm communities in the turnover of the most commonly detected carbon substrates in aquatic systems (Jorgensen and Jensen, 1994; Kaplan and Newbold, 1993; Hanisch, et al., 1996; Paerl, 1991; Volk, et al., 1997). While most studies have focused on planktonic organisms, a few exceptions have attempted to utilize biofilm bacteria in order to determine kinetic constants or substrate utilization profiles (Kaplan and Newbold, 1995; Frias, et al., 1994; Merlet, et al., 1991; Volk, et al., 1997; Kott, et al., 1997). These studies have been performed with BDOC columns or BAC filters; both of which were used in these experiments. The BDOC columns were used to measure carbon removal, growth rates, yields, and secondary carbon production from the communities maintained on biologically filtered drinking water. Drinking water has been the focus of most of these studies involved with bacterial biofilms due to the potential of detrimental growth within the distribution system as well as the positive aspects of biofilm formation in BAC filters. Regrowth within the distribution system is a concern due to the potential health hazards, deterioration of the waters aesthetic qualities (e.g. taste, odor, and color), and the persistence of indicator bacteria (e.g. coliforms) within the biofilm. Conversely, for pre-treatment of drinking water, biofilm growth in BAC filters is encouraged for the following reasons: (1) decreased chlorine demand in finished waters, (2) decreased regrowth potential due to BDOC removal, and (3) decreased disinfection by-product formation. In addition to these three positive aspects, it is advantageous to encourage

bacterial growth at a point where we can exercise some operational control and minimize it in the distribution system where the only practical control we have is increasing disinfection levels. Chlorination has not been a factor in any of these studies and since chlorine tends to be the disinfectant of choice, especially in North America, no attempts have been made to measure bacterial growth related parameters in the presence of chlorine.

Control Stage Carbon Level

Analysis of the negative control reactor data for total and culturable cells as well as DOC removal showed that the AOC "free" dilution water was free of utilizable carbon. This indicates that the GAC/BAC filter was very effective in removing utilizable carbon in the finished drinking water and that the responses measured in the supplemented control reactors were due to the added carbon substrates. The negative control reactor had a net removal of carbon not significantly different from zero while the BDOC columns removed an extra $32 \pm 38 \mu\text{g C/L}$ over and above what was removed by the GAC/BAC filter. Expressed as a rate, this was equivalent to $15.2 \mu\text{g C/L}\cdot\text{hr}$, or approximately 2% of the removal flux for the GAC/BAC filters (average was $760 \mu\text{g C/L}\cdot\text{hr}$). As a percentage of the excess substrate supplied to the reactors this ranged from 0-14% at the lower carbon loads (500 ppb C). GAC/BAC filtration removed an average of $18 \pm 4.7 \%$ of the influent TOC, typical for percentage removals in raw freshwater and marine systems (Kaplan and Newbold, 1995; Amon and Benner, 1996; Volk, et al., 1997; Tranvik, 1988). GAC/BAC filtration is usually performed much earlier in the treatment

process, and not in the finished water as was the case for this study. This likely indicates the formation of a bacterial community on the filters that is uniquely suited to the degradation of the carbon in the source water, which is not surprising when one considers that this filter had been operating for approximately 3 years prior to the initiation of this study and 7 years by the end of the study.

Based on the results from the control stage of the experiments the substrate group that provided the most favorable growth conditions was the amino acids. Carbohydrates supported a total effluent population of cells that was 80% of the amino acids whereas humics maintained a population that was only 47% of the amino acids. The same trend was seen for the culturable cells where the carbohydrates and humics supported only 75% and 32% of the amino acids cell density, respectively. Biofilm total cell density was similar with the carbohydrates supporting 75% and the humics supporting 15% of the amino acid population. The only parameter that did not follow this trend was the culturable biofilm cells from the carbohydrates experiments. Culturable were 27% greater in the carbohydrate biofilms than in the amino acid biofilms (humics support 5% of the amino acid cell density). Steady state effluent planktonic cell densities in all experiments were within the limits of those found in natural freshwater systems (Tranvik, 1994; Pollard and Greenfield, 1997; Rheinheimer and Gocke, 1994). This suggests that carbon loading levels were approximating those found in other aquatic environments, particularly in the case of the humics and mixture experiments. The unexplained error component comprised <5.0% of the total and culturable effluent cell data variance indicating that the error due to sampling was low. Variance components at the level of

experimental replication were approximately 10% of the total variance, indicating good replication. Without the high n-value achieved with this experimental protocol, it was unlikely that differences would have been detected between the population densities supported by the different substrates. At a power of 0.99, a post-hoc analysis based on the means ($\pm$ S.D.) measured in this study indicated that an n value of 13 was required. Reproducibility was good between experiments within substrates and between reactors within substrates for the entire study.

Test Stage Carbon Levels

Once chlorination began, the differences in the substrates were magnified to a larger degree than for the control stage of the experiments. Overall the amino acids were the best substrate for the control reactors. The amino acids produced higher culturable cell counts than the amount of total cells arising from the humic experiments. All of the control reactors exhibited an increase in effluent total cell counts from the lower to higher carbon loading levels, with only the humics increase being insignificant. The chlorinated communities follow a different trend, with total cells being equivalent across carbon levels for the amino acids and carbohydrates while the significant decrease seen in the carbohydrate culturable cells was not observed in the amino acid experiments. The humics were the only substrate that showed a significant increase in effluent total cells and culturable cells after chlorination was initiated (Figure 4.3).

While total and culturable cells were significantly correlated in all cases, the correlation was particularly weak for both the influent cells and the humic effluent cells. Culturability (i.e. cfu/tdc) was significantly impacted by chlorination for all of the

substrates tested, especially in the presence of humics and carbohydrates. This implies that an attempt to use culturable cells as an indicator of activity in the reactors would produce unreliable results with respect to yield, BCP, and SGRs. The range of culturability (5 – 50%) is considerably above the range of literature values in natural environments (0.01 – 10%). This could be due to the presumed lower community diversity in drinking water systems as compared to most other environments. Drinking water bacterial isolates tend to be Proteobacteria, most of which are able to grow on agar plates under laboratory conditions (Camper, et al., 1986; Kalmbach, et al., 1997; Shirey and Bissonnette, 1991; Manz, et al., 1993; Mikell, et al., 1996).

