Identification Studies of *Bacillus* Spores Using Fluorescence Spectroscopy

A thesis submitted in partial fulfilment of the
requirements for the degree of
Doctor of Philosophy in Medical Physics

JOSEPH KUNNIL MSc, GDip Tchg

Department of Physics and Astronomy

University of Canterbury

Christchurch, New Zealand

2005

Supervisor: Associate Professor Lou Reinisch PhD

This work is dedicated to the loving memory of my beloved parents

Abstract

Fluorescence spectroscopy was examined as a potential technique for identifying aerosol particles like bacterial spores. This technique was used for laboratory measurements on some common biological agent simulants. We have measured the intrinsic steady-state fluorescence emission spectra as a function of the excitation wavelength for several bacterial spores (washed and unwashed) in dry and aqueous suspensions at room temperature using excitation wavelengths from 200 to 600 nm. These measurements were compared to those of common, naturally occurring biological components like fungal spores and pollen and non spore samples like ovalbumin. The spectra of samples were combined into fluorescence profiles or fluorescence fingerprints. Different substrates were used for collection and detection of spores. Each bacterium produces a unique *in vitro* fluorescence profile when measured in dried and aqueous suspension and exhibits a strong maximum in its fluorescence emission spectrum near 330-340 nm. The fluorescence profiles were reproducible. The complexity of microorganisms made the interpretation of their spectral signature a difficult task. Principal components analysis (PCA) and cluster analysis were done as a data reduction technique for detection and identification from different backgrounds. PCA illustrates that linear combination of detected fluorescence intensities, which are present in different ratios in each of samples studied, can be used to discriminate biological agent simulants from other biological samples. The hydration effects, washing effects and the role of tryptophan on spore fluorescence were also investigated. The emission spectra of the dried spores showed a maximum near 330 nm, suggesting a hydrophobic environment for its tryptophan residues. The aqueous solution of tryptophan showed fluorescence shifted to 360 nm and in ethanol solution the maximum was shifted to 340 nm, suggesting a rather more polar average location of the tryptophan. To find the limit of detection we measured the quantum efficiency (QE) of a few samples. We concluded that spectroscopy techniques coupled with effective interpretation models are applicable to biological simulants agents.

Index Heading: Bacteria; Spores; Identification; Fluorescence; Fluorescence Quantum Efficiency; Principal Components Analysis; Cluster Analysis.

Acknowledgments

Many thanks go to Assoc. Prof. Lou Reinisch whose insight and guidance have been invaluable in the struggles of the last few years. The knowledge and the skills, which have been passed on by Prof. Lou, are gratefully acknowledged. Prof. Phil Butler (Head of Department) as co supervisor has also had an influence on these studies. Dr Easaw Chacko of Mathematics and Statistics Department as an associate supervisor has had a leading role in the progress of the research. My special thanks go to two of my friends, Mr. Thomas Francis of Chemistry Department and Mr. Murthy Mittinty of Mathematics and Statistics Department. I also acknowledge my colleagues Ekta Jhala and Sivananthan, S. Also, thankful to the School of Medicine, Christchurch Public Hospital for providing their fluorophotometer for the initial measurement. I gratefully acknowledge the Nova Sol for providing the funding for this research. I also acknowledge the help from the staff of Physical Library especially Mrs Angela Davies. During the course of this work I have been fortunate to have available the wide variety of services provided by all of the technical and office staff of this department. Thanks to the Department of Physics and Astronomy for the financial support throughout this work.

Finally I thank my wife Ann, daughter Vishakham and my son Vanamali for their love and support throughout the study.

CONTENTS

1 Abstract iii

2 Acknowledgments iv

3 List of Figures viii

4 List of Tables xvi

1 Introduction 1

1.1 Motivation 1

1.2 Bacterial Detection and Identification 2

1.3 Mathematical Models in Identification and Classification of Bacterial Spores 7

1.4 Overview of the Thesis 10

2 Introduction to Ultraviolet-Visible Spectroscopy and Data Analysis 13

2.1 General Principles of Spectroscopy 13

2.1.1 Transmittance and Absorbance 15

2.1.2 Luminescence 16

2.1.3 The Fluorescence Phenomenon and Frank-Condon Principle 17

2.1.4 Fluorophores and Biological Fluorescence 20

2.1.5 Comparison of Fluorescence Quantum Efficiency 23

2.2 Fluorescence Excitation-Emission Spectra via the Two-Dimensional Matrix Method 24

2.3 Fluorescence Instrumentation 25

2.4 Analysis Data Analysis using Principal Components Analysis and Cluster 27

3 Experimental Section 30

3.1 Spore Samples/Simulants/Interferents and Substrates 30

3.2 Spectral Measurements 33

3.3 Spores Fluorescence 36

3.3.1 Fluorescence Quantum Efficiency of Dry *Bacillus* Spores 36

3.3.2 Spores Detection from Tape Substrates/Dusts 39

3.3.3 Spores Detection from Quartz Substrates/Pollen 42

3.3.4 Effects of Washing on the Identification of *Bacillus* Spores 43

3.3.5 Autofluorescence of *Bacillus* Spores and Role of Tryptophan: A Comparative Study of Dry Spores and Spores in Aqueous and Ethanol Suspensions. 44

4 Spectroscopy and Data Analysis Results and Discussions 46

4.1 Spectral Visualization by EEM 47

4.2 Fluorescence Quantum Efficiency Measurements of *Bacillus* Spores 52

4.2.1 Absorbance and Fluorescence of Anthracene in Ethanol 52

4.2.2 Quantum Efficiency of *Bacillus* Spores 58

4.2.2.1 Quantum efficiency of unwashed and washed *B. globigii* (i) spores ……59

4.2.2.2 Quantum efficiency of unwashed and washed *B. globigii* (ii) Spores ……65

4.2.2.3 Quantum efficiency of unwashed and washed *B. globigii* (iii) Spores……. 69

4.2.2.4 Quantum efficiency of unwashed *B. cereus* spores 73

4.2.3 Discussions on the Quantum Efficiency Measurements of *Bacillus* Spores 77

4.3 Detection of *Bacillus* Spores from Different Substrates/Dust/Pollen and Effects of Washing on Fluorescence 82

4.3.1 Detection of Spores from Different Substrates 82

4.3.2 Detection and Identification of Dry Spores from Dusts Background on Tape Substrate 84

4.3.2.1 Spores Fluorescence 84

4.3.2.2 PCA and Cluster Analysis 87

4.3.3 Detection and Identification of *Bacillus* Spores from Background Pollen like Pig weed 97

4.3.3.1 Spore Fluorescence 97
4.3.3.2 PCA and Cluster Analysis 98
4.3.4 Effects of Washing on Identification of *Bacillus* Spores 104
4.3.4.1 Spore Fluorescence 104
4.3.4.2 PCA and Cluster Analysis 107

4.4 Autofluorescence of *Bacillus* spores and Role of Tryptophan: A Comparative study of Dry Spores and Spores in Aqueous and Ethanol Suspensions 112
4.4.1 Tryptophan Absorption and Fluorescence 112
4.4.2 Autofluorescence of *Bacillus globigii* (i) 115
4.4.3 PCA and Cluster Analysis 117

4.5 Discussions on Spore Fluorescence and Data Analysis 122
4.5.1 Spore Fluorescence 122
4.5.2 Data Analysis 131

5 Conclusions 135

6 References 140

7 APPENDIX 158

7.1 List of Publications 158

7.2 *Mathematica* Program for Smoothing the Spectrum 160

7.3 *Mathematica* Program for Plotting 2D Fluorescence Fingerprints 161

7.4 Principal Components Model 163

7.5 The R Program for PCA and Cluster analysis 166

7.6 PCA Output 172

List of Figures

Figure 2.1 The electromagnetic spectrum 14

Figure 2.2 Incident light/cell/transmitted light 15

Figure 2.3 Electronic energy level diagram and the transition affecting fluorescence and phosphorescence. s_0, s_1 and T_1 are unexcited singlet, excited singlet (fluorescence) and triplet (phosphorescence) chromophores states. 18

Figure 2.4 Illustrations of the absorption and fluorescence emission spectra of an anthracene solution as approximate mirror symmetry 19

Figure 2.5 Amino acids group and their absorption spectrum 21

Figure 2.6 Components of conventional fluorophotometer is shown with two single monochromator configurations (Retrieved from operational manual version 4.0 of SLM 8000C, Urbana, IL) 26

Figure 3.1 The SLM 8000C Spectrofluorometer (Retrieved from operational manual version 4.0 of SLM 8000C, Urbana, IL) 34

Figure 3.2 Schematic of how the dry spores are mounted to the quartz slide and placed in the fluorometer 38

Figure 3.3 Schematic of how the dry spores are mounted to the tape and placed in the fluorometer 41

Figure 4.1 Illustration of the fluorescence emission spectrum from dry *B. globigii* (i) spores on quartz substrate. The excitation is at 280 nm. The points show the measured data during a typical experiment. The solid line is from the Gaussian smoothing and parameterisation of the data. 48

Figure 4.2 Illustration of the fluorescence emission spectrum from dry *B. globigii* (i) spores on quartz substrate. The excitation is at 420 nm. The points show the measured data during a typical experiment. The solid line is from the Gaussian smoothing and parameterisation of the data. 49

Figure 4.3 Illustration of the fluorescence fingerprints from dry spores of *B. globigii* (i) measured on quartz substrate. The generation of the fingerprint is explained in the text. In 4.3(A) the conventional fluorescence fingerprint and (B) the modified fluorescence fingerprint are displayed. In both parts of the figure, the contours represent equal changes in the fluorescence intensity, and the peak intensity is normalised to 1. 51

Figure 4.4 Absorption spectrum of anthracene in ethanol at a 2.21E-04 mol/litre concentration. The absorbance at excitation wavelength 355.52 nm is 1.6153 O.D. 52

Figure 4.5 Fluorescence spectrum of anthracene in ethanol. Sample fluorescence spectrum of anthracene in ethanol is shown with excitation at 360 nm with emission maximum at 402 nm. 53

Figure 4.6 Sample fluorescence fingerprints of (A) ethanol and (B) anthracene in ethanol. The contour plot is the relative intensity of fluorescence as a function of the excitation wavelength and emission wavelength divided by the excitation wavelength. The excitation wavelengths were from 200-600 nm 54

Figure 4.7 Fluorescence of anthracene in ethanol as a function of absorbance. The integrated fluorescence intensity (counts) of anthracene in ethanol is shown versus absorption (O.D) at 320, 330, 340, 350 and 360 nm excitation wavelengths (with corresponding colours blue, red, yellow, light blue and brown respectively). The line is linear least squares fit. Each slope is proportional to the known QE of anthracene in ethanol at the excitation wavelength selected. 56

Figure 4.8 Response plot of the measured fluorescence intensity (counts) versus absorbance (O.D) against excitation wavelength range from 320 to 360 nm.... 57

Figure 4.9 Representative plots of absorbance spectrum of unwashed Bacillus spores studied for QE measurements. Samples: *B. globigii* (i) (Blue), *B. globigii* (ii) (red), *B. globigii* (iii) (yellow) and *B. cereus* (light blue). Scanning wavelength was from 200-600 nm. 58

Figure 4.10 Representative plots of absorbance spectrum of washed spores of *B. globigii* samples (i), (ii) and (iii). Absorbance (O.D) versus wavelength (nm) is shown. Scanning wavelength was from 200-600 nm. 59

Figure 4.11 Representative fluorescence spectrum of unwashed and washed *B. globigii* (i) spores, measured at excitation wavelengths 280, 360 and 400 nm. The emission ranges were 290-570, 370-700 and 410-700 nm. 60

Figure 4.12(A-B) The QE of unwashed *B. globigii* (i) spores at 280, 360 and 400 nm excitation wavelengths. In 4.12(A) the absorbance (O.D) as function of number spores is displayed and in 4.12(B) the integrated fluorescence intensity (counts) of *B. globigii* (i) spores is shown versus absorption (O.D). The lines are linear least squares fit represented by the corresponding colours blue, red and yellow respectively. The slopes are proportional to the QE of *B. globigii* (i). 62

Figure 4.13 The QE of washed *B. globigii* (i) spores at excitation wavelengths 280, 360 and 400 nm. The integrated fluorescence intensity (counts) of *B. globigii* (i) spores is shown versus absorption (O.D). The line is linear least squares fit and is represented by colours: blue, red and yellow. The slopes are proportional to the QE of *B. globigii* (i) spores. 64

Figure 4.14 Comparison of fluorescence between washed (red) and unwashed (blue) *B. globigii* (i) spores at 400 nm excitation wavelength. The corrected integrated fluorescence intensity (counts) of *B. globigii* (i) spores is shown versus absorption (O.D). The line is linear least squares fit. The slopes are proportional to the QE of *B. globigii* (i) spores at 400 nm. 65

Figure 4.15 The QE of unwashed *B. globigii* (ii) spores, at 280, 360 and 400 nm excitation wavelengths (represented by colours blue, pink and yellow respectively). (Top) the absorbance of the spore as function of number spores is displayed and (Bottom) the integrated fluorescence intensity (counts) of spores is shown versus absorption (O.D). The slopes are proportional to the QE of *B. globigii* (ii) spores. 67

Figure 4.16 Fluorescence versus absorbance plots of washed and redried spores of *B. globigii* (ii) at 280, 360 and 400 nm excitation wavelengths (represented by

colours blue, red and yellow respectively). The integrated fluorescence intensity (counts) of *B. globigii* (ii) spores is shown versus absorption (O.D). The line is linear least squares fit. The slopes are proportional to the QE of *B. globigii* (ii) spores. 68

Figure 4.17 The QE of unwashed *B. globigii* (iii) spores, at 280, 360 and 400 nm excitation wavelengths (represented by colours blue, red and yellow respectively). (Top) the absorbance of the spore as function of number spores is displayed and (Bottom) the integrated fluorescence intensity (counts) of *B. globigii* (iii) spores is shown versus absorption (O.D). The line is linear least squares fit. The slopes are proportional to the QE of *B. globigii* (iii) spores. ... 70

Figure 4.18 Fluorescence versus absorbance plots of washed *B. globigii* (iii) spores, at excitation wavelengths 280, 360 and 400 nm (represented by colours blue, red and yellow respectively). The integrated fluorescence intensity (counts) of *B. globigii* (iii) spores is shown versus absorption (O.D). The line is linear least squares fit. The slopes are proportional to the QE of *B. globigii* (iii) spores. ... 71

Figure 4.19 Comparison of fluorescence of unwashed and washed spores of *B. globigii* (iii) spores. The corrected integrated fluorescence intensity (counts) versus absorption (O.D) is shown for excitation wavelength 280 nm. The line is linear least squares fit. The slopes are proportional to the QE of *B. globigii* (iii) spores. 72

Figure 4.20 Representative fluorescence spectrum of unwashed *B. cereus* spores at excitation wavelengths 280 and 300 nm. The emission wavelength ranges were from 290-550 and 310-590 nm. 73

Figure 4.21 The QE of unwashed *B. cereus* spores. (Top) The absorbance of the spore as function of number spores. (Bottom) the corrected integrated fluorescence intensity of *B. cereus* is shown versus absorption (O.D). The line is linear least squares fit represented by colours blue and red for excitation 280 and 300 nm respectively. The slopes are proportional to the QE of *B. cereus*. . 75

Figure 4.22 Fluorescence fingerprints of substrates and spores coated on substrates using excitation wavelengths from 200-600 nm. In Figure (A-B) fingerprints

from tape and quartz slide is displayed. In Figure (C-D) fluorescence fingerprints of dry spores of *B. globigii* (i) on tape and quartz slide are displayed for comparison. Details of measurements are explained in the text. In all figures, the contours represent equal changes in the fluorescence intensity, and the peak intensity is normalised to 1. 83

Figure 4.23 Fluorescence fingerprints of (A) dust A; (B) *B. globigii* (i) and dust A; (C) *B. cereus* and dust A; (D) *B. popilliae* and dust A using excitation wavelengths from 200-600 nm is displayed. The fluorescence signal is divided into 15 equally spaced contour lines. Blue represents the maximum intensity. 85

Figure 4.24 Fluorescence fingerprints of (A) *B. globigii* (ii); (B) corn smut and (C) ovalbumin using excitation wavelengths from 200-600 nm is displayed. The fluorescence signal is divided into 15 equally spaced contour lines. Blue represents the maximum intensity..... 86

Figure 4.25 Scree plot for PCA analysis as explained in the text. The first two PCs would be analysed using this method 93

Figure 4.26 (A) PCA plot of three samples of *B. globigii* (i) (open squares), *B. cereus* (open circles) *and B. popilliae* (crosses) (9 measurements). Each sample is mixed with dust as explained in the text. PC1 versus PC2 is plotted. Plots of the PCA analysis of same three samples of *B. globigii* (i), *B. cereus*, *B. popilliae* (9 measurements) with the added sample of (B) *B. globigii* (ii) (solid diamond), (C) corn smut (solid diamond) and (D) ovalbumin (solid diamond) 94

Figure 4.27 Silhouette plot of the strength of clustering using PAM with three samples of *B. globigii* (i) (open squares), *B. cereus* (open circles), *B. popilliae* (crosses) (9 measurements) as shown in Figure 4-26. Silhouette plot of the strength of clustering using PAM with the same three samples of *B. globigii* (i), *B. cereus*, *B. popilliae* with added sample of (B) *B. globigii* (ii) (solid diamond), (C) corn smut (solid diamond) and (D) ovalbumin (solid diamond) 96

Figure 4.28 Fluorescence fingerprints of (A) *B. cereus*, (B) *B. globigii* (iii), (C) pig weed and (D) mixed samples of pig weed and *B. globigii* (iii) using excitation

wavelengths from 200-600 nm is displayed. The fluorescence signal is divided into 15 equally spaced contour lines. Blue represents the maximum intensity. 97

Figure 4.29 PCA plot of four samples of *B. cereus* (open circles), pig weed (triangles), *B. globigii* (iii) (plus) and mixture of *B. globigii* (iii) and pig weed (cross) (12 measurements). PCA was done with five excitation wavelengths 280, 320, 360, 400 and 440 nm and the corresponding emission wavelengths used from 310-410, 350-450, 390-490, 430-530 and 470-570 respectively. PC1 versus PC2 is plotted. 101