The biofilm sampling data does not show the same trend as that seen for the effluent data. All substrates produced significantly higher counts in control biofilm culturable cells when going from low to high level carbon loading, but as a fraction of the total biofilm cells, culturability decreased at the high carbon loading levels to 25-50% of the culturability seen at the low carbon level. Total cell counts indicated that there was an overall increase in the number of cells with increasing carbon load for all substrates. At the middle carbon level the humics were significantly higher than the carbohydrates and amino acids for total biofilm cells (control) whereas at the higher carbon level the humics were intermediate to the amino acids (high) and carbohydrates (low). For the chlorinated reactor, amino acids were the better substrate for biofilm cells at both carbon levels while the humics were significantly higher than the carbohydrates. For the culturable biofilm (and effluent) cells, amino acids provided some protection from the effects of chlorination since there was no difference between the number of cells that grew on R2A

agar between the chlorinated and control reactors at both of the chlorinated carbon levels. This is also apparent when considering the percentage of biofilm (and effluent) culturable cells with respect to total biofilm (and effluent) cells where the amino acids show 10% (21% in the effluent) culturability while the humics and carbohydrates showed 0.9% (1.5%) and 0.07% (5.4%) at the highest carbon level. Overall, the only difference in culturable biofilm cells for the control reactors was between the mid-level amino acids and high-level humics, which suggest that a poor relationship exists between culturable cells and potential activity in the biofilms.

While the amino acids and carbohydrates were almost totally utilized in the control reactors (98 and 90% respectively), humics removal was 78% of the influent load. Still this value is 2 to 10 times published literature values, which range from 1 to 45% (Volk, et al., 1997; Tranvik, 1988; Tranvik, 1993; Tranvik and Hofle, 1987; Tranvik and Kokalj, 1998; Amon and Benner, 1994). While most of these studies were performed in planktonic batch cultures, the values closest to ours, 25% from Volk et al. (1997) and 45% from Merlet et al. (1991) are from studies which used biofilm bacteria. We performed batch culture studies with the bacteria from the endpoint humics experiments (data not shown). Results from these experiments indicated that when the biofilm communities were disrupted by scraping and homogenizing, the now planktonic cells could only utilize $24 \pm 10\%$ ($n = 10$) of the total humic substrate over a 10 day incubation period. This is in contrast to the steady state biofilm removal of 78% for the control and 58% for the chlorinated biofilms during 2.1 hours as measured in the reactors.

It is not surprising that amino acids provided for highest population densities in the biofilms. Selection of the amino acids for this study was based on their environmental availability and relationship to central metabolism. Energetically, amino acids provide an easily utilizable carbon, energy, and nitrogen source, whereas carbohydrates only satisfy carbon and energy requirements. In this case the communities are forced to assimilate nitrate as the nitrogen source. The need to derepress nitrate assimilatory pathways and the energetic cost required for the reduction of nitrate to ammonia is high compared to the assimilation of glutamate in the amino acids mixture. Glutamate itself supplies the α -NH₂ group to amino acid anabolism as well as providing the nitrogen source for purine and pyrimidine biosynthesis. The product arising from transamination with glutamate is α -ketoglutarate which can directly enter the TCA cycle, providing an energetic and anapleurotic role. Serine and alanine feed into the Embden-Meyeroff-Parnas pathway at the level of pyruvate and can be used in a variety of anapleurotic pathways. Aspartate plays a key role in amino acid metabolism, acting as a precursor for lysine, methionine, threonine, isoleucine, asparagine, and to a minor extent alanine. Conversely, the carbohydrate communities would be expected to expend a significant amount of energy in both the assimilation of nitrate and the anabolic reactions required for the synthesis of all precursor metabolites such as amino acids and nucleotides.

The humics were a highly heterogeneous source of carbon, nitrogen, and trace elements. Trace level concentrations of a variety of amino acids and carbohydrates bound within the polyaromatic backbone of the humic structure would be expected to provide a relatively diverse nutritional supply, contingent on the ability of the communities to gain

access to the bound nutrients. In this regard, the function of the biofilm could aid to a considerable degree. Chlorination provided for some level of humic oxidation, likely leading to a consequent release of nutrients, while the biofilm itself certainly confers an increase in degradative capacity (Wolfaardt, et al., 1994; Wolfaardt, et al., 1994; Manem and Rittmann, 1992). A variety of explanations are plausible for this increased degradative capability. The increased surface area to volume ratio in the biofilm provides for a larger reactive surface area. Through forming biofilms, cells can alter their local environment to allow for more favorable growth conditions and potential interactions between species. Sequestration of the humic material within the biofilm could result in enhanced degradative capacity, such that the organisms are capable of utilizing a larger fraction of the influent material. However the increased degradative capability arises, the consequence is that the BDOC fraction in water is potentially much higher than previously thought. Utilization of the high molecular weight fraction has been shown for marine, streamwater and drinking water bacteria (Volk, et al., 1997; Amon and Benner, 1994; Amon and Benner, 1996; LeChevallier, et al., 1992).

This study and others have shown that biofilm communities have a much higher capacity to extract energy and carbon from recalcitrant sources. Whether one is interested in the regrowth potential (AOC) of drinking water or the removal of chlorine demand and the potential formation of disinfection by-products (BDOC), utilization of suspended planktonic organisms to derive results for systems where most of the activity occurs in biofilms does not provide realistic or reliable results.

Growth Rates, Production, and Yield

Most investigators would accept that treated drinking water is an oligotrophic aquatic environment. However, operationally, oligotrophic systems have generally been accepted to be those in which the rate of food supply is very low and when carbon is limiting, in the range of fractions of milligrams C/L•day (Andrews and Harris, 1986; Schut, et al., 1997). From this viewpoint, the dilution water entered the reactors at a maximum carbon loading rate of 1.0 mg C/L•day, whereas the supplemented reactors have carbon loading rates of approximately 6, 12, and 24 mg C/L•day. These values are comparable to the mean BDOC loading rate entering (i.e. finished drinking water) the GAC/BAC filters, which was 18 mg C/L•day. This high carbon loading rate entering the GAC/BAC filters is due to the short residence time (15 minutes). By the definition given above it would not be considered oligotrophic. Given the values above, the supplemented reactors were not receiving an amount of carbon that was excessively above what was in the drinking water itself. The quality of the substrate was better in the supplemented amino acids and carbohydrates experiments but likely of similar quality for the humics and mixture experiments. Since steady-state within the reactors was indicated by the previously mentioned measured parameters, reactor operation was at pseudo steady-state prior to increasing the carbon concentrations.

The concept of r - and K - selection theory in relation to microbial ecology can explain several of the findings with respect to SGRs, production, and observed yields. Andrews and Harris (1986) introduced the concept of r - and K - selection theory to microbial ecology in 1986. Within the context of r/K selection theory, population density

is of paramount importance with respect to the environmental niche that an organism occupies. Maximization of survival fitness occurs in response to either uncrowded (r – selection) or crowded (K – selection) environments. K refers to the density of individuals that a given environment can support while r refers to the maximum specific rate of increase of an organism. Thus, either the organism or the environment can be defined in terms of r/K selection theory. r -environments tend to be transitory and arise due to increased food supply, a severe reduction in the population density, and/or a shift in environmental conditions. A K – environment exists when the population is in equilibrium with its surroundings such that specific rates of increase are close to zero and food supply is maximally utilized.

Initiation of our experiments results in an r -environment situation due to the conditions in the reactors. The microorganisms, which colonize the annular reactors, originate from the oligotrophic and uncrowded environment (mid 10^5 total cells/ml) of the GAC/BAC filters. As soon as they enter the reactors, they are faced with a relative surfeit of carbon and colonize the reactors at a low population density. Substrate concentrations are relatively low and in the range of 2.1 – 3.5 μM for amino acids, 1.7 – 2.1 μM for carbohydrates, and approximately 25 μM C for humics. These concentrations are insufficient to support high cell densities and low cell densities tend to allow for higher growth rates given a constant supply of food. Once steady-state is reached, cell densities are still relatively low at the lowest feed concentration as compared to the higher cell densities found in the reactors at the higher feed concentrations (see Figures

4.1 and 4.2). This paradigm of population density, crowding, and rate of food supply can be used to explain the trend in growth rates seen in our experiments, as described below.

Control Communities. Specific growth rates (SGRs) for all substrates were significantly higher at the low carbon feed level as compared to the higher feed levels. The correlation between biofilm biomass and SGR was significantly negative for both the control and chlorinated communities ($r^2 = -0.637$, $p = 0.001$ and $r^2 = -0.570$, $p = 0.021$, respectively). There was no correlation between SGR and substrate concentration for either of the communities. In addition to r/K - selection theory predicting this consequence, a number of other investigators have found the same correlation for planktonic bacteria, where a higher standing stock of biomass is significantly correlated with a lower SGR (Morris and Lewis, 1992; Moriarty and Bell, 1993; Andrews and Harris, 1986; Pollard and Greenfield, 1997). SGR on carbohydrates and amino acids were equivalent at the higher carbon loading levels and both were significantly higher than for the humic grown communities. While the higher substrate concentrations support higher cell densities, SGR could be limited due to the lower per capita food supply. Table 4.4 shows that utilization rates normalized to biofilm biomass are 1.7 – 7.0 times lower for the control reactors as compared to the chlorinated reactors. This indicates that the amount of substrate per unit biomass is significantly lower in the control, compared to the chlorinated communities. Biofilm cell density increases by approximately an order of magnitude (see Figures 4.1 and 4.2) or more for the control reactors while carbon supply was only doubled. The control reactors removed most of the amino acids and

carbohydrates and a significant portion of the humics in order to support higher population densities growing at lower rates. Doubling times at the highest carbon feed level for the control reactors were 8.3 days for the mixture, 2.3 days for the carbohydrates, 15.4 days for the humics, and 2.5 days for the amino acids. As a comparison, the negative control reactor had a doubling time of 12.3 days. This approximates the value obtained by Van Der Wende, et al. (1989) (16.7 days) using annular reactors in unamended drinking water.

Yield decreased in the following order amino acids ($0.08 - 0.12 \mu\text{g cell C}/\mu\text{g C}$) > carbohydrates ($0.038 - 0.089$) > humics ($0.034 - 0.045$) with the highest yield values obtained at the lowest feed concentration. Decreased per capita food supply in conjunction with high cell density could have led to increased maintenance requirements and a concomitant decrease in the efficiency with which the individual organisms converted substrate to biomass. Higher biomass yield due to utilization of small molecules has been shown by Amon and Benner (1996) and Schweitzer and Simon (1995). Although the yield measured in these experiments tended to be at the low end, they are in the range of values reported in the literature (Tranvik and Hofle, 1987; Button, 1985; Hanisch, et al., 1996). Low cell yields would be preferred in biological filters, since the organisms arising from the filter would have to be removed prior to releasing the water into the distribution system. Therefore, encouraging high biomass in a biological filter would tend to decrease cell yield due to the rather inefficient conversion of substrate to biomass (however, see discussion on BCP below). The mixture provided for the highest yield, likely due to the diversity of nutrients present, but again the highest

yield (range, 0.13 – 0.25) was at the lowest carbon feed level. Yield was significantly positively correlated with BCP ($r^2 = 0.555$, $p = 0.005$).

BCP in the control communities tended to increase as carbon loading increased, although not significantly in the case of the humics. While the reported values are normalized to surface area, multiplying by the surface area to volume ratio of the annular reactors ($1800 \text{ cm}^2/0.630 \text{ L}$) and converting hours to days, results in production rates that can be compared to other environments. Sommaruga (1995) reported mean BCP rates of $587 \text{ } \mu\text{g C/L}\cdot\text{day}$ (maximum = 1071) in a hypertrophic Venezuelan lake, similar to values arising from the annular reactors in these experiments. BCP rates in $\mu\text{g C/L}\cdot\text{day}$ ranging from 635 – 1585 for amino acids, 295 – 803 for carbohydrates, 209 – 338 for humics, and 456 – 1738 for the mixture were measured in the control reactors. More typical BCP rates in natural planktonic communities range from 0 – $100 \text{ } \mu\text{g C/L}\cdot\text{day}$ (Fuhrman and Azam, 1980; Hanisch, et al., 1996; Simon and Azam, 1989).

A number of studies on BCP rates on epiphytic surfaces are within the range of our BCP estimates when normalized to surface area (see Results section and Figures 4.4 and 4.5). Thomaz and Wetzel (1995) measured rates in the range of $0.93 - 1.85 \text{ } \mu\text{g C/cm}^2\cdot\text{day}$ (at 20°C) on the surface of the aquatic macrophyte *Juncus effusus* detritus by ^3H -leucine incorporation rates. They showed that BCP increased with increasing biomass whereas growth rate decreased. They also estimated bacterial turnover times of 4.8 – 15.2 days for their samples, very similar to the turnover times measured in the annular reactors.