Figure 4.30 PCs 1, 2 and 3 for the 12 fluorescence measurements made with four different samples each with three different preparations. The circles' are *B. cereus*. The triangles' are pig weed. The pluses are the *B. globigii* (iii) and the crosses' are mixed sample of *B. globigii* (iii) and pig weed. 102

Figure 4.31 (Top) Dendrogram of the clustering of objects according to the first and second PCs as shown in Figure 4-29, displaying the relationships among spectra of samples based on correlation coefficient calculated by PCA. The strength of the four clusters shown as the silhouette widths S_i in Figure 4-31 (Bottom) .. 103

Figure 4.32 Fluorescence fingerprints of (A) *B. cereus*; (B) *B. popilliae* and (C) *B. thuringiensis* (all unwashed) using excitation wavelengths from 200-600 nm are displayed. The fluorescence signal is divided into 15 equally spaced contour lines. Blue represents the maximum intensity. 105

Figure 4.33 Fluorescence fingerprints from dry spores of (A) *B. globigii* (i) before washing, (B) *B. globigii* (i) after washing, (C) *B. globigii* (ii) before washing, (D) *B. globigii* (ii) after washing, (E) *B. globigii* (iii) before washing, and (F) *B. globigii* (iii) after washing. The excitation wavelengths from 200-600 nm were used……..... 106

Figure 4.34 Scree plot for PCA analysis as explained in the text. The first three PCs would be analysed using this method 108

Figure 4.35 PC1 versus PC2 for the 27 fluorescence measurements made. The circles are *B. globigii* (i). The filled circles are before washing and the open circles are after washing. The triangles are *B. globigii* (ii). The filled squares are before

washing and the open squares are after washing. The oval is drawn to demonstrate the clustering. The diamonds are the *B. cereus* samples, the x are the *B. popilliae* samples, and the + are the *B. thuringiensis* samples. 109

Figure 4.36 PCs 1, 2 and 3 for the 27 fluorescence measurements made. The circles are *B. globigii* (i). The filled circles are before washing and the open circles are after washing. The triangles are *B. globigii* (ii). The filled triangles are before washing and the open triangles are after washing. The squares are *B. globigii* (iii). The filled squares are before washing and the open squares are after washing. The diamonds are the *B. cereus* samples, the x are the *B. popilliae* samples, and the + are the *B. thuringiensis* samples. 110

Figure 4.37 (A) Dendrogram of the clustering of objects according to the first and second PCs as shown in Figure 4.35 displaying the relationships among spectra of samples based on correlation coefficient calculated by PCA. Samples with "un" were measured unwashed and samples with "wa" were measured at being washed twice and redried. (B) The strength of the three clusters shown with the brackets in the bottom of Figure 4.37(A). 111

Figure 4.38 Sample absorption spectrum of tryptophan in aqueous (blue) and ethanol (red) suspension. 112

Figure 4.39 Sample fluorescence spectrum of tryptophan in dried form, in aqueous and ethanol solutions at 290 nm excitation wavelength with emission peaks at 339, 360, 348 nm respectively. 113

Figure 4.40 Fluorescence fingerprints of tryptophan in dried form, aqueous and ethanol solutions using excitation wavelengths from 200-600 nm. In Figure 4.40(A) Florescence fingerprint of selected filter paper is displayed and in Figure (B-C) the fluorescence fingerprints of tryptophan; (B) in dry, (C) in aqueous solution and in (D) in ethanol solution were displayed. The excitation was stepped every 10 nm. The intensity is equally divided by 15 contour lines. The greatest intensity is blue and the least intensity is red. 114

Figure 4.41 (A) Fluorescence fingerprints of dry *Bacillus* endospores (*B. globigii* (i)) on filter paper. In Figure 4.41(B-C) the fingerprints of *B. globigii* (i) endospores

in aqueous and ethanol suspensions are displayed. The excitation was stepped every 10 nm from 200-600 nm. The intensity is equally divided by 15 contour lines. The greatest intensity is blue and the least intensity is red..... 116

Figure 4.42 PCA plot of samples of *B. globigii* (i) spores in dried form (open circles), spores in aqueous (triangles) and in ethanol (plus) suspensions (9 measurements). In (A) PC1 versus PC2 is plotted and in (B) PC1, PC2 and PC3 are plotted..... 120

Figure 4.43 (Top) Dendrogram of the clustering of objects according to the first and second PCs are shown. (Bottom) The strength of the three clusters shown. The cluster strength is defined in the text. 121

List of Tables

Table 2-1 Excitation and emission maxima of biological molecules that exhibit endogenous fluorescence (Extracted from Lakowicz, Principles of Fluorescence Spectroscopy) 20

Table 3-1 List of simulants and interferents investigated 32

Table 3-2 Excitation wavelengths used and the range of emission wavelengths used to collect the emission spectra. 35

Table 3-3 The different types of substrate used for the collection and detection of spores. 40

Table 4-1 Absorbance of anthracene in ethanol in different excitation wavelength.. 55

Table 4-2 Fluorescence data of anthracene in the selected excitation wavelengths .. 55

Table 4-3 The summary of the measured and calculated slope for various excitation wavelengths 57

Table 4-4 The optical measurements of unwashed *B. globigii* (i) spores 61

Table 4-5 The optical measurements of washed and redried spores of *B. globigii* (i) 63

Table 4-6 Measured absorbance (O.D) and integrated fluorescence intensity (counts) of unwashed *B. globigii* (ii) spores at 280, 360, 400 nm excitation wavelengths 66

Table 4-7 The optical measurement of washed *B. globigii* (ii) spores at excitation wavelengths 280, 360 and 400 nm 68

Table 4-8 The optical measurements of *B. globigii* (iii) spores (unwashed) at excitation wavelengths 280, 360 and 400 nm 69

Table 4-9 The optical measurements of washed and redried spores of *B. globigii* (iii) at excitation wavelength 280, 360 and 400 nm 71

Table 4-10 Measured absorbance and integrated fluorescence intensity (counts) of unwashed *B. cereus* spores at excitation wavelengths 280 and 300 nm 74

Table 4-11 shows the summary of optical property measurements of the *Bacillus* samples investigated. 76

Table 4-12 Correlation matrix for the PCA analysis of three samples each of..... 90

Table 4-13 Principal Components Scores (a_{ij}) 90

Table 4-14 Eigenvalue 91

Table 4-15 Correlation matrix for PCA analysis of *Bacillus* spores from pollen background..... 99

Table 4-16 Eigenvalue 99

Table 4-17 Eigenvalue 107

Table 4-18 Correlation matrix for PCA analysis for *B. globigii* comparative study. 118

Table 4-19 Eigenvalues..... 118

Chapter 1

Introduction

1.1 Motivation

Rapid, sensitive, and selective bacterial detection and identification at the species level is necessary to discriminate pathogenic and nonpathogenic microbial species. The development of technology to realize this level of distinction for microbial species would have a considerable impact on occupational and health care, defense, and environmental monitoring. The possible use of biological warfare agents is widely known. There are a number of potential biological agents. For example, *Bacillus anthracis, Yersnia pestis* and *Variola virus*, which are the pathogenic agents of anthrax, plague and smallpox, respectively. Among these, the most infamous pathogen is *Bacillus anthracis* [1]. The best defence against such material relies on the incorporation of a wide range of technologies. Timely warning and identification of the presence of such agents is key to this approach. Therefore a critical aspect of improving protection against these materials is the development of faster and more reliable detection systems. Given the severity of the threat posed by biological agents, detection of potentially harmful pathogens needs to be extremely sensitive since the presence of a relatively small number of pathogens can lead to infection.

A few species of bacteria have the ability to produce highly resistant structures known as endospores. Endospore-forming bacteria are gram-positive and two major endospore-forming bacteria are *Bacillus* (aerobic) and *Clostridium* (anaerobic) [2]. These resist a range of hazardous environments and are protected against heat, radiation and desiccation, and are metabolically dormant [3-7]. Endospores form within (hence *endo*-) special vegetative cells known as *sporangia* (singular *sporangium*). Endospores are commonly known as spores and both terms are used here. Each spore is composed of a protoplast, a protoplast membrane, cortex and three coat layers.

Bacillus central protoplast contain dipicolinic acid (DPA) and it has been suggested that it is involved in spore dormancy [8, 9]. It has been proposed to be tightly packed with the vital cell constituents, linked with proteins or amino acids, or intercalated with DNA. Slieman et al. [10] found that about 10 % of spores' dry weight is DPA. DPA exists as calcium complex.

The problems associated with the *Bacillus* genus are as diverse as the environments where they are found. Although the vegetative cells of endospores forming bacteria are commonly found in soil, endospores can be seen everywhere including the atmosphere, where they are carried by dust [11]. The bacterial spores are the cause of problems such as food spoilage and food-borne illnesses [12]. Diseases caused by sporing bacteria include botulism (*Clostridium botulinum*), tetanus (*Clostridium tetani*), acute food poisoning (*Clostridium perfringens*), and gas gangrene (*Clostridium perfringens* type A). Since the genus *Bacillus* contains *Bacillus thuringiensis*, an industrially important nonpathogenic pesticide and *Bacillus cereus*, a noninfectious food pathogen, rapid discrimination of the spores from each other is necessary for effective intervention and treatment of human disease.

1.2 Bacterial Detection and Identification

The interest in the application of new techniques to bacterial identification results largely from the limitations of more traditional identification procedures such as growth on nutrient agar or immunology. Traditionally, microbiologists performing bacterial identification rely on cultivation of organisms, despite the realization that most of them (>99%) are not cultivable by standard methods [13-15]. Generally, no single simple test provides a definite identification of an unknown bacterium. Hence a complex series of tests is often required before identification can be confirmed. The results of such a series are often difficult to interpret and not available on the time scale desired in the clinical laboratory. There is an urgent need for the rapid assay of chemical and biological unknowns, such as bioaerosols. These considerations have prompted a number of researchers to explore modern instrumental alternatives to classical identification procedures.

Modern identification procedures are characterised by rapid acquisition, high reproducibility of data and computer aided data recording and interpretation. Substantial progress has been made over the past decade. Bacterial identification by a variety of methods has become an essential diagnostic tool in areas such as healthcare, food, water quality testing and enzyme discovery [16-24]. Ultraviolet resonant Raman spectroscopy has been successfully applied to the identification of biopolymers, bacteria and bioaerosols like bacterial spores [25-29]. A number of unique approaches have been applied to the identification of bacteria. Several groups started fragmenting bacteria by pyrolysis or laser ablation and measuring the mass distribution of the fragments. The results of different authors are not identical, however, this techniques can be used for a reliable discrimination between biological and non biological samples [30-32].

Nucleic acid and antibody based assays are also used to rapidly and reliably detect pathogenic bacteria in foods. Genetic identification relies on polymerase chain reaction (PCR)-DNA sequencing to identify the species by their nucleotide sequences [33]. A PCR has been developed for the detection of *Clostridium botulinum* type A, a cause of human botulism [34]. A two primer set and oligonucleotide detection probe was used to specifically detect *Cl. botulinum* type A neurotoxin gene.

A technique for the identification of bacteria by PCR with fluorescein-labelled primers chosen from the conserved regions of the 16S rRNA gene flanking a variable region been developed [35]. The PCR product was denatured, separated on a nondenaturing gel, and detected by an automated DNA sequencer. The mobility of the single stranded DNA is sequence dependent and allows the identification of broad panel of bacteria. In the reported study, a broad range of gram-negative and gram-positive bacteria was tested by PCR single-strand conformation polymorphism, including, *Escherichia coli, Enterobacter spp., Klebsiella spp., Haemophilus spp., Neisseria spp., Staphylococcus spp, Streptococcus spp., Enterococcus* and *Bacillus spp.* The PCR technique shows promise for the rapid identification and a small number of bacteria in many situations. It was reported that oligonucleotide primers proved suitable for the routine identification of *Mycobacterium xenopi*, starting from one single colony on solid medium or from a liquid culture [36]. The primers were used in the PCR to amplify a specific 584-bp DNA fragment, located in the 16S RNA

gene of *M. xenopi* cultures. This set of primers was able to discriminate between the pathogen and other mycobacterial species as well as non-bacterial strains; it also was able to detect as little as 3 fg of *M. xenopi* DNA.

Even with the success of PCR and other studies, research is being conducted to improve bacterial detection with biochemiluminescence (BCL) [37]. The main attractions for the assay technology include exquisite sensitivity, high selectivity, speed and simplicity. Although BCL has been used for analytical purpose for more than 40 years it is only recently that it has found widespread acceptance. In biomedical research, one of the most important applications of BCL is to estimate microbial numbers and to access cellular states. Recent advances in development of ultrasensitive micromechanical thermal detectors have led to the use of microcalorimetric spectroscopy for detection and identification of microorganism [38].

Fluorescence spectroscopy has a strong potential to effectively detect and identify bacteria and bacterial spores. Multi-excitation wavelength with fluorescence emission method is not new for bacterial identification. The fluorescence technique has been considered on various occasions for bacterial pathogen identification [39-49]. The fluorescence from bacteria and bacteria spores is due to emissions of intrinsic fluorophores [50]. There are significant changes in the fluorescence emission spectra from bacteria and bacterial spores from different species. The emission peak from a single excitation wavelength is often regarded as a broad peak without significant features. Therefore, bacterial information with fluorescence spectroscopy generally involves a double discrimination technique. The double discrimination means that the identification is made by varying two different parameters. For example, the double discrimination can involve aminopeptidase profiling [39], the two parameters involve aminopeptidase treatments followed by fluorescence measurements (a chemical and an optical parameter). Another example for double discrimination technique is generating a fingerprint (profiles) in the form of an excitation-emission matrix (EEM) [40]. Here, the fingerprinting data are acquired in the form of matrix of fluorescence intensity as a function of multiple exciting and emitting wavelengths [51]. In the second example, the two different parameters that are varied are the excitation and emission wavelengths (two optical parameters).

Other light scattering techniques have been considered, from time to time, for bacteria detection and identification. The work done by Bronk et al. [52] with polarised light has shown that the S_{34} element of Mueller matrix to be sensitive to bacterial size, shape and index of refraction. The exact shape and size of the bacteria vary during the life cycle (growing to a maximum size and then dividing). Light scattering have been very effective in monitoring the shape and size changes of bacteria.

In this research, we have selected fluorescence spectroscopy for several reasons; the fluorescence technique is fast [40]. There is no need to grow the bacteria in the presence of antigens to determine the species. This traditional method of identification generally takes several hours or more. It is possible to measure a fluorescence spectrum in less than 1 s. Fluorescence is a resonance process and has a large cross section. The excitation is from an absorption band of the protein. Compared with non resonance phenomena such as Raman spectroscopy, this method of detection is extremely sensitive, requiring only a few bacteria to give a representative spectrum. Since the fluorescence signal is shifted in wavelength away from the exciting signal, very low level detection is possible. This means that small sample sizes on low sample concentration can be used and it is still possible to obtain a good signal to noise ratio. The small sample size also decreases risk to laboratory personnel during the development and testing of the technique. Fluorescence has a large number of parameters (*e.g.* excitation wavelength, emission wavelength, and fluorescence lifetime). This affords several possibilities to tailor the technique to the problem. Fluorescence spectroscopy can also be used in remote detection. This has obvious military applications, especially in the event of bacterial warfare.

Fluorescence is a valuable tool in probing different bacterial spores/materials. The surrounding material influences the emissions from intrinsic fluorophores. Subtle differences in the emission spectra can, therefore, be used to identify the environment of the fluorophores. When studying bacteria, these environmental differences can be linked to the species of the microbes, or the growth stages of the bacteria. Analysis of whole bacterial cells and spores with this technique has given rise to unique protein fingerprints that can be used for identification at the species and strain level [43]. More and more reports are coming from different groups [53-56]. Scully et al. used coherent spectroscopy instead of fluorescence for the identification

of bacterial spores [57]. Identification of aerosol like bacterial spores has been a major area of research in the last decade. Fluorescence detection of biological with and without interference with indoor and outdoor samples studies [58-63] were in the forefront. The recent reports [62, 63] describe the use of laser induced breakdown spectroscopy (LIBS) for biological simulants.

Fluorescence quantum efficiency (QE) measurements of biological agents and simulants are needed for the development of both point sensors and standoff detection technique that use ultraviolet autofluorescence for detection and identification of bacterial agent. We can use these data to determine the limits of detection and the limits of identification for the algorithms. The sensitivity of a detection system is limited by the QE of the microorganism to be detected. An estimate of QE can be obtained if optical constants of the microorganism are known [42]; the QE can then be used as a design parameter for the detector system. Fluorescence cross sections of aerosolized *B. globigii* have been measured by Faris et al. in a flowing aerosol stream [64]. Fluorescence cross sections of more *Bacillus* samples were measured by Stephens [65] on a cloud of particles suspended in an electrodynamic particle trap. The most recent report of the QE measurements of biological samples also appeared in the literature [56].

1.3 Mathematical Models in Identification and Classification of Bacterial Spores

The complexity of the spectra from bacterial spores makes the interpretation of their spectral signature a difficult task. The current methods for identification and classification of biological samples have their shortcomings. A possible alternative is the use of mathematical model. Mathematical algorithms are used to manipulate data collected for a sample to investigate the interrelationship between the variables measured. One type of mathematical model is multivariate analysis. This includes a range of statistical techniques used to analyse more than two variables simultaneously. Examples of multivariate techniques include principal components analysis (PCA) and factor analysis, which both attempt to represent the original data with reduced number of variables. Other multivariate techniques attempt to group variables based on their similarities/dissimilarities (cluster analysis) or assign variables to one of two groups (discriminant analysis).

Most algorithms to date use qualitative and, to a lesser extent, statistically selected variables with binary or probability based classification schemes. The development of a physically based model that uses biochemical and morphological features that are related to the measured fluorescence emission spectra has been hampered by the fact that fluorescence spectroscopy of microorganism is greatly affected by the absorption and scattering of excitation light and the emitted light, making interpretation of the measured spectral information challenging [66]. The problems in high dimensional spectroscopy are the partial or complete overlap of signals from the different constituents in the sample. Hence the classification, detection and analysis of the spectra are often difficult. In high dimensionality, the parameter space (or feature space) is very sparse. The efficiency of classifiers decreases with high dimensionality of the space. This fact called Bellman's "curse of dimensionality" [67] imposes a pre-processing step in the analysis.