While the SGR and yields decreased in our experiments, increased BCP was likely due to the higher biomass density present at the higher feed concentrations. Since the growth rates were independent of the external substrate concentration, the biofilm communities were not growing according to the conventional Monod kinetic model used for free-living bacteria. These results are similar to those found by Moller, et al. (1995) who used a direct microscopic image analysis technique to determine that growth rates of attached *Pseudomonas putida* in chemostat studies were independent of the dilution rate. This has implications for the modeling of bacterial growth in drinking water distribution systems. Most of the models assume that growth is first order (i.e. Monod kinetics) in the biofilm, while this study indicates that this is not correct. This also implies that knowledge of the growth rates of the organisms in a drinking water distribution system or biological filter is unpredictable and requires direct measurements as were done in this study.

Chlorinated Communities. With the exception of the amino acids, SGRs in the presence of chlorine were always higher than the SGRs in the parallel control reactor. The correlation between biomass and SGR was significant and negative (see above). SGRs were not significantly different between carbon levels after initiation of chlorination. The chlorinated amino acids biofilm communities were seemingly unaffected as biofilm biomass was not significantly impacted and substrate utilization was the same as for the control reactors. Accordingly, one would expect the growth rates to be equivalent for the chlorinated and control communities. However, the SGR in the chlorinated reactor was significantly lower than that of the control reactor. This is likely

due to the small but significant impact of chlorination on the effluent cells, whereby the effluent cell concentration in the chlorinated communities is 63% and 35% of the control effluent population at the mid and high carbon loading stage respectively. SGRs were equivalent at the mid-level feed concentration, but the chlorinated reactor SGR was 20% of the control reactor at the high carbon loading stage.

Another possible explanation is an interaction between the dual stress of nutrient limitation and chlorination. Biofilm and effluent cell concentrations for the carbohydrates and humics are significantly impacted by chlorination (Figures 4.2 and 4.3). The substrate is relatively more plentiful (see Table 4.3 and 4.4), which results in higher SGRs (Figures 4.4 and 4.5). The carbohydrate-fed reactors, with the highest concentration of substrate in the effluent as well as the lowest biomass in both the effluent and biofilm showed the highest SGRs under chlorination with the exception of the mixture reactor at the highest carbon feed level. Doubling times at the highest carbon feed level for the chlorinated reactors were 1 day for the mixture, 1.1 days for the carbohydrates, 4.3 days for the humics, and 12.1 days for the amino acids.

Observed yields in the chlorinated communities were negatively correlated with SGR ($r^2 = -0.525$, $p = 0.037$) and positively correlated with biomass ($r^2 = 0.566$, $p = 0.022$) whereas no correlations were detected for the control reactors. Thus, as SGR increased and biomass decreased the communities became progressively less efficient at converting substrate into increased amounts of biomass. This could be due to the injury caused by the oxidative stress from chlorination and the need to increase the amount of substrate used for repair and maintenance vs. growth. The trend seen for the control

reactors where yield decreased as carbon loading increased was also seen for the chlorinated reactors. Cell yield for the humics was equivalent in the chlorinated reactor and the control reactor at the high carbon loading level. Given that both of the populations are stressed (one by chlorination and the other by nutrient limitation) maintenance energy requirements would have a larger affect on the more slowly growing communities, potentially resulting in a lower yield in the nutrient stressed organisms as compared to the chlorine stressed organisms (Russell and Cook, 1995).

Increased yield could also be due to an interaction between chlorine and humics, whereby humic oxidation results in a subsequent release of more readily utilizable substrates and a change in the character of the humic material. This phenomenon has been shown to occur when water is ozonated or chlorinated (Hanna, et al., 1991; Goel, et al., 1995; Huck, et al., 1991; Huck, et al., 1994). A highly significant and positive correlation between yield and BCP was noted for the chlorinated reactors ($r^2 = 0.899$, $p < 0.0005$).

BCP in the chlorinated reactors was also correlated with biomass ($r^2 = 0.619$, $p = 0.011$). BCP was 366 and 347 $\mu\text{g C/L}\cdot\text{day}$ for amino acids, 136 and 3 for the carbohydrates, 110 and 326 for the humics, and 113 and 969 $\mu\text{g C/L}\cdot\text{day}$ for the mixture at the two carbon feed levels of 1000 ppb and 2000 ppb, respectively. The high SGR in the chlorinated carbohydrate community was insufficient to offset the low biomass and yield such that BCP was the lowest measured in these experiments. The next lowest BCP value of 58 $\mu\text{g C/L}\cdot\text{day}$ was measured in the negative control reactor. This suggests that biological filtration is as effective as chlorination in controlling bacterial growth and

production. BCP in the humics-grown communities was low at the 1000 ppb C feed level and increased to a value not significantly different from the control reactor BCP at the 2000 ppb C level. This could be attributed to the higher SGRs in the chlorinated humics (3.6 times the control) communities since the yields were similar (0.036-chlorinated and 0.035-control) but biomass was lower in the chlorinated system. The mixture was the only substrate that showed a progressive increase in BCP for both the control and chlorinated reactors. A combination of high yield ($0.13 \mu\text{g cell C}/\mu\text{g C}$), high biofilm biomass ($4.95 \times 10^6 \text{ cells}/\text{cm}^2$), and high growth rate (doubling time = 1 day) in the chlorinated reactor was insufficient to offset the higher biomass ($1.11 \times 10^8 \text{ cells}/\text{cm}^2$), equivalent yield ($0.11 \mu\text{g cell C}/\mu\text{g C}$) and low growth rate (doubling time 8.3 days) of the control reactor. BCP under chlorination was 56% of that measured in the control reactor.

Comparison of Growth Rates Measured by Different Methods

In these experiments, we had a unique opportunity to measure growth rates by a number of different methods (see Table 4.2). All of the methods resulted in growth rates that were not significantly different from each other. Friedrich, et al. (1999) corrected growth rates according to microautoradiography measurements. Free (doubling time, range 4.42 – 26.88 days) and attached (0.15 – 1.8 days) aquatic communities had lower doubling times when growth rates were corrected according to activity. When we corrected the growth rates for activity, we found no difference compared to the other methods. This is likely due to the model used for measuring growth rates in the reactors.

Since we have to account for the activity of the organisms in the influent, effluent and biofilm, the correction factors applied cancelled each other out. In addition, we assumed that the active portion of biofilm cells was equivalent to the active portion of cells in the effluent which may or may not be true (Sommaruga, 1995; Manz, et al., 1993; Kalmbach, et al., 1997; van Loosdrecht, et al., 1990; Caldwell and Lawrence, 1986; Fletcher, 1986). If the actual activity in the biofilm was higher than our estimate then the actual growth rates would be lower than measured whereas an overestimation of active biomass would underestimate the growth rates.

Conclusions

- 1) Replication was reproducible between (variance of ca. 10%) and within (<5%) experiments in the annular reactors when sufficient time for steady state was allowed between carbon loading levels.
- 2) Growth rates were independent of bulk fluid substrate concentration (zero order kinetics) and were positively correlated with biofilm biomass.
- 3) Substrate preference in the control communities was amino acids > carbohydrates > humics, when considering only the cell densities in the effluent and biofilm.
- 4) With respect to cell densities, free chlorine residuals in the range of 0.15- 0.20 mg chlorine/L had a relatively insignificant impact in the presence of amino acids, a significant and negative impact in the presence of carbohydrates, and a relatively positive impact in the presence of humic substances.

- 5) Specific growth rates and observed yields decreased with increasing biofilm biomass. Accordingly, the highest growth rates and yields were observed at the lowest feed concentrations for all substrates tested.
- 6) Chlorination had a positive impact on the specific growth rate when compared to control communities at the same carbon loading rates.
- 7) Humic utilization in chlorinated and control biofilm communities was significantly higher than previously reported values for both planktonic and biofilm systems.
- 8) Bacterial carbon production rates were equivalent to or exceeded the highest values previously reported for natural environments.
- 9) BIOLOG plates may be a simple method for estimating bacterial growth rates by measuring the dye reduction rates.
- 10) Considering SGRs, BCP, and yield, biological filtration of the drinking water was as effective as (or better than) chlorination in controlling growth within the annular reactors. This can be seen by comparing the negative control reactor with the chlorinated reactors.

References Cited

- Amon, R.M.W., and Benner, R. (1994). Rapid cycling of high-molecular-weight dissolved organic matter in the ocean. *Nature*, 369:549-551.
- Amon, R.M.W., and Benner, R. (1996). Bacterial utilization of different size classes of dissolved organic matter. *Limnol. Oceanogr.*, 41:41-51.
- Andrews, J.H., & Harris, R.F. (1986). r- and K - selection and microbial ecology. In K.C. Marshall (Ed.), *Advances in Microbial Ecology*, vol. 9. (pp. 99-147). New York: Plenum Press.
- Bradford, S.M., Palmer, C.J., and Olson, B.H. (1994). Assimilable organic carbon concentrations in southern California surface and groundwater. *Wat. Res.*, 28(2):427-435.
- Bryers, J.D. (1987). Biologically active surfaces: processes governing the formation and persistence of biofilms. *Biotech. Prog.*, 3(2):57-68.
- Button, D.K. (1985). Kinetics of nutrient-limited transport and microbial growth. *Microbiol. Rev.*, 49:270-297.
- Caldwell, D.E., and Lawrence, J.R. (1986). Growth kinetics of *Pseudomonas fluorescens* microcolonies within the hydrodynamic boundary layers of surface environments. *Microb. Ecol.*, 23:99-112.
- Camper, A.K., LeChevallier, M.W., Broadaway, S.C., and McFeters, G.A. (1986). Bacteria associated with granular activated carbon particles in drinking water. *Appl. Environ. Microbiol.*, 52(3):434-438.
- Chrzanowski, T.H., Sterner, r.W., and Elser, J.J. (1995). Nutrient enrichment and nutrient regeneration stimulate bacterioplankton Growth. *Microb. Ecol.*, 29:221-230.
- Coveney, M.F., and Wetzel, R.G. (1992). Effects of nutrients on specific growth rate of bacterioplankton in oligotrophic lake water cultures. *Appl. Environ. Microbiol.*, 58:150-156.
- Currie, D.J. (1990). Large-scale variability and interactions among phytoplankton, bacterioplankton, and phosphorous. *Limnol. Oceanogr.*, 35:1437-1455.
- Dunn, O.J., & Clark, V.A. (1974). *Applied statistics: Analysis of Variance and Regression*. New York: John Wiley and Sons.

- Ellis, R.J., Thompson, I.P., and Bailey, M.J. (1995). Metabolic profiling as a means of characterizing plant-associated microbial communities. *FEMS Microbiol. Ecol.*, 16:9-18.
- Fletcher, M. (1986). Measurement of glucose utilization by *Pseudomonas fluorescens* that are free-living and that are attached to surfaces. *Appl. Environ. Microbiol.*, 52:672-676.
- Frias, J., Ribas, F., and Lucena, F. (1994). Substrate affinity from bacterial strains and distribution water biofilms. *J. Appl. Microbiol.*, 76:182-189.
- Friedrich, U., Schallenberg, M., and Holliger, C. (1999). Pelagic bacteria-particle interactions and community-specific growth rates in four lakes along a trophic gradient. *Microb. Ecol.*, 37:49-61.
- Fuhrman, J.A., and Azam, F. (1980). Bacterioplankton secondary production estimates for coastal waters of British Columbia, Antarctica, and California. *Appl. Environ. Microbiol.*, 39:1085-1085.
- Garland, J.L., and Mills, A.L. (1991). Classification and characterization of heterotrophic microbial communities on the basis of patterns of community-level sole-carbon-source utilization. *Appl. Environ. Microbiol.*, 57(8):2351-2359.
- Glaze, W.H., Andelman, J.B., Bull, R.J., Conolly, R.B., Hertz, C.D., Hood, R.D., and Pegram, R.A. (1993). Determining health risks associated with disinfectants and disinfection by-products: Research needs. *J. AWWA, March*:53-56.
- Goel, S., Hozalski, R.M., and Bouwer, E.J. (1995). Biodegradation of NOM: Effect of NOM source and ozone dose. *J. AWWA, April*:90-105.
- Gremm, T.J., and Kaplan, L.A. (1997). Dissolved carbohydrates in streamwater determined by HPLC and pulsed amperometric detection. *Limnol. Oceanogr.*, 42:385-393.
- Haack, S.K., Garchow, H., Klug, M.J., and Forney, L.J. (1995). Analysis of factors affecting the accuracy, reproducibility, and interpretation of microbial community carbon source utilization patterns. *Appl. Environ. Microbiol.*, 61:1458-1468.
- Hanisch, K., Schweitzer, B., and Simon, M. (1996). Use of dissolved carbohydrates by planktonic bacteria in a mesotrophic lake. *Microb. Ecol.*, 31:41-55.
- Hanna, J.V., Johnson, W.D., Quezda, R.A., Wilson, M.A., and Xiao-Qiao, L. (1991). Characterization of aqueous humic substances before and after chlorination. *Environ. Sci. Technol.*, 25:1160-1164.