Many techniques have been developed for dimensionality reduction. Briefly, we will mention: feature selection or feature reduction, i.e. building a new reduced set of features from the original large set. This can be done through linear methods such as PCA [68-70] which has found many applications in optical spectroscopy. PCA

has been used for the analysis of ultraviolet-visible (UV/VIS) fluorescence data [71, 72] and its usefulness has also been demonstrated for liquid chromatography [73] and time resolved infrared spectroscopy [74]. Identification of bacteria by using PCA in analysing their fluorescence spectra has been presented before by Leblanc et al. [75]. The report of most recent use of multivariate analysis with biological spectroscopy detection appears in the literature [62, 63], with LIBS fluorescence in the UV-VIS region to distinguish between biological and interferences.

PCA is a powerful tool for the extraction of low dimensional information from heavily over determined high dimensional data. Both the experimental noise and hidden dependencies among a set of experimental conditions may confound the inferential process. The PCA technique is based on the assumption that the variance of spectral data may be used as a measure of spectral content. PCA is an exploratory multivariate statistical technique for examining relationships among several quantitative descriptors or variables [76, 77]. PCA is a decomposition technique that produces a set expression pattern known as principal components (PCs). Linear combinations of these patterns can be assembled to represent the behaviour of all of the objects in a given data set [78]. Briefly, each PC is a linear combination of the original variables with coefficients equal to eigenvectors of the covariance or correlation matrix. The eigenvectors are orthogonal, so that the PCs represent jointly perpendicular directions through the space of the original variables.

If the spectra of the biologics have repeatable differences, the data points from each sample should cluster together, showing successful discrimination of a sample-specific fluorescence signature. It is quite common for PCA to be used to identify variables that contribute to a particular process and then use cluster analysis to group samples which are affected by each process [79, 80]. Cluster analysis is a method for displaying the similarities and dissimilarities between a pairs of objects in a set. Clustering techniques have been applied to a wide variety of research problems [81, 82]. In general, whenever one needs to classify a complex of information into manageable meaningful set, cluster analysis is of great utility. In fact, we think that no clustering method can be efficient in all the situations that can be encountered in practice.

Many methods permit to build classes and or partitions on a set of data. The most classical ones optimise some criteria over a set of partitions with a given

number of classes [83]. Criteria commonly used are the split of classes, the diameter of the partitions, and the sum of (squared) distances between elements in different classes [84]. A projection into a Euclidian space using a multidimensional scaling or a factorial method is often commonly realized and then a centre method (k medoids) is also applied to the set of the project to the point. More rarely, a sequential clustering method is applied [85]. It consists in building a class as homogeneous as possible, then to iterate the procedure on the data set X minus this class, until there is no class sufficiently homogeneous.

1.4 Overview of the Thesis

The work described in this thesis has grown out of our laboratory developing project *"Pathogen Detection and Identification Using Two-dimensional Fluorescence Spectroscopy*". As cited in Weichert et al.'s article [86], fluorescence spectroscopy is a powerful tool for the detection and identification of bacteria and bacterial spores. Whether used for remote detection by a mobile light detection and ranging (LIDAR) [87] or with a small, unattended ground sensor [88], a database of fluorescence spectra of various biological and non-biological spores is necessary.

It was the goal of this study to investigate the feasibility of using the fluorescence identification procedure and mathematical models for identification and detection of biological pathogens. The objectives of this study have been of considerable scope and depth. While there has been one main goal, there have been many smaller supporting sub-objectives. During the course of the project, our goals were modified, as more was understood about the nature and relative importance of the tasks at hand. Some of the goals were accomplished, while others were deemed unfeasible within the scope of the current project. In addition to the main objective, many sub-tasks were outlined in the original proposal. The tasks and goals are arranged below in order of equal or decreasing importance:

a) Evaluate the use of two-dimensional spectroscopy for rapid identification and detection of biological pathogens

b) Measure the QE of *Bacillus* spores

c) Clustering of biological pathogens using PCA of fluorescence data

d) To find the changes in the autofluorescence between dried and wet *Bacillus* spores using EEM

The sub objectives are to draw a conclusion regarding the application of two-dimensional spectroscopy and PCA to determine the quality parameters that are important in identification of *Bacillus* spores. Using *in vitro* techniques, we investigated different bacterial spores found in the different background and determine if the fluorescence spectra are unique and can be identified. In retrospect, the tasks and goals described above may have been overly ambitious, given the amount of time and funding available. Therefore, they were modified according to their overall impact upon the project's main objective.

In this research, we followed the techniques of generating fluorescence fingerprints [51] to investigate biological samples, in particular bacterial spores for detection and identification [43, 44, 48]. Preliminary work borrowed from the US Department of Defence studies, on the remote detection and identification of biological samples, was the foundation for this project [41, 42]. The fluorescence from bacteria and bacterial spores is due to the emissions of intrinsic fluorophores [39, 40, 89-91]. The number and the environment of the fluorophores influence these emissions. Thus, the fluorescence probes the interior and composition of the biological samples. In earlier studies; several species of bacteria and bacterial spores were studied with fluorescence excitation and emission spectroscopy [41]. With dilute room temperature suspensions, reproducible characteristics in the fluorescence spectra from several different species of bacteria were found. These characteristics are generally independent of the conditions of growth and thought to be useful as a rapid mean of species identification. However the exact shape and size of this peak change with the environment of the tryptophan. In general there is an excitation peak near 280 nm, with a strong emission peak near 340 nm. This peak is primarily due to tryptophan [92-94]. These characteristic changes in the environment are the key to the species discrimination with fluorescence spectroscopy. Differentiation between different bacteria becomes more reliable, if fluorescence spectra are measured at several different excitation wavelengths. The fluorescence spectra of fungal spores and *B. subtilis* vegetative cells are nearly identical at 266 nm excitation, but there is a distinct difference in the spectra if excitation occurs at 351 nm [55].

The main purpose of this project, the detection and identification of a large number of biological simulant, has been accomplished. We have measured two dimensional excitation-emission spectra of all of the biological samples under study in dry spore forms and in aqueous suspension. We used EEM technique to visualise fluorescence fingerprints of spores investigated. A fluorescence EEM is a two dimensional contour plot that displays the fluorescence intensities as a function of a range of excitation and emission wavelengths. Each contour represents points of equal fluorescence intensities**.** The multivariate statistical techniques of PCA and partition around medoid (PAM) and hierarchical cluster analysis were performed on

both washed and unwashed *Bacillus* spores in order to group these samples based on their spectral fingerprints.

In addition to the qualitative fluorescence spectra, we have calculated the QE of some of the species and compared with the standard sample of anthracene in ethanol. The fluorescence of a biological molecule is characterised by its QE and its life time [91]. The QE is the ratio of the number of photons emitted to the number absorbed. The life time is defined as the average time the biological molecule spends in the excited state prior to return to the ground state. Fluorescence spectroscopy is the measurement and analysis of various features that are related to the fluorescence QE and or life times of a biological molecule. The fluorescence intensity of a biological molecule is a function of its concentration/number, its extinction coefficient (absorbing power) at the excitation wavelength, and its QE at the emission wavelength [95]. We checked for changes in the fluorescence spectrum before and after washing the spores. We noticed a difference in the fluorescence intensity for washed and unwashed bacteria spores with unwashed fluorescing more strongly than the washed due to a component of growth medium.

Chapter 2

Introduction to Ultraviolet-Visible Spectroscopy and Data Analysis

This chapter is a general review of UV-VIS spectroscopy and data analysis by PCA, that have found much application in identification and classification of biological samples investigated in this study. In the first part, we will briefly introduce UV spectroscopy and in the second part a brief account of PCA and clustering techniques is given, that have found much importance in data reduction and classification of microorganisms.

There are several good books available for a detailed discussions of general theory of luminescence, including the two monographs [91, 96]. Luminescence measurements have had widespread applications in the analytical studies of large, unsaturated, highly conjugated, organic molecules such as polynuclear aromatic compounds. A complete review on fluorescence techniques for studying protein structure was given by Eftink [97], as well as an earlier review by Beechem and Brand [98] specifically treating tryptophan fluorescence in proteins.

For data analysis, PCA is now a well established tool for interpretation chemical and biological data. The application of PCA analysis in bacterial identification was used in several research projects [24, 62, 63]. All basic features of the method are thoroughly documented together with numerous chemical examples [69, 99]. Hence, in this section we provide only a brief description of the features necessary for the interpretations in the present work.

2.1 General Principles of Spectroscopy

Optical spectroscopy is the study of the interaction of electromagnetic radiation with matter that occur at the UV, VIS, near-infrared (NIR) and infrared (IR) wavelengths. All chemicals and biological samples absorb energy from at least one region of the spectrum of electromagnetic radiation (Figure 2.1). The energy at which absorption occurs depends on the available electronic, vibrational and

rotational energy levels of the molecule. When absorption is from the UV-VIS region of the spectrum, transitions occur between electronics energy levels. It is these transition that form the basis of UV-VIS spectrometry.

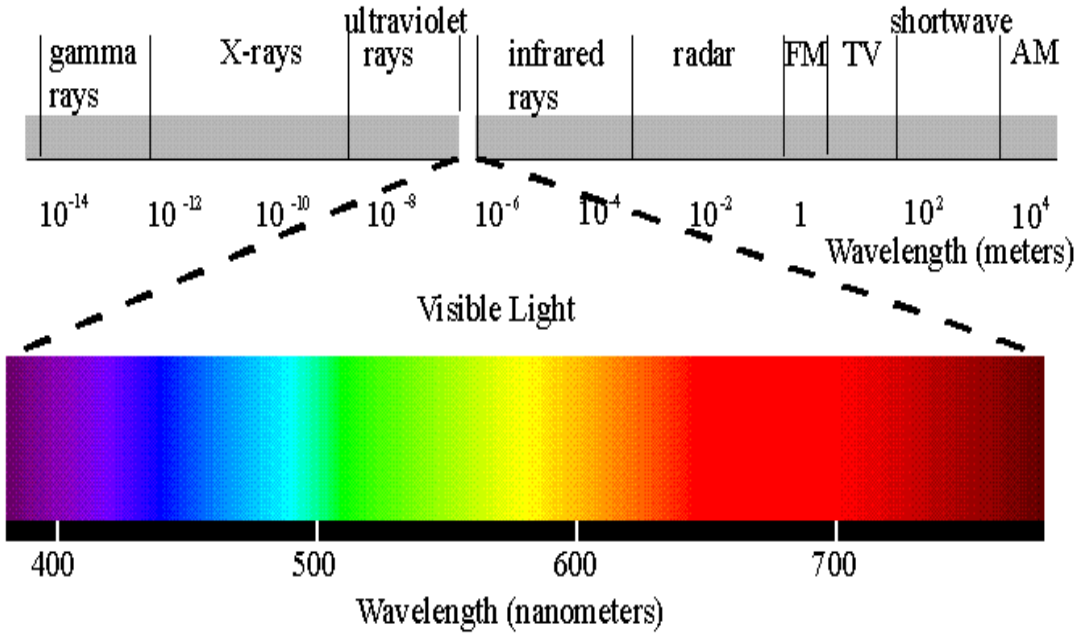


Figure 2.1 The electromagnetic spectrum

(Retrieved from www.yorku.ca/ eye/spectru.htm)

The electromagnetic spectrum can be defined in terms of wavelength or frequency. The wavelength of light is defined as the distance between two crests, or troughs of a wave and is usually expressed by the symbol λ. The unit used in the UV-VIS region of the spectrum is a nanometer (nm). The wavelength of light is related to energy in terms of speed of light, c, and Planck's constant, h

$$E = hc / \lambda \qquad (2.1)$$

This relationship shows that the shorter the wavelength, greater the energy of light.

2.1.1 Transmittance and Absorbance

Molecules absorb light of particular wavelength. When the energy that corresponds to a wavelength equals the energy different between the ground and exited state, the light is absorbed. This selective absorption of light is the basis of UV-VIS spectrometric analysis. Absorbance of a sample is obtained indirectly by measuring the light which is not absorbed by the sample, i.e., the transmitted light (Figure 2.2). The Beer-Lambert law [95] relates the absorption intensity to the incident and transmitted intensities, as well as concentration of the absorbing material such that

$$Abs = - \log (I/I_0) = \sigma c l \qquad (2.2)$$

where σ is the absorption coefficient of the absorbing substance, c is the concentration of absorbing substance and l is optical the path length of the cell. The Beer-Lambert law is the basis for all quantitative measurements in UV-VIS spectrometry as it relates absorbance directly to concentration.

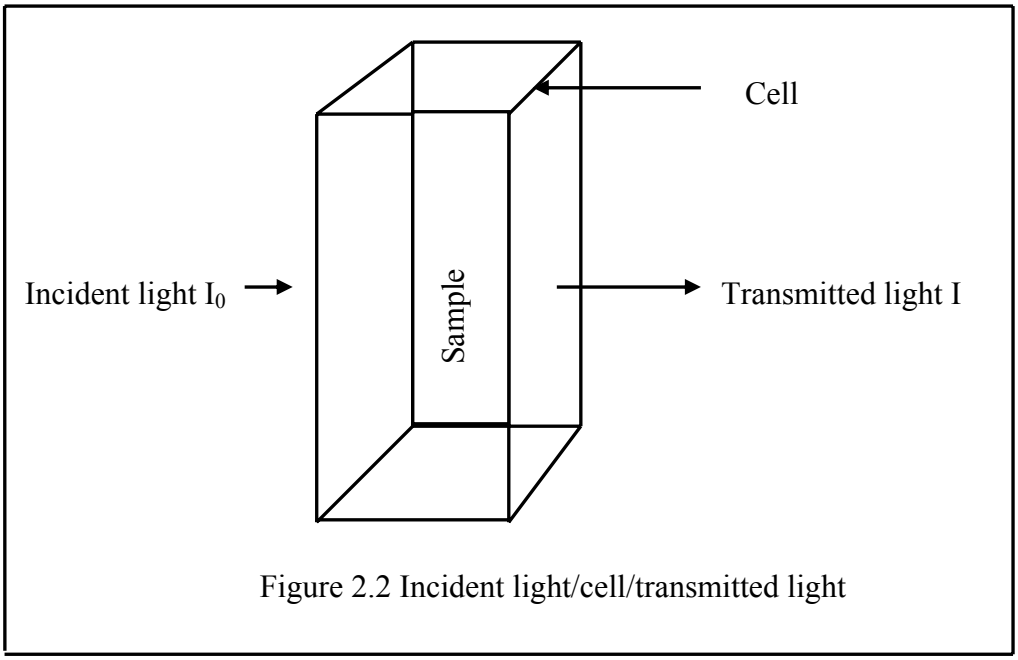


Figure 2.2 Incident light/cell/transmitted light

2.1.2 Luminescence

Optical spectroscopy probes the energy levels of a molecule [95]. The energy level of a molecule is defined as its characteristics state, which is related to the molecular structure of the molecule and to the energetics and dynamics of any chemical processes that the molecule may undergo. An electron transition induced by absorption of photon by a molecule can be represented by a scheme of energy-levels [91, 96, 100-103].

Figure 2.3 displays an energy level diagram with ground (s_0) and excited electronic state as well as vibrational energy levels within each electronic state of a molecule [91]. On the left part of the figure is shown a set of singlet states, s_k, i.e., the states with zero total spin. On the right part of the figure is shown the scheme of triplet states, T_k, i.e., the states for which the total spin of the molecule equals 1. Triplet state T_k possesses the same electronic configuration as a corresponding singlet state s_k. These states differ by the relative orientation of the spins of two electrons on the outer orbital. According to Hund's rule, the triplet energy level is always lower than the corresponding singlet one.

Luminescence is the broad term that describes the emission of photons from electronic excited states of molecules. There are two major categories of luminescence: fluorescence and phosphorescence. Fluorescence is the emitted light due to transitions between singlet states of a molecule. Phosphorescence is the emitted light due to transitions between triplet and singlet states. Luminescence may be interpreted as the detectable loss of energy by an excited molecule. Fluorescence is simply regarded as a very rapid (1-10 ns) emission of light in contrast to phosphorescence, which is a delayed release of the absorbed energy. In more subtle theoretical terms, fluorescence is said to originate from a singlet excited state of a molecule and phosphorescence from a triplet excited state. Due to the longer lifetime of a triplet excited state, collisional deactivation is predominant in a solution at room temperature and phosphorescence is rarely observed.

2.1.3 The Fluorescence Phenomenon and Frank-Condon Principle

When a molecule is illuminated at an excitation wavelength that lies within the absorption spectrum of that molecule, it will absorb the energy and be activated from its ground state (s_0) to an excited singlet state (s_1), with the electron in the same spin as its ground state (see the Figure 2.3). The molecule can then relax back from the excited state to the ground state by generating energy, either nonradiatively or radiatively, depending on the local environment.

It was found empirically (Kasha's rule) that in condensed media the excitation energy of the s_2 state (Figure 2.3), or any other higher state, quickly dissipates due to the inelastic collisions and the molecule nonradiatively reaches the zero vibrational level of the s state. Nonradiative transitions between electronic states of the same spin multiplicity are called *internal conversion*. In the internal conversion process the energy is transferred from a point on the potential energy hypersurface of a state (s_2, for example) to the potential energy hypersurface of s_1. After that the excess vibrational energy, dissipates in the course of *vibrational relaxation* and the system reaches the thermally equilibrated state of s_1 molecules. In this process the excitation energy is transferred to the external medium. It is assumed that the rate constants of internal conversion and vibrational relaxation exceed 10^{12} s^{-1} (Figure 2.3) [91].