- Hitzl, W., Rangger, A., Sharma, S., and Insam, H. (1997). Separation power of the 95 substrates of the BIOLOG system determined in various soils. *FEMS Microbiol. Ecol.*, 22:167-174.
- Huck, P.M. (1990). Measurement of biodegradable organic matter and bacterial growth potential in drinking water. *J. AWWA*, June:78-86.
- Huck, P.M., Fedorak, P.M., and Anderson, W.B. (1991). Formation and removal of assimilable organic carbon during biological treatment. *J. AWWA*, 83:69-80.
- Huck, P.M., Zhang, S., and Price, M.L. (1994). BOM removal during biological treatment: A first-order model. *J. AWWA*, July:61-71.
- Jorgensen, N.O.G., and Jensen, R.E. (1994). Microbial fluxes of free monosaccharides and total carbohydrates in freshwater determined by pad-hplc. *FEMS Microbiol. Ecol.*, 14:79-94.
- Jorgensen, N.O.G., Tranvik, L., Edling, H., Graneli, W., and Lindell, M. (1998). Effects of sunlight on occurrence and bacterial turnover of specific carbon and nitrogen compounds in lake water. *FEMS Microbiol. Ecol.*, 25:217-227.
- Kalmbach, S., Manz, W., and Szewzyk, U. (1997). Dynamics of biofilm formation in drinking water: Phylogenetic affiliation and metabolic potential of single cells assessed by formazan reduction and in situ hybridization. *FEMS Microbiol. Ecol.*, 22:265-279.
- Kaplan, L.A., & Bott, T.L. (1990). Nutrients for bacterial growth in drinking water: Bioassay evaluation. National Technical Information Service.. PB 89-213 995/AS, 1 p.
- Kaplan, L.A., & Newbold, J.D. (1993). Biogeochemistry of dissolved organic carbon entering streams. In T.E. Ford (Ed.), *Aquatic Microbiology, An Ecological Approach*. (pp. 139-165). Blackwell Scientific Publications.
- Kaplan, L.A., and Newbold, J.D. (1995). Measurement of streamwater biodegradable dissolved organic carbon with a plug-flow bioreactor. *Wat. Res.*, 29:2696-2706.
- Klingler, J.M., Stowe, R.P., Obenhuber, D.C., Groves, T.O., Mishra, S.K., and Pierson, D.L. (1992). Evaluation of the biolog automated microbial identification system. *Appl. Environ. Microbiol.*, 58:2089-2092.
- Konopka, A., Oliver, J., and Turco, R.F.Jr. (1998). The use of carbon substrate utilization patterns in environmental and ecological microbiology. *Microb. Ecol.*, 35:103-115.
- Kott, Y., Ribas, F., and Lucena, F. (1997). Comparison between the evaluation of bacterial regrowth capability in a turbidimeter and biodegradable dissolved organic carbon bioreactor measurements in water. *J. of App. Microbiol.*, 83:347-352.

- LeChevallier, M.W., Babcock, T.M., and Lee, R.G. (1987). Examination and characterization of distribution system biofilms. *Appl. Environ. Microbiol.*, 53:2714-2724.
- LeChevallier, M.W., Becker, W.C., Schorr, P., and Lee, R.G. (1992). Evaluating the performance of biologically active rapid filters. *J. AWWA*, April:136-146.
- LeChevallier, M.W., Schulz, W., and Lee, R.G. (1991). Bacterial nutrients in drinking water. *Appl. Environ. Microbiol.*, 57:857-862.
- Lucena, F., Frias, J., and Ribas, F. (1990). A new dynamic approach to the determination of biodegradable dissolved organic carbon in water. *Environ. Tech.*, 12:343-347.
- Manem, J.A., and Rittmann, B.E. (1992). Removing trace-level organic pollutants in a biological filter. *J. AWWA*, April:152-157.
- Manz, E., Szewzyk, U., Ericsson, P., Amann, R., Schleifer, K.-H., and Stenstrom, T.-A. (1993). In situ identification of bacteria in drinking water and adjoining biofilms by hybridization with 16S and 23S rRNA-directed fluorescent oligonucleotide probes. *Appl. Environ. Microbiol.*, 59:2293-2298.
- McFeters, G.A. (1990). Enumeration, occurrence, and significance of injured indicator bacteria in drinking water. In G.A. McFeters (Ed.), *Drinking Water Microbiology*. (pp. 478-492). New York, Inc.: Springer-Verlag.
- Merlet, N., Merlet, Y., Prevost, M., Desjardins, R., and Bablon, G. (1991).. Removal of organic matter in BAC filters: The link between BDOC and chlorine demand. Denver, CO: American Water Works Association Research Foundation. pp 1111-1127. WQTC Annual Meeting, Orlando, FL.
- Mikell, A.T.Jr., Smith, C.L., and Richardson, J.C. (1996). Evaluation of media and techniques to enumerate heterotrophic microbes from karst and sand aquifer springs. *Microb. Ecol.*, 31:115-124.
- Minitab. 12.1. Minitab. 1998. *Release 12 ed.* Windows 95.
- Moller, S., Kristensen, C.S., Poulsen, L.K., Carstensen, J.M., and Molin, S. (1995). Bacterial growth on surfaces: Automated image analysis for quantification of growth rate-related parameters. *Appl. Environ. Microbiol.*, 61:741-748.
- Moriarty, D.J.W., & Bell, R.T. (1993). Bacterial growth and starvation in aquatic environments. In S. Kjelleberg (Ed.), *Starvation in Bacteria*. (pp. 25-53). New York: Plenum Press.
- Morris, D.P., and Lewis, W.M.Jr. (1992). Nutrient limitation of bacterioplankton growth in Lake Dillon, Colorado. *Limnol. Oceanogr.*, 37:1179-1192.

- Paerl, H.W. (1991). Ecophysiological and trophic implications of light-stimulated amino acid utilization in marine picoplankton. *Appl. Environ. Microbiol.*, 57:473-479.
- Paerl, H.W., and Pinckney, J.L. (1996). A mini-review of microbial consortia: Their roles in aquatic production and biogeochemical cycling. *Microb. Ecol.*, 31:225-247.
- Pollard, P.C., and Greenfield, P.F. (1997). Measuring *in situ* bacterial specific growth rates and population dynamics in wastewater. *Wat. Res.*, 31:1074-1082.
- Putnam, S.W., and Graham, J.D. (1993). Chemicals versus microbes in drinking water: a decision sciences perspective. *J. AWWA*, March:57-61.
- Rheinheimer, G., and Gocke, K. (1994). The use of bacteriological variables for the characterization of different water bodies. *Int. Revue ges. Hydrobiol.*, 79:605-619.
- Ribas, F., Frias, J., Ventura, F., Mohedano, L., and Lucena, F. (1995). Assessment of biological activity and fate of organic compounds in a reactor for the measurement of biodegradable organic carbon in water. *J. Appl. Bact.*, 79:558-568.
- Russell, J.B., and Cook, G.M. (1995). Energetics of bacterial growth: Balance of anabolic and catabolic reactions. *Microbiol. Rev.*, 59:48-62.
- Schut, F., Prins, R.A., and Gottschal, J.C. (1997). Oligotrophy and pelagic marine bacteria: facts and fiction. *Aquat. Microb. Ecol.*, 12:177-202.
- Schweitzer, B., and Simon, M. (1995). Growth limitation of planktonic bacteria in a large mesotrophic lake. *Microb. Ecol.*, 30:89-104.
- Servais, P., Billen, G., and Hascoet, M.-C. (1987). Determination of the biodegradable fraction of dissolved organic matter in waters. *Wat. Res.*, 21:445-450.
- Shirey, J.J., and Bissonnette, G.K. (1991). Detection and identification of groundwater bacteria capable of escaping entrapment on 0.45- μ m-pore size membrane filters. *Appl. Environ. Microbiol.*, 57:2251-2254.
- Simon, M., and Azam, F. (1989). Protein content and protein synthesis rates of marine planktonic bacteria. *Mar. Ecol. Prog. Ser.*, 51:201-213.
- Sobsey, M.D., Dufour, A.P., Gerba, C.P., LeChevallier, M.W., and Payment, P. (1993). Using a conceptual framework for assessing risks to health from microbes in drinking water. *J. AWWA*, March:44-48.
- Sokal, R.R., & Rohlf, F.J. (1995). *Biometry, The Principles and Practice of Statistics in Biological Research*. New York: W.H. Freeman and Company.

Sommaruga, R. (1995). Microbial and classical food webs: A visit to a hypertrophic lake. *FEMS Microbiol. Ecol.*, 17:257-270.

Thomaz, S.M., and Wetzel, R.G. (1995). [^3H]Leucine incorporation methodology to estimate epiphytic bacterial biomass production. *Microb. Ecol.*, 29:63-70.

Tranvik, L. (1994). Colloidal and dissolved organic matter excreted by a mixotrophic flagellate during bacterivory and autotrophy. *Appl. Environ. Microbiol.*, 60:1884-1888.

Tranvik, L., and Kokalj, S. (1998). Decreased biodegradability of algal DOC due to interactive effects of UV radiation and humic matter. *Aquat. Microb. Ecol.*, 14:301-307.

Tranvik, L.J. (1988). Availability of dissolved organic carbon for planktonic bacteria in oligotrophic lakes of differing humic content. *Microb. Ecol.*, 16:311-322.

Tranvik, L.J. (1993). Microbial transformations of labile dissolved organic matter into humic-like matter in seawater. *FEMS Microbiol. Ecol.*, 12:177-183.

Tranvik, L.J., and Hofle, M.G. (1987). Bacterial growth in mixed cultures on dissolved organic carbon from humic and clear waters. *Appl. Environ. Microbiol.*, 53:482-488.

Vahjen, W., Munch, J.-C., and Tebbe, C.C. (1995). Carbon source utilization of soil extracted microorganisms as a tool to detect the effects of soil supplemented with genetically engineered and non-engineered *Corynebacterium glutamicum* and a recombinant peptide at the community level. *FEMS Microbiol. Ecol.*, 18:317-328.

van der Kooij, D. (1992). Assimilable organic carbon as an indicator of bacterial regrowth. *J. AWWA*, 84:57-65.

van der Kooij, D., Visser, A., and Hijnen, W.A.M. (1982). Determining the concentration of easily assimilable organic carbon in drinking water. *J. AWWA*, 74:540-545.

Van Der Wende, E., Characklis, W.G., and Smith, D.B. (1989). Biofilms and bacterial drinking water quality. *Wat. Res.*, 23:1313-1322.

van Loosdrecht, M.C.M., Lyklema, J., Norde, W., and Zehnder, A.J.B. (1990). Influences of interfaces on microbial activity. *Microbiol. Rev.*, 54:75-87.

Volk, C.J., Volk, C.B., and Kaplan, L.A. (1997). Chemical composition of biodegradable dissolved organic matter in streamwater. *Limnol. Oceanogr.*, 42:39-44.

Wolfaardt, G.M., Lawrence, J.R., Roberts, R.D., and Caldwell, D.E. (1994). The role of interactions, sessile growth, and nutrient amendments on the degradative efficiency of a microbial consortium. *Can. J. Microb.*, 50:331-340.

Wolfaardt, G.M., Lawrence, J.R., Robarts, R.D., Caldwell, S.J., and Caldwell, D.E. (1994). Multicellular organization in a degradative biofilm community. *Appl. Environ. Microbiol.*, 60(2):434-446.

CHAPTER 5

SUMMARY

In the last twenty years, knowledge in the area of microbial ecology has expanded greatly. This can be attributed to increased awareness of the role that bacteria play in nutrient cycling and the consequent effects that this has in shaping and modifying both the physical and chemical environment of the biosphere. Methodological advances have been instrumental in the acquisition of this information. Microscopic observation of "wee animalcules" began with Anton van Leeuwenhoek (1632 – 1723) more than two centuries ago and continues today. Advancements in computer technology and image analysis now allow microbiologists to investigate bacteria in unprecedented numbers, as well as analyze the data in a manner not previously possible.

Still, the data must be collected and analyzed and far too often in environmental microbiology, the datasets collected are small and under analyzed. This can arise due to a number of reasons, including (1) time constraints, (2) expense, (3) experimental design, and/or (4) statistical naiveté. The experimental design for this study was developed with statistical analysis at the forefront (e.g. see Figure 1.1). Due to the nature of biological data and especially environmentally related biological data, both large datasets and extended sampling periods are required.

Biological stability of drinking water is becoming increasingly important as safe and clean drinking water sources are disappearing with accelerating frequency as

industrialization continues. All of us have an interest in the delivery of safe drinking water; it is an absolute requirement that cannot be neglected or relegated a secondary priority. Bacteria can play both a deleterious and beneficial role in the delivery of safe drinking water. In both cases, the organisms main mode of growth is sessile, attached to the distribution system pipes in the first case and attached to filter media in the second case. The development of successful strategies to control bacterial growth in areas where we can maximize their beneficial role and minimize their detrimental role depend on our understanding of the life history strategies and growth of sessile bacterial communities.