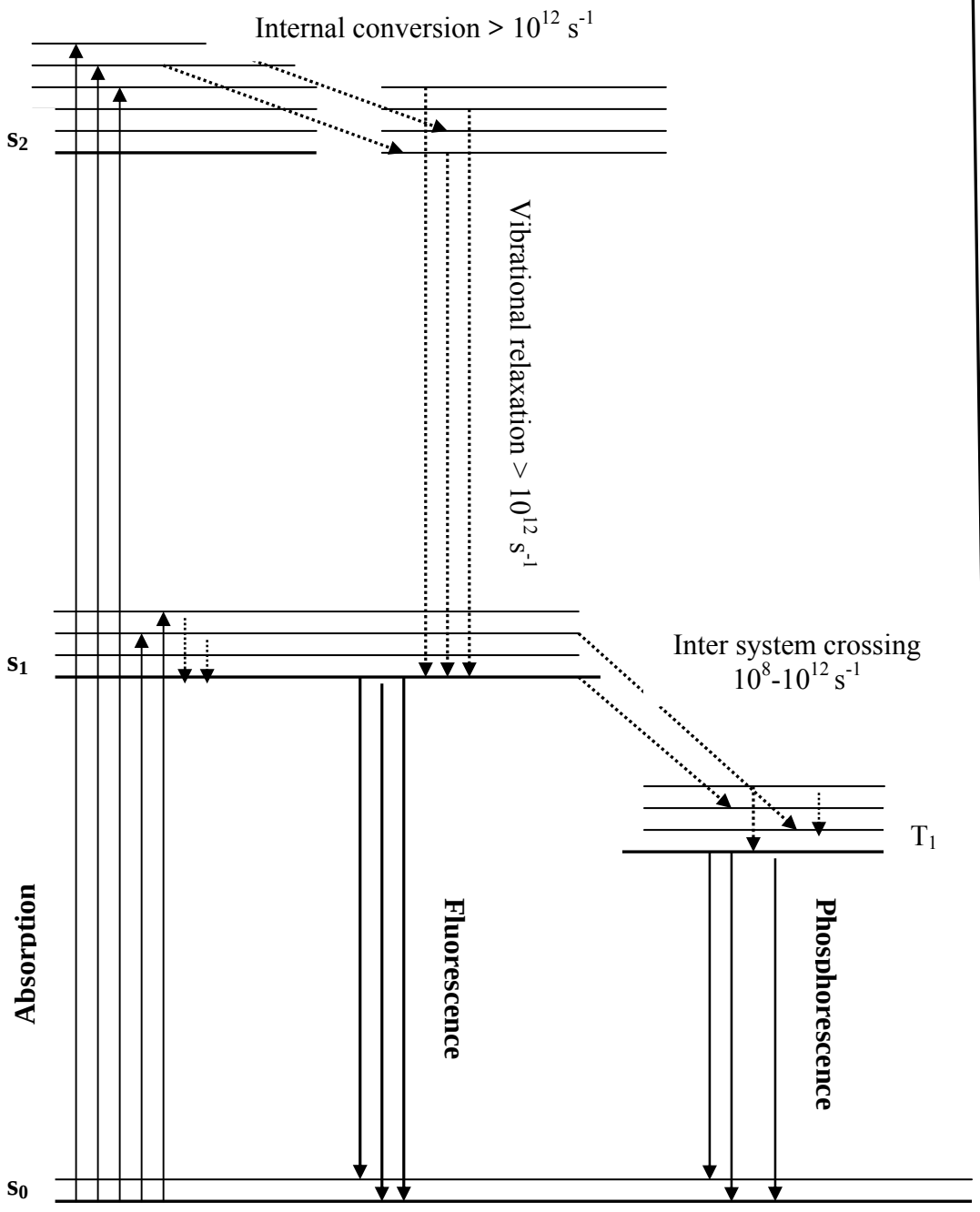


Figure 2.3 Electronic energy level diagram and the transition affecting fluorescence and phosphorescence. s_0, s_1 and T_1 are unexcited singlet, excited singlet (fluorescence) and triplet (phosphorescence) chromophores states.

Radiative transitions in a molecule between levels of the same multiplicity give rise to *fluorescence*. It is evident that the emission occurs as a result of the transition of the molecule from s_1 to the s_0 state (see the Figure 2.3). The Frank-Condon principle is formulated as follows [96, 100, 101]. Electron transitions are so fast (10^{-15} s) in comparison to the movements of nuclei (10^{-12} s), that during electron transitions nuclei can change neither their velocities nor their positions. The principle reflects the fact of impossibility of a fast conversion of electronic energy into energy of nuclear vibrations.

The phenomenon of fluorescence displays several general characteristics for a particular molecule [91]. First, due to the losses in energy between absorption and emission that occurs as a result of nonradiative transitions, fluorescence occurs at emission wavelengths that are always red-shifted (Stoke' shift) relative to the excitation wavelength. Second, the emission wavelengths are independent of the excitation wavelength. Third, the fluorescence spectrum of a molecule is generally a mirror image of its absorption spectrum (see Figure 2.4 for illustration).

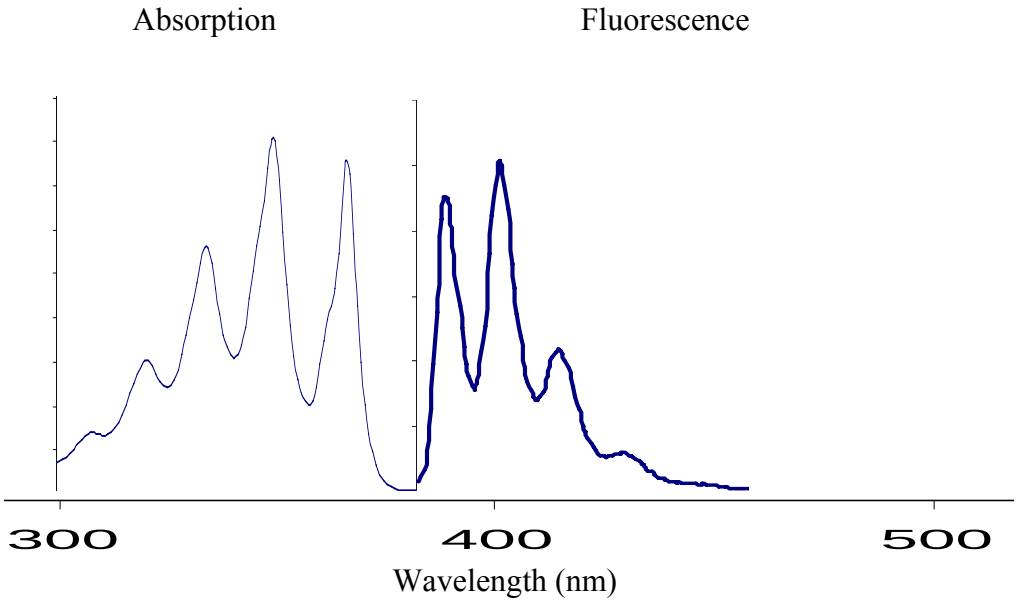


Figure 2.4 Illustrations of the absorption and fluorescence emission spectra of an anthracene solution as approximate mirror symmetry

2.1.4 Fluorophores and Biological Fluorescence

Microorganisms are known to be capable autofluorescence due to the presence of fluorophores [50]. Detection of biological materials is possible as there is a wide variety and high concentration of biological components that exhibit intrinsic fluorescence: Nicotinamide adenine dinucleotide and other reduced pyridine nucleotides, lumazines, pterins and flavoproteins. Nucleic acid polymers, proteins and various lipids exhibit higher energy fluorescence, making these markers potentially useful for the detection of fingerprints. The difference in intrinsic fluorescence emission of biological materials can be differentiated. Many biological materials exhibit similar or indistinguishable components. Simultaneous excitation of a sample, with multiple energies characteristic of the excitation for fluorescent components with the subsequent collection and detection of emitted, reflected and scattered light energies (both associated with and independent of the fluorophores, respectively), is fundamental for the detection of biological material on a surface.

Table 2-1 lists the biological molecules that exhibit endogenous fluorescence, along with their excitation and emission maxima [91, 104, 105].

Table 2-1 Excitation and emission maxima of biological molecules that exhibit endogenous fluorescence (Extracted from Lakowicz, Principles of Fluorescence Spectroscopy)

Endogenous fluorophores	**Excitation maxima (nm)**	**Emission maxima (nm)**
Amino acids		
Tryptophan	280	350
Tyrosine	275	300
Phenylalanine	260	280
Enzymes and coenzymes		
FAD, flavins	450	535
NADH	290, 351	440, 460
NADPH	336	464

These fluorophores include amino acids, structural proteins, enzymes, co-enzymes etc. Their excitation maxima lie in the range 250-450 nm, whereas their

emission maxima lie in the range 280-700 nm. Among these, amino acids are the basic structural units of a protein and proteins play crucial roles in virtually all of the biological processes [106]. Three amino acids (tryptophan, tyrosine and phenylalanine) with aromatic side chains are fluorescent. At excitation wavelengths longer than 295 nm only tryptophan is fluorescent [91]. From 280 to 295 nm, both tyrosine and tryptophan are fluorescent; however, energy transfer from tryptophan to tyrosine is common. At excitation wavelengths shorter than 280 nm, all three amino acids can be excited [91]. Figure 2.5 shows the amino acids group and their absorption [107].

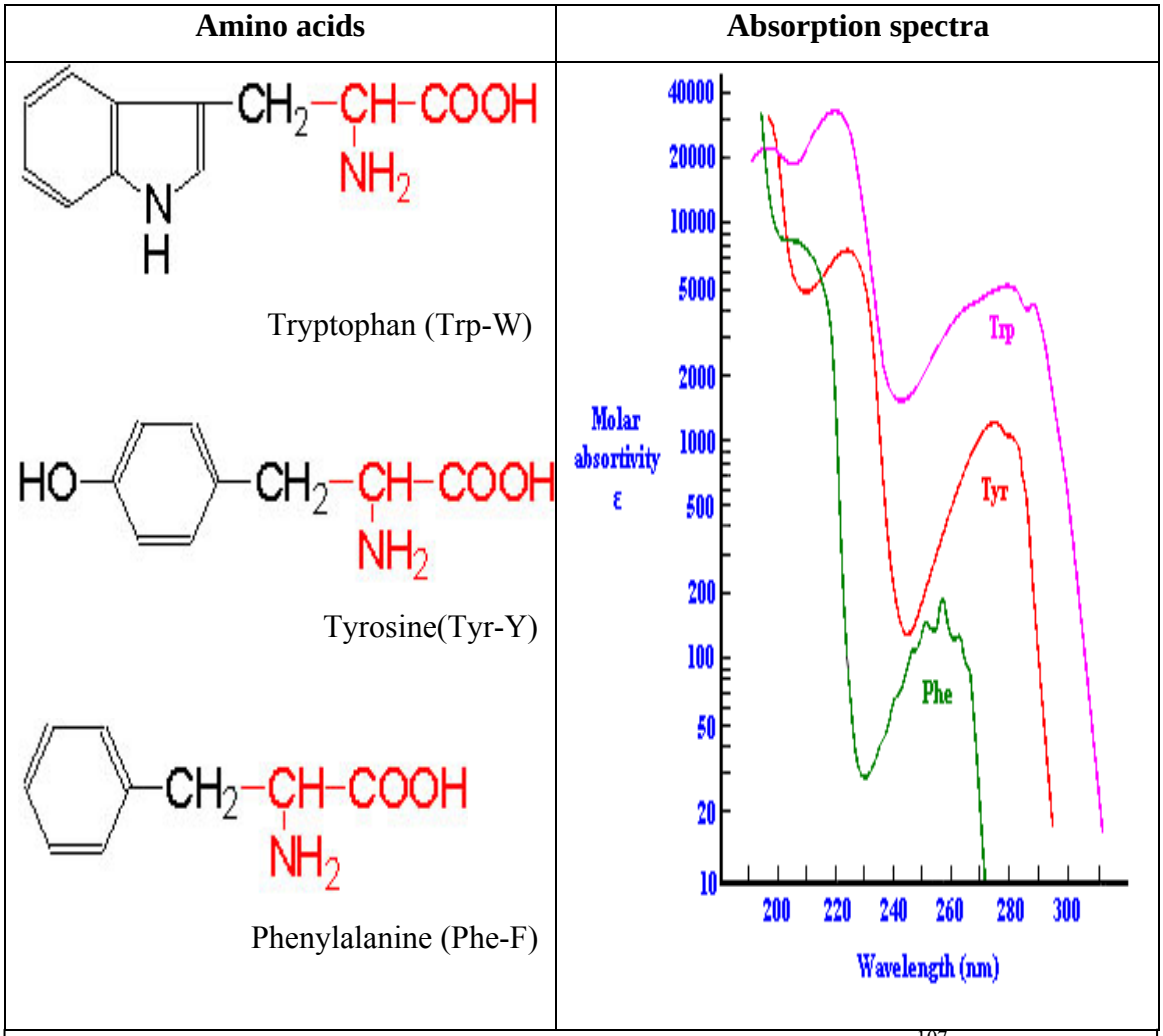


Figure 2.5 Amino acids group and their absorption spectrum [107]

The technique of UV-VIS spectroscopy is the most common in analytical science and is used in a wide variety of application, including clinical, environment, pharmaceutical, education and research. Since spectrometry carries the ability to select a measuring wavelength, it is widely used when both selectivity of measurement and sensitivity are required. The applications of UV-VIS spectrometry are numerous and the following a brief discussion is given for its biological applications especially for proteins.

UV-VIS spectroscopy is a valuable tool in probing different materials. The surrounding material influences the emissions from intrinsic fluorophores. Subtle differences in the emission spectra can, therefore, be used to identify the environment of the fluorophore. When studying bacteria, these environmental differences can be linked to the species of the microbes or the growth stages of the bacteria [108, 109]. The protein molecule, with respect to its spectroscopic properties, may be considered as a complex system of chromophore groups, which differ in structure and position of spectra. The protein absorption spectrum is, at first approximation, the superposition of spectra of chromophores composing the protein molecule. The most extensive research effort has been focused on the 190-220 nm spectral range, in which the peptide group absorption and the absorption of many amino acid residues are observed, as well as on the range of approximately 280 nm in which tyrosine and tryptophan absorption occurs. Both ranges are used for conducting experiments in protein spectroscopy. In emission spectra the contribution of the tryptophan chromophore is the most important one.

L- tryptophan is a nonpolar amino acid with aromatic (indole) side chain (see Figure 2.5). This aromatic R-group is hydrophobic and very bulky in size. This side chain absorbs light in the near UV region. UV absorption can be used to quantify the concentration of proteins. Considerable work has been done on the use of UV radiation of biological aerosols to detect the fluorescence emission from aromatic compounds inherent to bacteria [64]. Tryptophan is the target molecule of greatest interest in detecting microbes. Fluorescence of tryptophan has been extensively used to study structure and conformational change of proteins as well as binding of substrates of proteins [98, 110]. Tryptophan in solution in water has two excitation optima, one about 225 nm and one about 280 nm, and emission maxima for either

excitation occurring with peaks about 350 nm. When tryptophan is bound in a cell or spore the emission maxima are shifted to shorter wavelength and the peak is near 290 nm.

2.1.5 Comparison of Fluorescence Quantum Efficiency

The fluorescence QE is the ratio of photons absorbed to photons emitted through fluorescence. Using the Beer-Lambert law [95] one can express fluorescence intensity of an optically dilute, homogeneous medium (sum of the optical densities at the excitation and emission wavelength, everywhere is less than unity) [91] as a linear function of the concentration of the fluorophores in that medium. The number of quanta emitted per unit time is proportional to the number of quanta absorbed per unit time and fluorescence QE:

$$N_f = I_0 (1-10^{-D}) Q_f \tag{2.3}$$

where I_0 is the intensity of incident light; D is the absorbance of the solution, which is the product of the extinction coefficient of the molecule ε, the concentration of the molecules c, and the optical path length l ($D = \varepsilon cl$); Q_f is fluorescence QE which is the ratio of the number of quanta emitted from an excited state to the number of quanta absorbed during the transitions from the ground to this excited state per time unit. If concentration c is low and hence D is small, then the equation (2.3) becomes

$$N_f \sim 2.303 \varepsilon c l I_0 Q_f \tag{2.4}$$

Although fluorescence spectroscopy of optically dilute, homogeneous media is well understood, fluorescence spectroscopy of biological samples is complicated by their highly absorbing and scattering properties [66].

The most reliable method for recording QE is the comparative method of Williams et al. [91, 111] which involves the use of well characterised standard samples with known QE values [112]. A simple ratio of the integrated fluorescence intensities / optical densities of the two samples will yield the ratio of the QE values. Since QE for the standard sample is known, it is trivial to calculate the QE for the test samples using the formula [91, 111],

$$Q_{f\,(sample)} = Q_{f\,(standard)}\,[Slope_{\,(sample)} / Slope_{\,(standard)}] * n_1^2/ n_2^2 \qquad (2.5)$$

where the term Slope is the slope of the plot of the integrated fluorescence at a selected excitation wavelength versus the absorbance of the test sample as well as the standard sample, n_1 and n_2 the refractive index of the sample and the standard sample respectively.

2.2 Fluorescence Excitation-Emission Spectra via the Two-Dimensional Matrix Method

In using fluorescence as an analytical technique, one is often confronted by a sample containing an unspecified number of fluorescent species. If the components are well separated spectrally, it is possible by trial and error to selectively excite one or more of the components and to make reasonable interpretations of the data in terms of known fluorescence spectrum. Some excellent examples of this technique of selective excitation can be seen in the monograph by C.A. Parker [102]. The advantages of multidimensional data sets have been realised due to the availability of small, yet powerful, computers. The very nature of such large data sets makes it impossible to produce the information manually. Numerous data reduction technique and algorithms have been introduced to aid the analyst in filtering through data sets. Techniques such as factor analysis, pattern recognition and PCA have been widely used to reduce the bulk of data to a smaller set of essential elements [69, 113].

Two-dimensional matrixes have found many applications, especially in rank annihilation factor analysis [114-118]. Warner and co-worker have exploited fluorescence EEM in the analysis of mixtures of fluorescent compounds [51, 119]. The analysis utilises the fact that each fluorescent component is characterised by a unique dependence of its fluorescent intensities upon two distinct parameters, the excitation wavelength λ_{exc} and the observed emission wavelength λ_{em}. The data matrix consists of an EEM, [**M**], whose elements M_{xy} represent the fluorescent intensities measured at wavelengths λ_{em} for excitation at λ_{exc} [120]. For dilute mixtures, these intensities depend upon the sum of the product functions associated with each component [95]. If the wavelengths are properly sequenced, then the row of **M** represents fluorescence

spectrum at a particular wavelength of excitation, while a column represents a fluorescent excitation spectrum at a particular monitoring wavelength.

In the present application, EEM can be used to plot two-dimensional (2D) fluorescence spectra of *Bacillus* spores for detection and identification. The input matrix **M** is obtained by measuring fluorescence spectra at systematically increasing excitation wavelengths so that each row vector of **M** is fluorescence spectrum (not normalised) at a single λ_{exc}, and each column of **M** is an excitation wavelength at a single emission wavelength λ_{em}. Thus for a single component optically dilute sample, the matrix can be expressed with an adequate approximation [121] as

$$\mathbf{M} = \sigma\, \mathbf{XY} \tag{2.6}$$

where **X** and **Y** are column and row vectors corresponding to normalised fluorescence- excitation and fluorescence spectra of the component respectively, and σ is a scaling factor related to the relative concentration of the components. Note that this EEM technique has certain features in common with the techniques of PCA which will be explained in a separate section of data analysis.

2.3 Fluorescence Instrumentation

In order to perform quantitative spectroscopy of bacterial spores, the fluorescence emission spectra need to be measured. The instrumentation of total luminance spectroscopy ranges in complexity from computer controlled conventional fluorometers [122-125] to the video fluorometer [119, 126-128]. In Figure 2.7 the general components of a conventional fluorophotometer (SLM 8000C, Urbana, IL) is shown with two single monochromator configurations. The light source, generally a xenon lamp, provides photons for excitation of samples. It has a continuous output in the UV and VIS wavelengths. The unique properties of laser radiation that make it useful for a fluorescence light source have been reviewed [129]. In most fluorometers the excitation light is passed through a wavelength selector, the monochromator. Various types of filters and diffraction gratings have been used as monochromator. Instruments that use diffraction gratings can be used to scan a range of wavelengths. The photomultiplier tube (PMT) is the typical detector used in conventional

fluorometers. The PMT is sensitive to UV and VIS light. The geometry of the instrument shown in Figure 2.6 is typical of fluorometer. The emission light is monitored 90 degree relative to the excitation light.

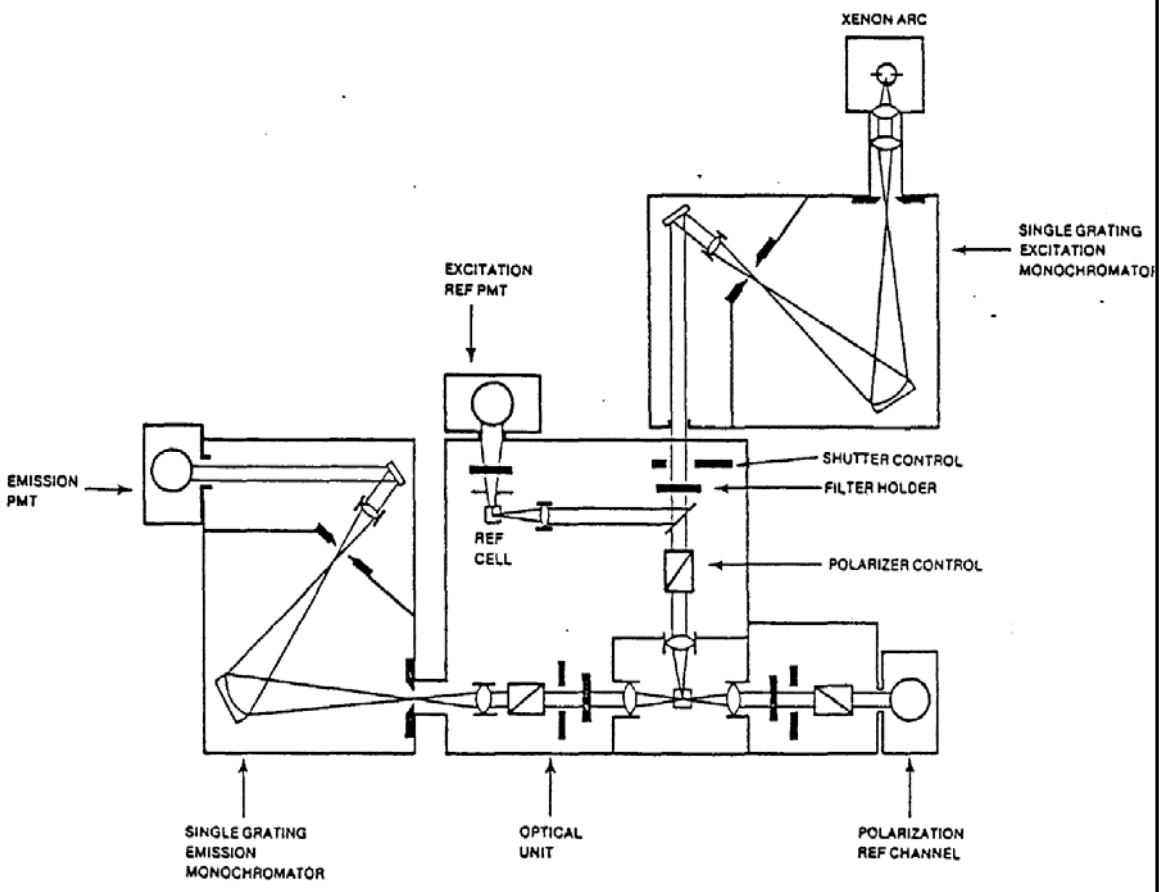


Figure 2.6 Components of conventional fluorophotometer is shown with two single monochromator configurations (Retrieved from operational manual version 4.0 of SLM 8000C, Urbana, IL)

2.4 Data Analysis using Principal Components Analysis and Cluster Analysis

A general theory of PCA model is given in Appendix 7.4. PCA was first introduced by K. Pearson [130], developed independently by H. Hotelling [131] and hence, sometimes, it is called Hotelling transform. PCA plays an important role in the remote sensing applications such as change detection, classification, enhancement and data detection. Both the experimental noise and hidden dependencies among a set of experimental conditions may confound the inferential process. It is non trivial to eliminate either of these complicating factors. It may be beneficial to pre-process the data before analysis in order to identify the independent information content of different experimental conditions. The PCA technique is based on the assumption that the variance of spectral data may be used as a measure of spectral content [132]. PCA is an exploratory multivariate statistical technique for examining relationships among several quantitative descriptors or variables [76, 77, 133]. This approach is especially valuable in exploratory analyses to assess resemblances and differences among samples under study.

The standard PCA scheme in spectroscopy represents each experimental spectrum as a vector $\mathbf{X}_x$ (x = 1….n) in the p dimensional space, where p is the number of data points in each spectrum and n is the number of spectra. The data matrix **M,** of the dimension n by p, is constructed from all the data sets. Diagonalization of the second moment matrix

$$\mathbf{Y} = \mathbf{M}^{\mathbf{T}}\mathbf{M} \tag{2.7}$$

where $\mathbf{M}^{\mathbf{T}}$ is the transpose of **M,** gives n nonzero eigenvalues and their corresponding eigenvectors. By finding the n eigenvectors and eigenvalues of **M,** and by arranging the eigenvectors in the descending order of eigenvalues, one can construct an ordered orthogonal basis. Each original spectrum can be represented by a linear combination of n basic vectors, or *components*. By selecting the eigenvectors having the *largest eigenvalue* and neglecting those with the smallest ones, one can represent all the data sets by using a linear combination of just a few (n_c) PCs (eigenvectors). It represents the samples/variable in eigenvector space with number

of coordinates corresponding to the number of PCs [68, 69]. Because $n_c < n < p$ (in most practical cases $n_c << p$), the PCA provides a convenient way to reduce the dimension of the representation. Usually, the initial step in PCA is some preprocessing of raw data. With fluorescence intensities, normalization is necessary to outweigh the effect of different peak heights due to different signal to noise ratios in the spectra. Since the largest peaks also have the largest absolute variance, systematic variation in small peaks can be masked in PCA. This problem can be solved by using logarithmic or auto scaled data.

A decision needs to be made as to how many of the PCs will be used to investigate the relationship between the original variables. A number of approaches are recommended in various texts [134-136], but essentially the analyst can choose the number of PCs to retain. Some of the available methods include: 1) choosing the total amount of the variance to be accounted for by the PCs. Usually this is some value greater than about 60% of the total variance. In some cases a value as high as 90% is chosen [136]; 2) choosing to retain PCs above the "elbow" of the scree plot. The scree plot is a plot of variance versus component number (see Figure 4.25) [136]; 3) when conducting PCA on correlation matrix, choosing to retain PCs that have variance greater than one, because PCs with variance less than one explain less variance than any of the original variables [134, 135]; 4) choosing a convenient number of PCs to retain. Typically two PCs are retained because these plot easily on a two dimensional plot making interpretation of the results easier [134, 135]. The first and second methods were used in this research.

After PCA is used to reduce the dimensionality of the data, we must determine if we are still able to group the measurements correctly. There are several ways to cluster data. We considered hierarchical and partitioning algorithms. These partition a data set of *n* objects into *k* clusters. The number of clusters, *k*, must be given. Two common methods are k-means and k-medoids (more commonly called Partition Around Medoids or PAM) [137]. The k-means divides the objects into *k* clusters. The centroid of each cluster is found and the distance from each centroid to each object is computed. The objects are regrouped around the closest centroid. A new centroid is calculated and the process iterates. The distance between an object and a centroid can be found many ways. Typically, the Euclidean distance is used.

While the k-means method of clustering is generally considered to be very efficient, in that the number of iterations required is small, the method often terminates at local minimae instead of finding the global minimum. Also, k-means does not handle out-liers or very noisy data. PAM is much more effective with noisy data or out-liers. The method is very similar to k-means. Instead of using the centroid of a cluster, the distances are calculated to the object in the cluster that is closest to the centroid. PAM also works well for small data sets.

When clustering together data sets, we can question the strength of the clusters. We can question whether one cluster should be divided into two clusters or if two clusters should be combined into one cluster. The PAM program also plots the silhouettes of the cluster strength.

The cluster strength is defined as

$$S_i = [B_i - A_i]/\text{Max}\,[A_i, B_i] \tag{2.8}$$

where A_i is the mean distance from object *i* to all the other objects in the same cluster and B_i is the average distance from object *i* to objects in the closest, different cluster. If a cluster is very small and far from the other clusters, S_i will approach 1.0. If a cluster is very large and near to a neighbouring cluster, then S_i will approach 0. Generally, cluster strengths less than approximately 0.5 should be questioned. However, there is no absolute cut-off where a cluster is deemed inappropriate. Distances can be calculated in many different ways, we chose to use Euclidian distances in our calculations.

Chapter 3

Experimental Section

This chapter deals with the central part of this research of detecting and identifying *Bacillus* simulants using 2D spectroscopy. In Section 3.1 we will explain the fluorescence spectral measurements and visualization by EEM of various samples investigated. For convenience, the data analysis part of this research will be presented in results and discussion chapter.

3.1 Spore Samples/Simulants/Interferents and Substrates

Most *Bacillus* species are versatile chemoheterotrophs capable of respiration using a variety of simple organic compounds (sugars, amino acids, organic acids). A few species, such as *Bacillus megaterium*, require no organic growth factors; others may require amino acids, B-vitamins, or both. The majority are mesophiles, with temperature optima between 30 and 45 degrees, but the genus also contains a number of thermophilic species with optima as high as 65 degrees. *Bacillus* species are easily isolated and readily grown in the bacteriology laboratory. The simplest technique that enriches for aerobic spore formers is to pasteurize a diluted soil sample at 80 degrees for 15 minutes, then plate onto nutrient agar and incubate at 37 degrees for 24 hours up to several days. The plates are examined after 24 hours for typical *Bacillus* colonies identified as catalase-positive, gram-positive, endospore-forming rods.

The selection of *Bacillus* spores used for this study was based on previous work [41, 42] and the characteristics and availability of the spores. The spores that simulate pathogens were used to assess the feasibility of 2D spectroscopy that could detect and identify spores posing an immediate threat. The spores forming *Bacillus* spores used for this experiment furnished as a generous gift from the US Army SBCCOM. A list of samples investigated in this research is given in Table 3-1. The uses of multiple samples of the same species yield information on the amount of variability in the fluorescence intensities of the samples. None of the details of spore production

methods were made available to us, except that none of the spores were treated with any special flow agent or anti-clumping compounds. All the spore samples were measured as they were provided and in a dry state. We have measured the two dimensional excitation/emission spectra of samples from spore forming bacteria, fungal spores, non bacterial samples on different substrates.

Table 3-1 List of simulants and interferents investigated

Sample	**Source**	**Characteristics**	**Purpose**
AFIP *Bacillus globigii* (*B. globigii* (i)) unwashed	ATCC 9372	Bacterial spore	*B. anthracis* simulant
AFIP *Bacillus globigii* (*B. globigii* (i)) washed	ATCC 9372	Bacterial spore	*B. anthracis* simulant
Merck *Bacillus globigii* (*B. globigii* (ii)) unwashed	From Agnes' aerosol lab Lot D-1-1732-6 – 1-18-56	Bacterial spore	*B. anthracis* simulant
Merck *Bacillus globigii* (*B. globigii* (ii)) washed	From Agnes' aerosol lab Lot D-1-1732-6 – 1-18-56	Bacterial spore	*B. anthracis* simulant
Dugway *Bacillus globigii* (*B. globigii* (iii)) unwashed	New (2001, Dugway-Denmark) B.GLOBIGII-Lot 06057- 18880 Batch No. 012	Bacterial spore	*B. anthracis* simulant
Dugway *Bacillus globigii* (*B. globigii* (iii)) washed	New (2001, Dugway-Denmark) B.GLOBIGII-Lot 06057- 18880 Batch No. 012	Bacterial spore	*B. anthracis* simulant
Bacillus cereus (*B. cereus*)	ATCC	Bacterial spore	*B. anthracis* simulant
Bacillus subtilis (*B. subtilis*)	ATCC	Bacterial spore	*B. anthracis* simulant
Bacillus popilliae (*B. popilliae*)	ATCC	Bacterial spore	Insectiside/interferent
Bacillus thuringiensis (*B. thuringiensis*)	ATCC	Bacterial spore	Insectiside/interferent
Ovalbumin	Bal Milled 1hr – from ECBC aerosol team.	Protein	Toxin simulant
Corn smut	Lot 8553-10 from Green Labs. Lenior, NC	Fungal spore	Interferent
Pig weed	Amaranthus retroflexus Lot 26X52-6b from Carolina Bioligical Supply	Pollen	Interferent
Dust	Home dusts		Interferent
Tryptophan	Fisher Biotech	Chemical	Amino acid
Anthracene	BDH Chemicals Ltd	Chemical	Standard
Ethanol	Scharlau ET 0015	Hydrophobic	Solvent

3.2 Spectral Measurements

The fluorescence emission spectra of samples investigated were measured in a SLM 8000C spectrofluorometer with xenon lamp (450 W) light source. Such lamps are generally useful because of their high intensity at all wavelengths ranging from 250 nm. A schematic of the experimental setup is shown in Figure 3.1. The instrument shown is equipped with monochromators to select both the excitation and the emission wavelengths. The excitation monochromator contains two aberration corrected concave gratings, which decreases stray light (i.e., light of wavelengths different from the chosen wavelength). Gratings are 1500 grooves per mm with a reciprocal linear dispersion of 2 nm per mm, an input focal length of 0.32 m. The emission monochromator is the same as excitation monochromator except that it includes one holographic grating which further decreases stray light. Both monochromators are motorised to allow automatic scanning of wavelength. The fluorescence was detected with photomultipliers that detect light in the current mode. The schematic diagram in the Figure 3.1 also shows the components of the optical module which surrounds the sample holder. The slits on both monochromators were set for a 4 nm bandwidth.

Alternately, fluorescence and phosphorescence spectra were also measured using a Hitachi fluorescence spectrophotometer, model F- 4500 with xenon lamp (150 W) light source. This instrument was configured with monochromators with eagle design with stigmatic concave diffraction grating 900 L/ mm, blazed at 300 nm for excitation and 400 nm for emission. The slits on both monochromators were set for a 5 nm bandwidth. We use a scan rate of 1200 nm per minute.

For calibration of monochromators we used light from a mercury lamp whose light is allowed to directly enter the cuvette holder and locate dominant mercury lines using the emission monochromator and ensure that the line is assigned the correct wavelength. After calibration of the emission monochromator, the excitation monochromator was calibrated against this standard, with some of the excitation light scattered into the emission monochromator. We further verified the calibration with a dilute solution of anthracene in ethanol placed in the cuvette holder to scatter

the exciting light. By measuring the fluorescence of anthracene using excitation wavelengths from 320 to 360 nm we determined the system response.