As a prelude to the study of bacterial growth in biofilms, it was necessary to perform the measurements presented in Chapter 2. In Chapter 2, I combined epifluorescence microscopy, computer technology and sophisticated statistical analyses to observe, collect, and analyze an extremely large dataset of bacterial cell sizes. In the course of attempting to derive a relationship between biovolume (or cells) and biomass for our particular system, I noted that the distribution of biovolume abundance was very similar to distributions that describe biomass/size abundance relationships at trophic levels above bacteria. Further investigation led to and resulted in the studies reported on in Chapter 2. This resulted in the identification of a probability distribution that could effectively model 94.4% of the data acquired in this study. Probability plots produced from the data potentially provide us with the ability to predict biomass from relatively few biovolume measurements as compared to the large number of measurements needed in generating the models.

During the course of these experiments, I confirmed previous speculation on the nature of biomass distributions in bacterial communities. Due to the highly skewed and

non-linear relationship between biovolume and biomass in natural assemblages of bacteria, models based on single value conversion factors were concluded to be inappropriate. Utilization of single value conversion factors for bacterial biovolume to biomass calculations results in under- or overestimation of the carbon sequestered in natural communities. This all depends on where most of the biomass falls within the distribution of cell sizes. Underestimation would occur when a high percentage of the population is composed of small cells. Conversely, overestimation would occur when a high percentage of the community is composed of large cells. In order to achieve a realistic value for bacterial biomass, calculations must take into account the skewed and non-linear relationship between biovolume and biomass.

Enumeration of the active component in the bacterial communities arising from the different treatments applied in these experiments comprised Chapter 3. I compared the active population to the total population by a number of different methods. Enumeration of the metabolically active component of microbial communities is required for the accurate estimation of productivity, biomass turnover and nutrient cycling. While an accurate, reproducible, and easy method for enumerating this active fraction still eludes microbiologists, microautoradiography (MAR) is likely the most sensitive. This method suffers from the fact that it is very time consuming, and as such, only one experiment from each of the replicate experiments was used to measure activity in the population of cells arising from the biofilms in this study.

Comparative analysis of MAR with a number of frequently used fluorescent activity stains showed that they worked well and produced similar results for the experimental system used in this study. A problem with fluorescent stains is that they

only measure potential activity and not active involvement in growth, production, and nutrient turnover. This arises due to the method in which these stains detect activity. For example, membrane potential is required in order for an organism to maintain its integrity and exclude the outside environment. Stains that target the membrane potential only tell us if the organism is capable of activity and not if it is actively involved in protein synthesis and nutrient turnover at the specific moment of sampling, it may be dormant or dying. One exception could be the respiratory stain CTC, which indicates that a cell is involved in active respiration. This implies that at the time of sampling the cell was actively reducing oxygen and therefore could be considered as actively involved in substrate turnover (disregarding potential endogenous metabolism). In contrast to the potential activity stains, MAR tells us directly if the organism is actively involved in substrate uptake and protein synthesis at the time of sampling.

All of the treatments used in these experiments had an effect on the activity of the populations arising from the biofilms. The most significant effect was due to chlorine treatment with less significant effects produced by carbon source and carbon concentration treatments. While there were significant differences in the active fraction of cells according to these treatments, they were remarkably consistent in that a highly significant fraction (>50%) of the population was active in all cases, with the exception of the carbohydrate-grown organisms subjected to chlorination. The activity measurements estimated by MAR were also used in Chapter 4 for calculating growth rates of the biofilm organisms.

Bacteria have a unique ability to show explosive growth given the right conditions of food supply and space. Jacques Monod developed an empirical relationship between bacterial growth and substrate concentration in the 1950s that is still used for modeling bacterial growth in chemostats and batch cultures today. Despite evidence to the contrary, many microbiologists and engineers persist in attempting to force this model on bacterial growth in all environments. Monod developed his model with axenic cultures in the laboratory environment of culture flasks at high substrate concentrations. In the case of environmental studies such as this one, undefined, mixed bacterial populations are competing for a limited supply of carbon at low concentrations. The natural environment is not a chemostat or batch culture, and bacterial growth on surfaces does not necessarily conform to the kinetic model developed by Monod.

In Chapter 4, I show that the assumption of first-order kinetics for growth of biofilm bacteria is incorrect at the low concentrations of substrate used in these experiments. A number of models have been developed to predict biofilm growth, most of which have met with limited success. All of these models assume that growth on surfaces occurs in the same manner as growth in the fluid phase as modeled by Monod kinetics. If the organisms in a distribution system biofilm or biological filter are growing with zero order kinetics, these models that use first order kinetics for predicting growth will continue to be ineffective. Population density as predicted by r/K -selection theory was more highly correlated with growth rates than was substrate concentration. In this case, as cell or biomass areal density increases, growth rates decrease. This can be attributed to an interaction between substrate loading rate and the amount of food available to individual cells in the biofilm. While higher rates of food supply support an

increased population density, the amount of food available to individual cells decreases. Chlorine treatment exerted maximal control in the carbohydrate-grown biofilms by limiting population density and this ensured that there was sufficient carbon available to those organisms that could maintain a niche in the reactors. Ultimately, this produced growth rates in the chlorinated reactors that were higher than the parallel control reactors. What this means in practical terms is that in order to know the rate at which the organisms are growing, direct measurements, such as I performed in these experiments is required.

In summary, the treatment that had the largest impact on all measured parameters was chlorination. It was surprising to find that substrate concentration had a relatively small impact in these experiments. Given the indoctrination of first order kinetics at almost all levels of microbiology from growth to enzyme catalysis, the elucidation of zero order growth kinetics was unexpected. Zero order kinetics at the substrate concentrations used in these experiments for biofilm growth could explain the inadequacies of current models used for predicting bacterial growth in distribution systems. While substrate concentration did have an impact on biovolume, it had little affect on the parameters that describe the distribution of biovolume, somewhat similar to the results for growth rates.

Substrate type had an impact on all measured parameters including probability distributions, volumes, rates, population densities, production and activity. This was expected. What was not expected was the ability of the biofilm communities to utilize the humic substrate to such a large extent. Humic material is usually considered a recalcitrant substrate. Extremely high steady state degradation rates of 58 and 78% of the influent

substrate concentration were measured in the chlorinated and control biofilms, respectively. This indicates that the community organization provided by forming a biofilm enhances degradation when compared to the relatively unstructured mode of existence shown by free-living bacteria.

MONTANA STATE UNIVERSITY - BOZEMAN



3 1762 10421356 4