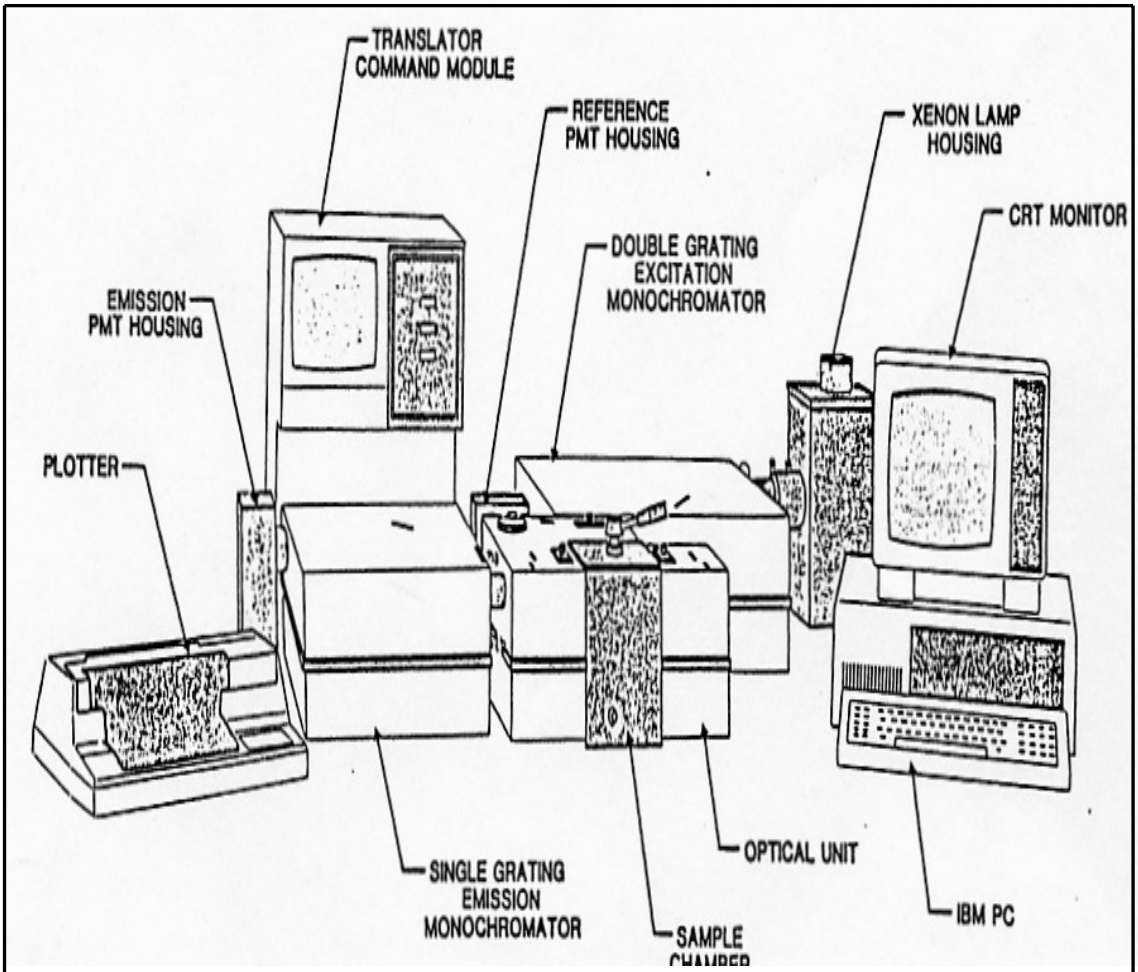


Figure 3.1 The SLM 8000C Spectrofluorometer (Retrieved from operational manual version 4.0 of SLM 8000C, Urbana, IL)

The excitation wavelength used to create each emission spectrum was varied in 10 nm increments from 200 to 600 nm. Light diffraction at the grating can occur as a first, second or higher order process and these diffraction orders frequently overlap. Hence, for this reason the emission spectrum was stepped every 1 nm from 10 nm longer than the excitation wavelength to 10 nm shorter than twice the excitation

wavelength (for Hitachi fluorometer, the emission spectrum was stepped every 0.2 nm) and that no added filters were used for wavelength selection of the monochromators. The wavelength longer than 700 nm was not measured since the photomultiplier was not sensitive at wavelengths longer than 700 nm. In Table 3.2 we show all the excitation wavelengths and the exact ranges of emission wavelengths used in the data collection.

Table 3-2 Excitation wavelengths used and the range of emission wavelengths used to collect the emission spectra.

Excitation Wavelength (nm)	**Emission Range Measured (nm)**	**Excitation Wavelength (nm)**	**Emission Range Measured (nm)**	**Excitation Wavelength (nm)**	**Emission Range Measured (nm)**
200	210 -390	340	350-670	480	490-700
210	220-410	350	360-690	490	500-700
220	230-430	360	370-700	500	510-700
230	240-450	370	380-700	510	520-700
240	250-470	380	390-700	520	530-700
250	260-490	390	400-700	530	540-700
260	270-510	400	410-700	540	550-700
270	280-530	410	420-700	550	560-700
280	290-550	420	430-700	560	570-700
290	300-570	430	440-700	570	580-700
300	310-590	440	450-700	580	590-700
310	320-610	450	460-700	590	600-700
320	330-630	460	470-700	600	610-700
330	340-650	470	480-700		

3.3 Spores Fluorescence

The starting point of our research was to measure various spores' fluorescence. *Bacillus* spores fluorescence measurements involve different sets of preparation and procedures according to the goal of the experiments. In the following several sections, we are giving the details of samples preparations, the QE measurements of spores, types of substrates used, collection and detection of spores in different backgrounds and fluorescence signal measurements in a fluorometer.

In the first section of spores' fluorescence study, we investigated a very important parameter for the detection of *Bacillus* spores, i.e. the measurements of QE of dry spores. In this section we will show various spore's fluorescence intensity measured as a function of the number of spores counted in each sample, optical data obtained in terms of absorbance and corrected integrated fluorescence, and finally the measured values for QE in comparison to a standard sample of anthracene in ethanol.

In the subsequent sections, we will show the detection of various *Bacillus* simulants, non bacterial spores and fungal spores on a plastic tape and quartz substrates. We will continue the investigation with detection of three *Bacillus* species from three different dusts backgrounds. As a follow up we will show the detection of *Bacillus* spores from pollen background. Again detection of spore fluorescence in a different perspective, i.e. effects of washing on fluorescence will also be covered in this section. Also we investigated the behaviour of *Bacillus* spores in dry forms, spores in aqueous and ethanol suspensions.

3.3.1 Fluorescence Quantum Efficiency of Dry *Bacillus* Spores

We measured the fluorescence signal as a function of the number of spores. The fluorescence QE of the samples under study was measured relative to a standard solution of a fluorophore such as anthracene. The fluorescence yield was from a measurement of absorption and fluorescence radiance. The absorption is a strong function of wavelength and a series of measurements of absorption and fluorescence are performed for different concentration/numbers, then the measured fluorescence

intensity / optical density, implies the relationship. The signal to noise ratio will be small if we measure the fluorescence of the small number of spores. We can use these data to determine the limits of detection and the limits of identification for the algorithms.

The spores forming bacteria *Bacillus globigii*: *B. globigii* (i), *B. globigii* (ii) (Merck), *B. globigii* (iii) (Dugway) and *B. cereus* were used for this experiment. The uses of multiple samples yield information on the amount of variability in the QE for *B. globigii*. We have performed a series of fluorescence measurements of *B. globigii* samples in unwashed and washed redried form and *B. cereus* in unwashed conditions. The unwashed dried spores were fixed to the sticky side of a quartz substrate (Esco, New Jersey, USA). The quartz was cut about 0.9 cm wide and 3 cm high. The quartz with dry spores was masked with aluminium foil having 2 mm hole to view the spores.

Images of the spores were taken *in situ* using an Olympus system microscope model CX41 (Melbourne, Australia) and a digital camera and were transferred to a computer to count the number of pixels for one spore and the number of dark pixels, thereby obtaining an estimation of the number of spores used for each sample. After removing the aluminium foil, the quartz with spores was placed diagonally in the quartz cuvet and placed in the spectrofluorometer (SLM 8000C), so that the front face of the quartz slide (with the spores) was positioned at a 30-degree angle to the excitation and emission optical axes as shown in Figure 3.2. This particular quartz and glue have a very small fluorescent signal that does not overlap very strongly with the spore fluorescence.

For washed redried measurements, clean samples of *B. globigii* spores were prepared by centrifuging at 10000 rpm for five minutes, washing twice and redried on a quartz surface. Experiments were repeated as before for a series of measurements.

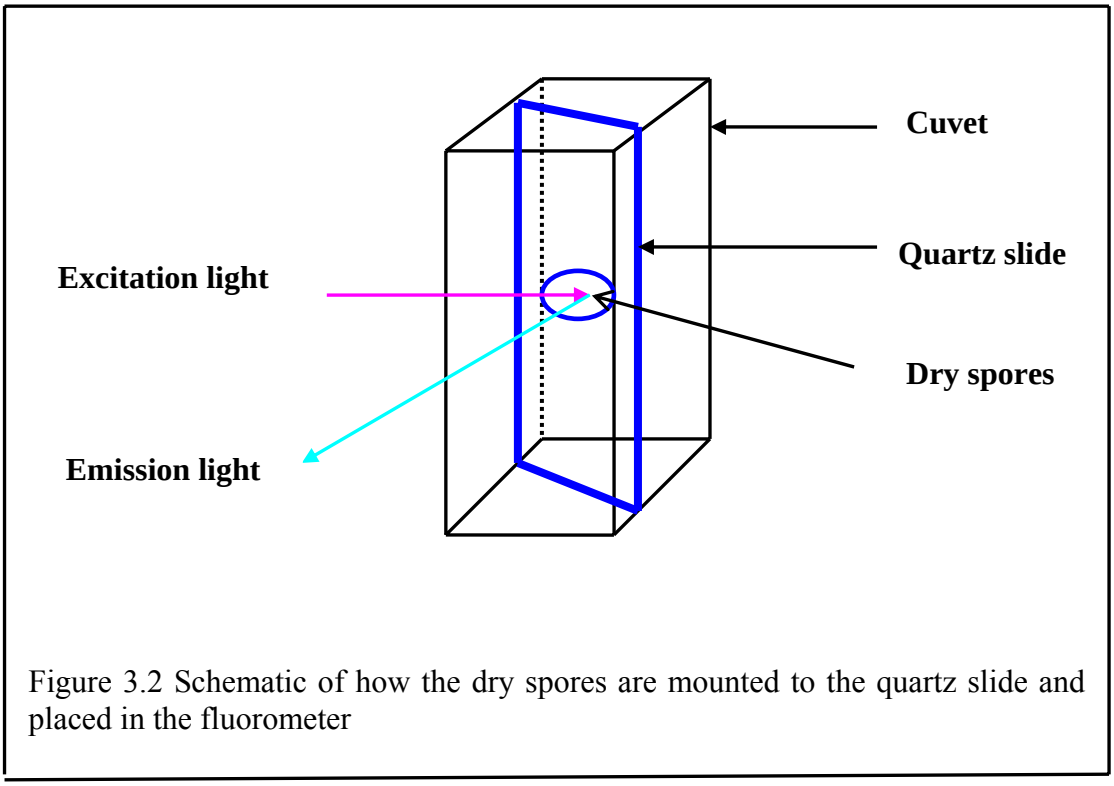


Figure 3.2 Schematic of how the dry spores are mounted to the quartz slide and placed in the fluorometer

Absorbance of dried samples were measured with varying concentrations/numbers in a scanning range 200-600 nm using a GBC UV/VIS Cintra 40 spectrometer (Victoria, Australia). The slit was set with a 2 nm width.

Anthracene has a molar extinction coefficient of 9,700 $M^{-1}cm^{-1}$ at 356.25 nm [138] and fluorescence yields of 0.27 in ethanol [139, 140]. The anthracene concentrations used in this experiment depend on the wavelength of exciting light. It is important that the absorbance of the solution is not more than about 0.2 O.D so that the fluorescence intensity is uniform throughout the solution. We chose a concentration range of 1- $2x10^{-5}$ M solution of anthracene. Solutions of anthracene were prepared by weighing dry anthracene powder in a measured volume of spectrograde ethanol, namely 2.4 mg/50 ml, 2.4 mg/100 ml etc. The absorption spectra of the standard sample anthracene in ethanol were measured with varying concentrations in a scanning range 200-600 nm using the spectrometer. Here also the slit was set with a 2 nm width. We measured the fluorescence of ethanol and fluorescence intensities of the prepared anthracene solutions with varying concentrations. The details were given in the previous section.

Various plots of absorbance versus the number of spores used for each sample in this section were plotted for calculating the absorption coefficients of the samples. For measuring QE, we limited our measurements with three values of QE at excitation wavelengths 280, 360 and 400 nm for *B. globigii* samples and for *B. cereus* only two values were calculated at 280 and 300 nm excitation wavelengths, since only at these excitation wavelengths, the samples show measurable fluorescence intensities. As explained before, each samples' corrected fluorescence intensities at the selected excitation wavelength, were calculated. The fluorescence QE of the samples relative to anthracene were determined using the following formula,

$$QE_{sample} / QE_{anthracene} = \int I_{sample} \; A_{anthracene} \, d\lambda / \int I_{anthracene} \; A_{sample} \, d\lambda \qquad (3.1)$$

where $\int I$ is the fluorescence intensity integrated over the emission spectrum and A is the absorbance at the selected excitation wavelength. Plots of integrated intensities versus absorbance were drawn for each sample and compared with the slopes of the standard sample anthracene, whose corrected fluorescence intensity versus absorbance at various excitation wavelengths were calculated.

3.3.2 Spores Detection from Tape Substrates/Dusts

One of the most likely biological components of an aerosol would be a bacterial endospore that can be seen in different backgrounds. For the collection and detection of spores, one must determine the optimal substrate. The chosen substrates must have a very small fluorescent signal that does not overlap very strongly with the spore fluorescence. In Table 3.3, we show the different types of substrates used in this investigation for the collection of spores and fluorescence and phosphorescence measurements.

Table 3-3 The different types of substrate used for the collection and detection of spores.

Substrate materials	Source
3M Super Strength Mailing Tape	3M , St. Paul, Minneapolis, MN
Quartz slides	Esco, New Jersey,USA
Filter paper	GF/A, Glass Microfibre Filters, Whatman, Kent, UK

To begin with fluorescence measurements, we started with a plastic tape for collection and detection of simulants investigated. For detecting autofluorescence signals from simulants with and without background, we found that 3M Super Strength Mailing Tape was useful. Initially we measured fluorescence simulants from this tape without any background, such as dust or pollen. Separate measurements were done three times, each with spores of three preparations of *B. globigii* (i), *B. cereus*, *B. popilliae*, ovalbumin (egg white), corn smut (fungal spores) and pig weed (pollen grain). The dried spores were fixed to the sticky side of Scotch (3M, Minneapolis, MN) extra strength mailing tape. The tape was cut about 1 cm wide and 3 cm high. The tape with spores was placed diagonally in the a quartz cuvet and placed in the spectrofluorometer (F- 4500) so that the front face of the tape (with the spores) was positioned at a 45 degree angle to the excitation and emission optical axes as shown in Figure 3.3 [141]. This particular tape was selected after making measurements of the fluorescence from many different glues and tapes. This particular tape and glue have a very small fluorescent signal that does not overlap very strongly with the spore fluorescence (see Figure 4.22A).

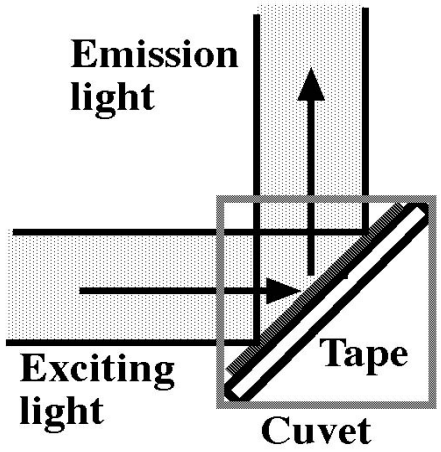


Figure 3.3 Schematic of how the dry spores are mounted to the tape and placed in the fluorometer [141].

The rapid detection and identification of bacteria and bacterial spores have many applications. One of the more recent applications is for the detection and identification of bacterial spores that might be used in bioterrorism. One would like to have a method to quickly collect and identify the potential danger from an unknown powder spilling from an opened letter. If the spores were to be distributed in an act of bioterrorism, the collection of the spores for analysis would most likely also involve the collection of dust and other particles in addition to the spores. We have, therefore, added domestic dust to three samples: *B. globigii* (i), *B. cereus* and *B. popilliae*. A simple, but effective, method of collecting the spilled powder or spores for analysis is to use a piece of tape. Spilled powder could be picked up with the sticky side of tape. The tape could then be put into a device where the powder is analysed.

Several pieces of tape are cut in 1 x 3 cm pieces to fit into a 1 cm path length quartz cuvet. Nine pieces of tape are used to collect domestic dust from three different locations in a house (living room, bathroom and bedroom). We call these samples, dust A, B and C respectively. The tape with dust from each of the three locations is lightly coated with one each of three different *Bacillus* spores: *B. globigii*

(i), *B. cereus* and *B. popilliae*. Thus, there are nine samples with each combination of dust and bacterial spore. The fluorescence of each dust and spore combination is then measured using a fluorometer ((F- 4500) as explained in previous section. These nine samples are used as the training set for PCA analysis that will be explained in data analysis section.

In addition, we measure the fluorescence signals from three other samples without added dusts. We measure the fluorescence signals from a sample of *B. globigii* spores prepared from an outside source (*B. globigii* (ii)). The outside source is Merck Chemicals (St. Louis, MO USA Lot D-1-1732-6). We also measure corn smut, a fungal spore (Lenoir, NC, Lot 8553-10) and ovalbumin (egg white and a growth medium ingredient) (Sigma, St. Louis, MO USA).

3.3.3 Spores Detection from Quartz Substrates/Pollen

Spores are likely to be seen mixed with particles from various environments. The main hurdles in spore's fluorescence detection are the presence of background fluorescence due to various aerosol including pollen grains from plants. Fluorescence measurements of various *Bacillus* spores were found to be very easy if we use slides of quartz as substrate. Spores of *B. globigii* samples, *B. cereus*, *B. popilliae*, and *B. subtilis* were again measured on a quartz slide to check the effect of different substrates on spore's fluorescence. Also, to detect *Bacillus* spores from a pollen background, pig weed (Amaranthus retroflexus Lot 26X52-6b from Carolina Biological Supply) and a mixed sample of pig weed and spores of *B. globigii* (iii) were also used.

Samples were prepared similar to the previous sections. Here, we used quartz slides for collection and detection of *Bacillus* simulants. The quartz was cut into small slides about 0.9 cm wide and 3 cm long. The quartz slides were coated with a very thin layer of low fluorescence adhesive. The spores were fixed to a quartz plate and the coated slide was placed diagonally in a quartz cuvet and placed in the spectrofluorometer (SLM 8000C) so that the front face of the quartz slide (with the spores) was positioned at approximately a 30 degree angle to the excitation optical axes and approximately a 60 degree angle to the emission optical axes (see Figure 3.2). The chosen quartz had very little fluorescence signal that did not overlap very

strongly with the *Bacillus* spore fluorescence (see Figure 4.22 B&D). Separate fluorescence measurements were done for the samples (simulants) under study with and without added background pollen, on a quartz substrate in fluorometer as explained in the previous sections.

3.3.4 Effects of Washing on the Identification of *Bacillus* Spores

We were interested in determining the effects of washing the spores on our ability to correctly identify the spore. Spores can be produced with a varying amount of washing. The spores might have adhered bits of the growth medium or fragments of the vegetative cells clinging to the outside of the spore. If the fluorescence signal of a spore is dominated by these remnants of the spore production process, the fluorescence would better indicate how the spore was grown more than the identity of the spore. In such a case, it might be necessary to wash all spores before trying to identify them. On the other hand, the bits of vegetative cells or growth medium that is dried with the spores might produce only a small background noise in the fluorescence information. In this case, the PCA is particularly appropriate to ignore that information when using cluster analysis to identify the spore.

Four different species of *Bacillus* spores: *B. cereus, B. popilliae, B. thuringiensis and B. globigii* were used in this experiment. The six spore samples with three of the samples *B. globigii* used before and after washing provide nine different samples. For each of the nine samples, we prepared three quartz slides to measure the fluorescence. The quartz was cut into small slides about 0.9 cm wide and 3 cm long. The quartz slides were coated with a very thin layer of low fluorescence adhesive. The spores are fixed to a quartz plate and the coated slide was placed diagonally in a quartz cuvet and placed in the spectrofluorometer (SLM 8000C) so that the front face of the quartz slide (with the spores) was positioned at approximately a 30 degree angle to the excitation optical axes and approximately a 60 degree angle to the emission optical axes (see Figure 3.2).

The six different preparations of the spores were: *B. popilliae, B. cereus, B. thuringiensis, B. globigii* (i), *B. globigii* (ii), and *B. globigii* (iii). The three preparations of *B. globigii* (i, ii and iii) were produced in different laboratories and at different times. None of the details of spore production methods were made available

to us, except that none of the spores were treated with any special flow agent or anti-clumping compounds. All the spore samples were measured as they were provided and in a dry state.

In addition, the three preparations of *B. globigii* were measured before and after washing. The samples were purified by repeated centrifugation and water washing. Clean samples of *B. globigii* spores were prepared by centrifuging at 10000 rpm for five minutes, washed with a large quantity of deionized water twice and redried on a quartz surface in air for 24 hours. The supernatant was removed and the pellet was kept. The use of multiple samples yields information on the amount of variability in the fluorescence of the samples. The unwashed samples were found containing significant amount of growth media or other debris. The supernatant was a red-brown colour and the spores were a lighter shade after being washed and we could still distinguish the spores before or after washing the spores.

Fluorescence of all samples was measured as explained in the previous sections. It is important to note that we did see changes in the fluorescence upon washing the spores. In fact, some of the spores appeared different to the naked eye after washing, but no further testing was done to check the nature of spores after being washed and redried. Before washing, most of the spores were isolated from one another or in small clumps of a few spores. This was evident from the microscopic images (not shown).

3.3.5 Autofluorescence of *Bacillus* Spores and Role of Tryptophan: A Comparative Study of Dry Spores and Spores in Aqueous and Ethanol Suspensions.

In an attempt to understand more about the role of tryptophan in autofluorescence of *Bacillus* spores, we carried out a comparative study of fluorescence of tryptophan and *Bacillus* spores, in their dried form on filter paper and fluorescence measurements in aqueous and ethanol solutions.

Sample preparation was done separately. For measuring fluorescence of a dry sample, the filter paper (GF/A, Glass Microfibre Filters, Whatman, Kent, UK) was found to be very useful [142]. The filter paper was cut about 1cm wide and 3 cm high. The dried tryptophan powder was smeared on the filter and then made wet with

water to minimise the concentration and afterwards, re-dried. The filter paper with tryptophan powder was placed at 30 degrees to the incident and fluorescence light in a 1 cm path length quartz cuvet (see Figure 3.2) and placed in the spectrofluorometer (SLM 8000C). The chosen filter had a minimal fluorescence signal that did not overlap very strongly with the *Bacillus* spore fluorescence (see Figure 4.40A).

Solutions of L-tryptophan were prepared in Milli Q water and in spectrograde ethanol with concentration of 10mg/L each. Fluorescence wavelength and QE depend on pH and temperature [143]. The fluorescence measurements were performed in room temperature in a more conventional manner at 90 degrees to the exciting light. Because tryptophan is degraded by exposure of light, extra care was taken to shield the tryptophan solutions from direct light.

Although bacteria are often found in an aqueous environment and are frequently measured in aqueous suspensions, bacterial spores can be found in a completely dry environment. Therefore, we wanted to probe the fluorescence of the dried *Bacillus* spores and compare the fluorescence spectra with spores in suspension.

As a special case, separate fluorescence measurements of *B. globigii* (i) were done in their dried form, with spores in aqueous (Milli Q water) and in ethanol suspension. Fluorescence spectra of dry spores of *B. globigii* (i) were measured on a filter paper similar to our previous study [142]. We tried to measure fluorescence of dry spores on filter paper, instead of using plastic or tape, because we were interested in comparing the dry spores' fluorescence with the fluorescence of the dry powder of tryptophan, which was measured on filter paper. For suspension measurements, the spores were suspended for at least two hours before the measurements were made, to allow them to fully hydrate. Measurement details were the same as above.

Chapter 4

Spectroscopy and Data Analysis Results and Discussions

In this chapter we are presenting various spectroscopy and data analysis results and discussions. There are many issues related to the spectroscopy of *Bacillus* spores, the identification of their characteristic fingerprints and data analysis. The spectrometers and the measurement of the spectra do not represent a major technical challenge. However, the sampling process, sample preparation, and their natural variability in terms of shape and chemical composition can be expected to introduce considerable differences in their spectral response. The reproducibility study, the results of spectral measurements (both absorbance and fluorescence) and data analysis using PCA reported herein have been limited to 1) the QE of *Bacillus* spores, 2) spores detection from different substrates, 2) detection of spores from different dusts backgrounds, 3) detection of spores from pollen, 4) effect of washing on fluorescence of samples investigated and 5) fluorescence of tryptophan, fluorescence of dry *B. globigii* spores, spores suspensions in sterilized de-ionized water and in ethanol.

Absorption and fluorescence spectrum of various *Bacillus* spores under study were measured. These measurements and data analysis were done to answer some of the questions posed at the start of the research:

Do fluorescence QE measurements of dry spores give a detection limit for detection and identification unit?

Do various *Bacillus* simulants show different fluorescence fingerprints?

Do agent simulant show different spectra when measured in dry spores or spores suspended in aqueous or ethanol?

Does PCA provide an excellent tool for pattern recognition?

Dry *Bacillus* spores fluorescence spectra were obtained on different substrates with different spore's numbers/concentration to produce spectra of good signal to noise. The EEM technique was applied for visual representation of

fluorescence spectra. Fluorescence QE was calculated as a function of number of spores in comparison with a standard solution of anthracene in ethanol. Also spectra of spores from different backgrounds were obtained. For comparative study with tryptophan, spectra of dry spores of *B. globigii* (i), spores in aqueous and ethanol suspensions were obtained. PCA and cluster analysis were done to discriminate the spore fluorescence from different environments. Results are given as an analysis of 2D fluorescence fingerprints of various *Bacillus* spores simulants. Results are also given from PCA and cluster analysis performed on both washed and unwashed *Bacillus* spores.

4.1 Spectral Visualization by EEM

The measured fluorescence spectra were transferred to a Windows XP computer where they are analysed with *Mathematica* (Wolfram Research, Urbana, IL). For viewing the fluorescence spectrum by EEM techniques; first, we divide the wavelengths of emission spectrum by the wavelength of excitation spectrum. The range of each spectrum is then approximately 1.0 to 2.0. Each spectrum is smoothed with a three point Gaussian analysis [43]. The three point Gaussian analysis fits the measured spectrum with the following sum of Gaussians.

$$P(x) = \Sigma_i A_i \exp[-(x - x_{oi})/2\sigma^2] \quad (4.1)$$

where x_{oi} are equally spaced every 4 nm across the data range, and the width of each Gaussian is equivalent 4 nm. The amplitudes A_i are determined from least-square fit from *Mathematica*. The parameterisation of the data is not useful in understanding or deconvolving the spectra. It simply gives an analytical equation to use in the subsequent analysis (permitting interpolation) and smooths the measured data. Any well-behaved function that approaches zero in both of the "tails" (limit as x approaches $-\infty$ and $+\infty$) can be used in this analysis. We arbitrarily selected Gaussians.

In Figures 4.1 and 4.2 we illustrated a single measured fluorescence emission spectrum of spores of *B. globigii* (i), on a quartz substrate at excitation wavelengths 280 and 420 nm with emission ranges 310- 550 nm and 450-700 nm respectively. The measured data are shown as points and the solid line shows the parameterized

and smoothed curve that we use to represent the data in the subsequent analysis. The parameterization is well behaved and approaches to zero. The Gaussian functions also force the fit to zero intensity at wavelength beyond the measured region.

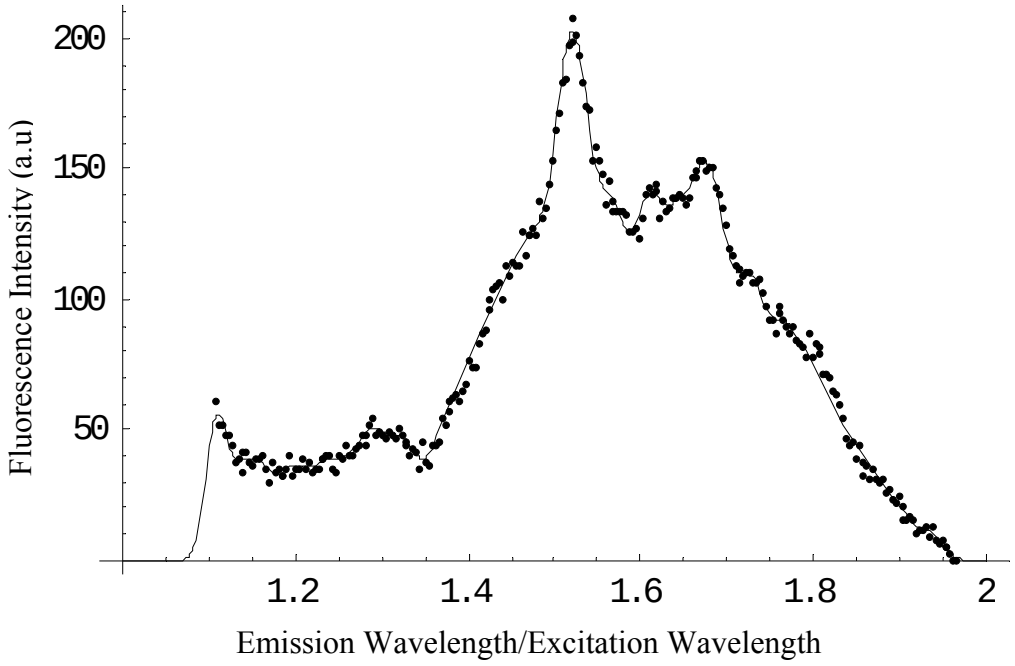


Figure 4.1 Illustration of the fluorescence emission spectrum from dry *B. globigii* (i) spores on quartz substrate. The excitation is at 280 nm. The points show the measured data during a typical experiment. The solid line is from the Gaussian smoothing and parameterisation of the data.

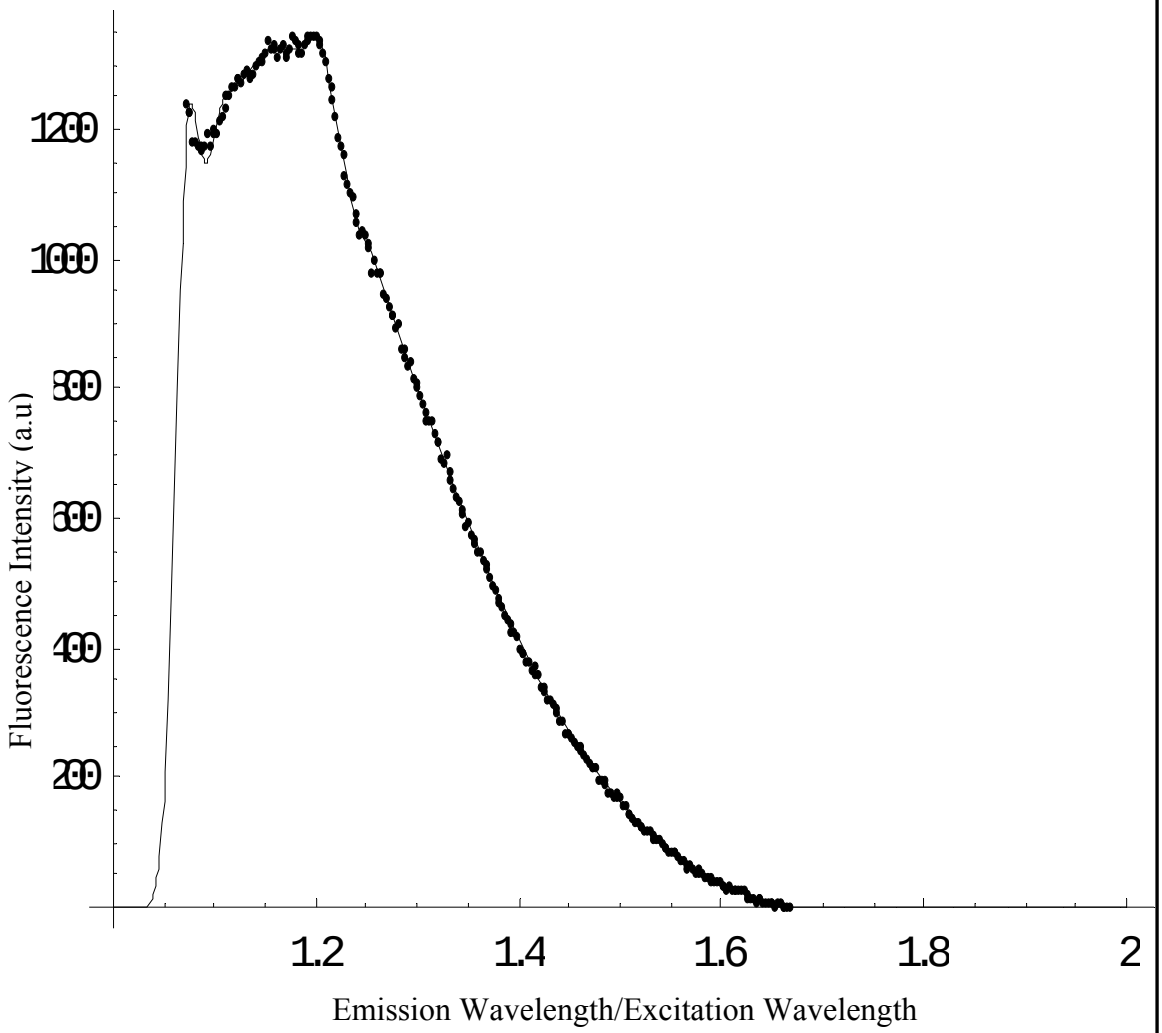


Figure 4.2 Illustration of the fluorescence emission spectrum from dry *B. globigii* (i) spores on quartz substrate. The excitation is at 420 nm. The points show the measured data during a typical experiment. The solid line is from the Gaussian smoothing and parameterisation of the data.

Mathematica then interpolates between the matrixes of emission spectra in calculating the contour plot. To make the contour plot, the fluorescence intensities are determined at 41 equally spaced points on the ratio axis for each excitation wavelength. The 41 equally spaced points on the ratio axis represent steps of 10 nm to 35 nm on a wavelength axis, depending upon the excitation wavelength. Although this method is non conventional, the variable step side weights each spectrum equally. Each spectrum contributes 41 points to be used in the least-squares fit. It also represents closer to a constant step side on an energy axis instead of equal steps on the wavelength axis. This result in a 41 x 41 matrix of intensities as explained earlier [51]. The intensities are shown as a contour plot. In the contour plot there is an intensity (z value) for each x and y point on the surface. In the fit procedure, the sum of the squared difference between the calculated z and the measured z value is minimised.

The contour plot of the fluorescence intensity as function of excitation wavelength and emission wavelength is usually termed a fluorescence fingerprint [51] (see Figure 4.3(A)). Fluorescence emission can be measured only at wavelength longer than the excitation wavelength. Typically, the emission is further restricted to be at wavelength shorter than twice the excitation wavelength. Generally, the second harmonic transmission through the monochromator dominates the fluorescence signal at wavelength longer than twice the excitation wavelength. The "stair-step" edges on the profile are caused by the discrete fluorescence spectra measured. We therefore plot the contour intensity curves as a function of excitation wavelength and the ratio of the emission wavelength divided by the excitation wavelength from 1.0 to 2.0. This presentation of the data avoids the usual "distortion" that one observes when fluorescence intensity contours are plotted on an excitation and emission wavelength plane.

In Figure 4.3(A), the conventional presentation of fluorescence fingerprints of dry spores of *B. globigii* (i) on a quartz slide can be compared with the modified (Emission wavelength / Excitation wavelength in the x-axis) method of presentation for the same data in Figure 4.3(B). In both cases the fluorescence maximum is normalised to one and the contour represent equal steps in the fluorescence intensity.

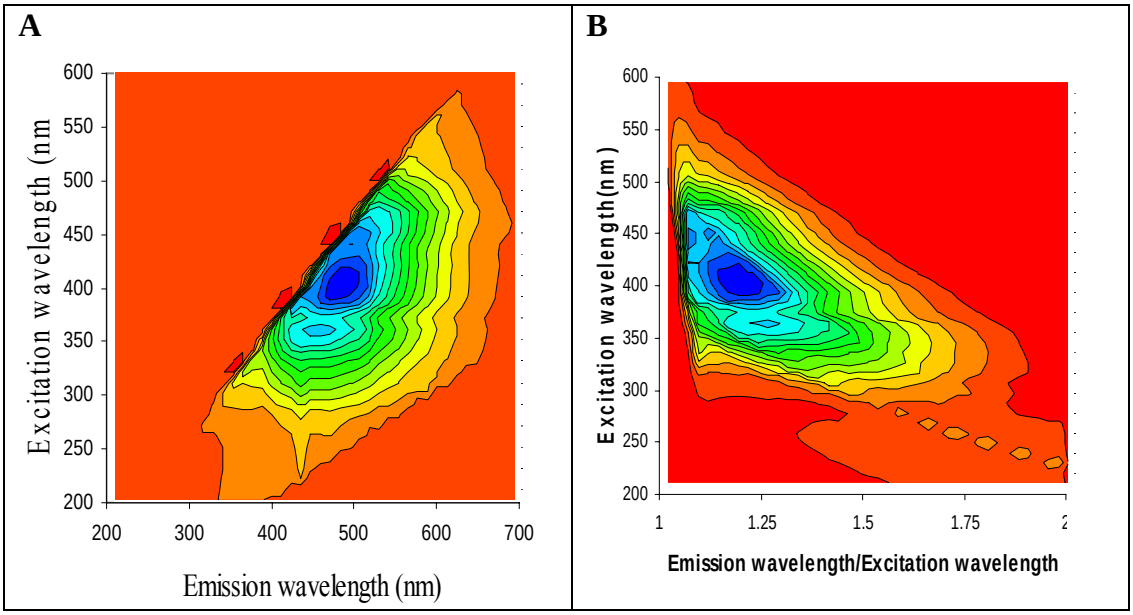


Figure 4.3 Illustration of the fluorescence fingerprints from dry spores of *B. globigii* (i) measured on quartz substrate. The generation of the fingerprint is explained in the text. In 4.3(A) the conventional fluorescence fingerprint and (B) the modified fluorescence fingerprint are displayed. In both parts of the figure, the contours represent equal changes in the fluorescence intensity, and the peak intensity is normalised to 1.

4.2 Fluorescence Quantum Efficiency Measurements of *Bacillus* Spores

In this section we present the results of fluorescence QE measurements of four samples of *Bacillus* dry spores measured in comparison with QE of the standard solution of anthracene in ethanol. The results are divided in two sub sections. In section 4.2.1, the various optical measurement of anthracene in ethanol and in section 4.2.2, the fluorescence and QE measurement of both unwashed and washed spores investigated are shown.

4.2.1 Absorbance and Fluorescence of Anthracene in Ethanol

In Figures 4.4 and 4.5 we show a sample absorption and fluorescence spectrum of anthracene in ethanol. The fluorescence emission maximum at 355.52 nm excitation wavelength is at 402 nm. As expected for a simple molecule like anthracene, mirror symmetry between absorption and fluorescence spectra can be observed.

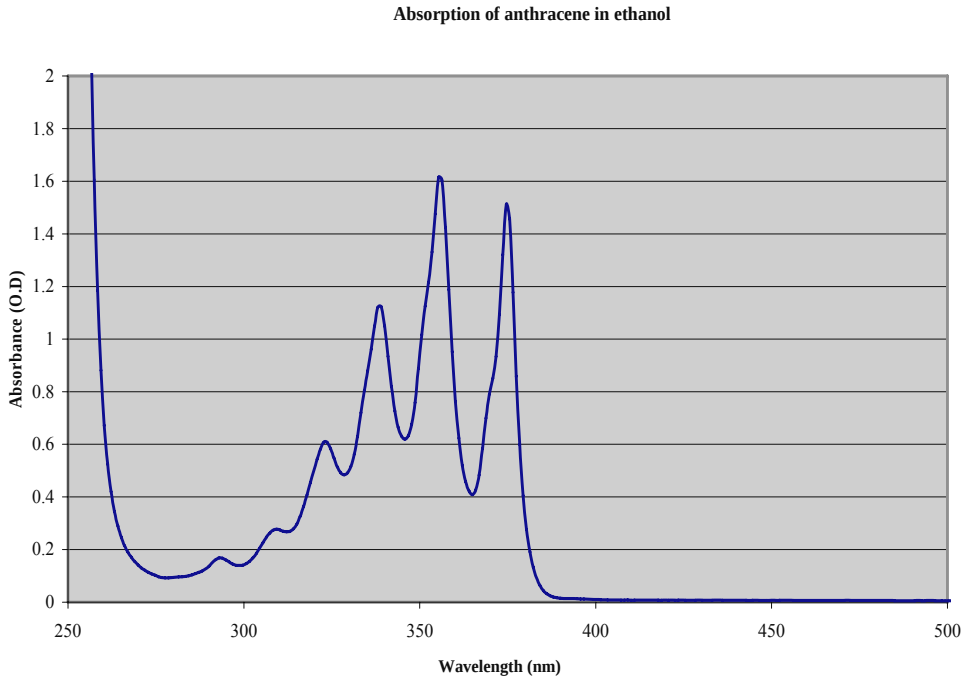


Figure 4.4 Absorption spectrum of anthracene in ethanol at a 2.21E-04 mol/litre concentration. The absorbance at excitation wavelength 355.52 nm is 1.6153 O.D.

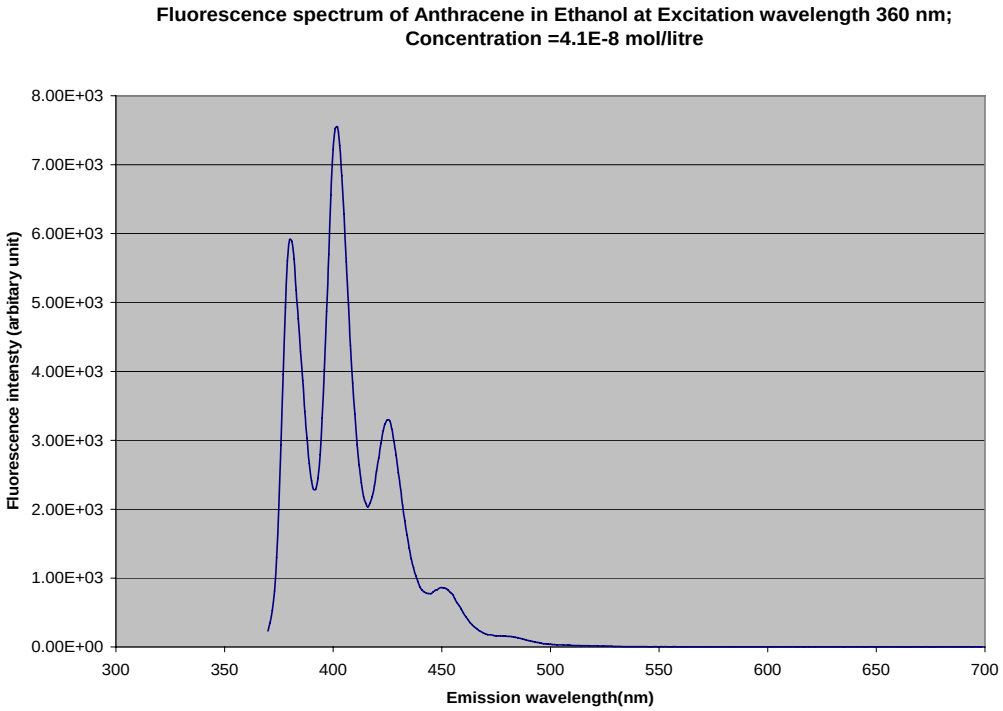


Figure 4.5 Fluorescence spectrum of anthracene in ethanol. Sample fluorescence spectrum of anthracene in ethanol is shown with excitation at 360 nm with emission maximum at 402 nm.

In Figure 4.6(A-B) we show the measured fluorescence 2D plot from a sample of ethanol and anthracene in ethanol. The fluorescence spectrum of anthracene was measured. Excitation was done from 200-600 nm with a 4 nm bandwidth. The integrated fluorescence was measured as a function of the concentration (absorbance).

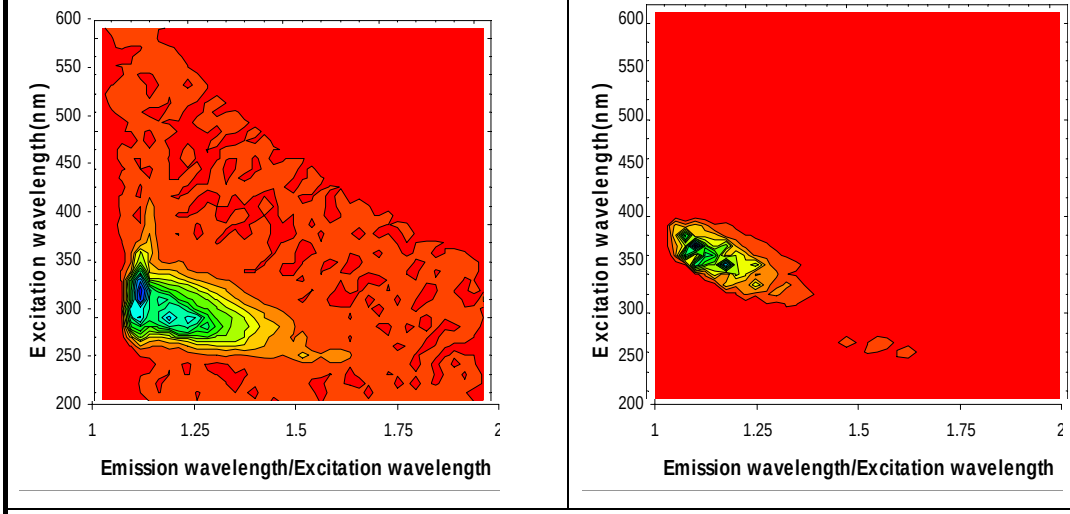


Figure 4.6 Sample fluorescence fingerprints of (A) ethanol and (B) anthracene in ethanol. The contour plot is the relative intensity of fluorescence as a function of the excitation wavelength and emission wavelength divided by the excitation wavelength. The excitation wavelengths were from 200-600 nm

Below in Tables 4-1 and 4-2 we show the optical data of anthracene in ethanol measured to plot fluorescence versus absorbance (O.D) plot.

Table 4-1 Absorbance of anthracene in ethanol in different excitation wavelength

Trial No	**Absorbance (O.D) at excitation wavelength**				
	320 nm	**330 nm**	**340 nm**	**350 nm**	**360 nm**
1	0.0566	0.0571	0.1186	0.0883	0.0985
2	0.0346	0.0350	0.0786	0.0569	0.0650
3	0.0203	0.0204	0.0532	0.0372	0.0426
4	0.0190	0.0192	0.0465	0.0332	0.0388
5	0.0170	0.0173	0.0400	0.0284	0.0331
7	0.0107	0.0108	0.0182	0.0144	0.0153
9	0.0178	0.0176	0.0321	0.0248	0.0275
10	0.0566	0.0571	0.1186	0.0883	0.0985

Table 4-2 Fluorescence data of anthracene in the selected excitation wavelengths

Trial No	**Corrected integrated fluorescence intensity (counts) at excitation wavelength**				
	320 nm	**330 nm**	**340 nm**	**350 nm**	**360 nm**
1	8.77E+04	1.06E+05	2.46E+05	2.12E+05	2.86E+05
2	4.93E+04	2.16E+04	1.48E+05	1.31E+05	1.79E+05
3	4.82E+04	5.80E+04	1.34E+05	1.16E+05	1.56E+05
4	4.15E+04	5.05E+04	1.18E+05	1.02E+05	1.37E+05
5	2.68E+04	3.36E+04	8.14E+04	7.25E+04	9.86E+04
7	1.21E+04	1.45E+04	3.45E+04	2.98E+04	4.00E+04
9	2.23E+04	2.70E+04	6.45E+04	5.54E+04	7.50E+04
10	8.77E+04	1.06E+05	2.46E+05	2.12E+05	2.86E+05

In Figure 4.7 we show the slope measurement of anthracene in ethanol. Corrected integrated fluorescence intensity (counts) versus absorption (O.D) of anthracene in ethanol at excitation wavelengths 320, 340, 350 and 360 nm were plotted to correct the instrument response for fluorescence data. Table 4-3 shows the summary of the measured and calculated slope for various excitation wavelengths. We chose the linear portion of the plot in Figure 4.8 from excitation wavelength 320 to 360 nm to calculate the QE of *Bacillus* spores as this region represents the actual fluorometer response to various excitation wavelengths.

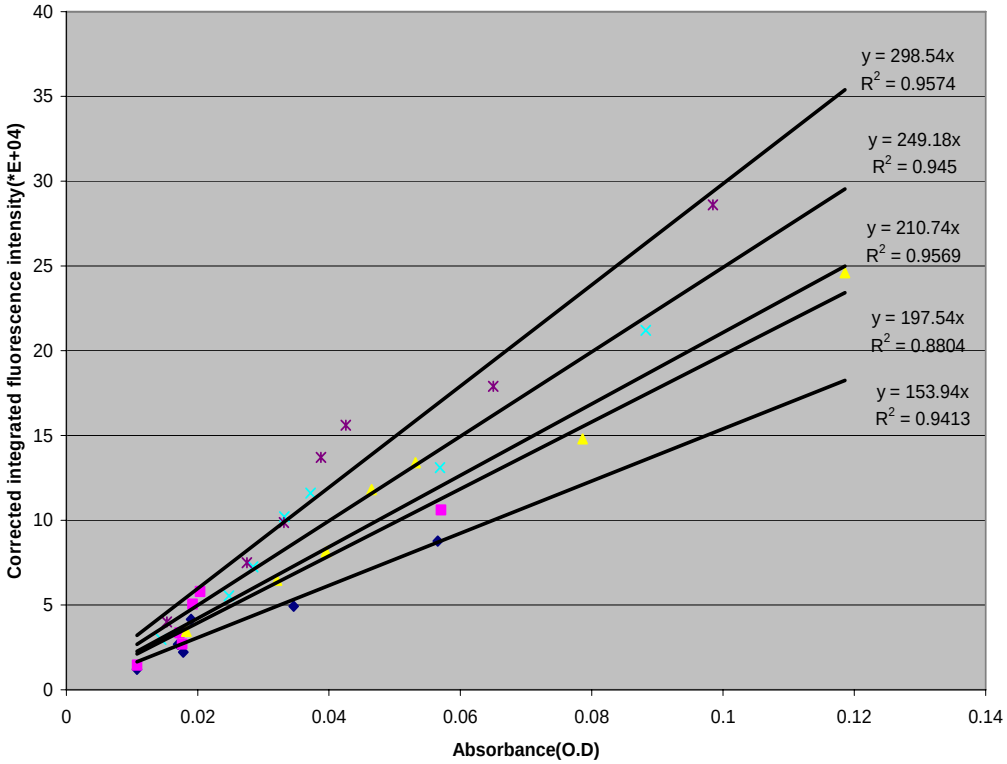


Figure 4.7 Fluorescence of anthracene in ethanol as a function of absorbance. The integrated fluorescence intensity (counts) of anthracene in ethanol is shown versus absorption (O.D) at 320, 330, 340, 350 and 360 nm excitation wavelengths (with corresponding colours blue, red, yellow, light blue and brown respectively). The line is linear least squares fit. Each slope is proportional to the known QE of anthracene in ethanol at the excitation wavelength selected.

Table 4-3 The summary of the measured and calculated slope for various excitation wavelengths

Excitation wavelength (nm)	Slope (Counts/O.D)
280(Calculated)	17.4E+04
300(Calculated)	85.54E+04
320	153.94E+04
330	197.54E+04
340	210.31E+04
350	249.18E+04
360	298.54E+04
365.5 (Calculated)	308.86E+04
400 (Calculated)	426.45E+04

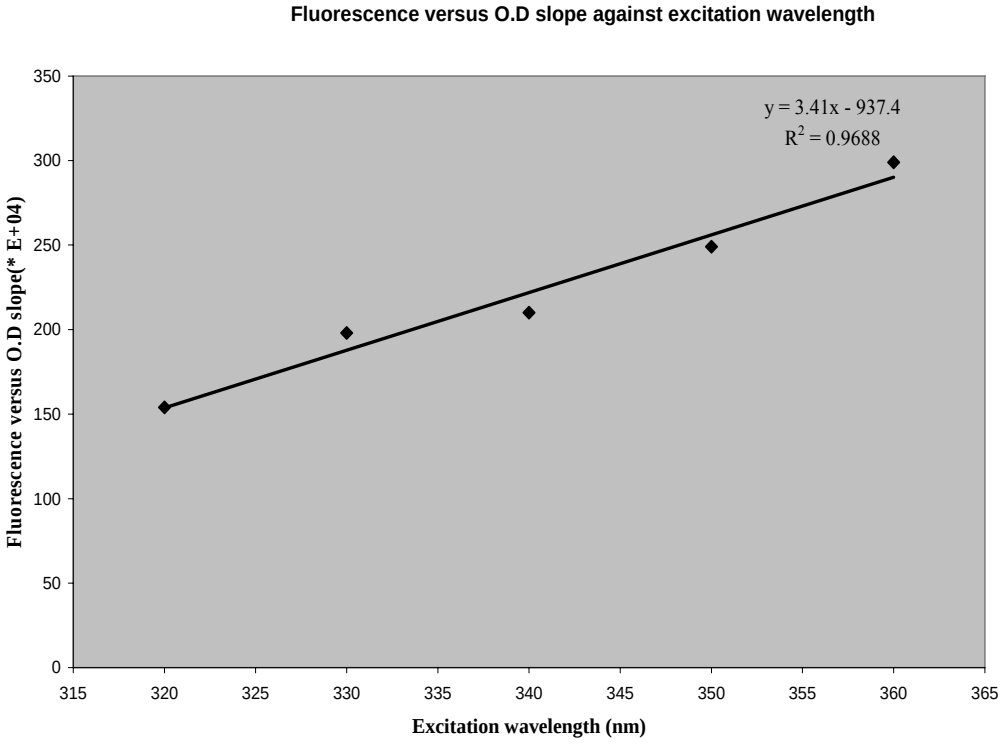


Figure 4.8 Response plot of the measured fluorescence intensity (counts) versus absorbance (O.D) against excitation wavelength range from 320 to 360 nm

4.2.2 Quantum Efficiency of *Bacillus* Spores

In this section we are presenting results of QE measurements made with the samples investigated**.** It is of interest to measure the QE for various microorganisms because this value determines limits of sensitivity. In the case of spores we can measure a comparative value of this quantity with a standard and did so for unwashed and washed spores of *B. globigii* (i), *B. globigii* (ii), *B. globigii* (iii) and *B. cereus* (unwashed). This can be done since the absorption appears to rise sharply in the UV region as indicated by the samples absorption plots in Figures 4.9 and 4.10. Because of sharp rise in the graphs to a peak at 280, 360 and 400 nm, we judged that most of the O.D was due to absorption by the spores at that wavelength. The difference in peak height was due to the difference in number/concentration of spores used.

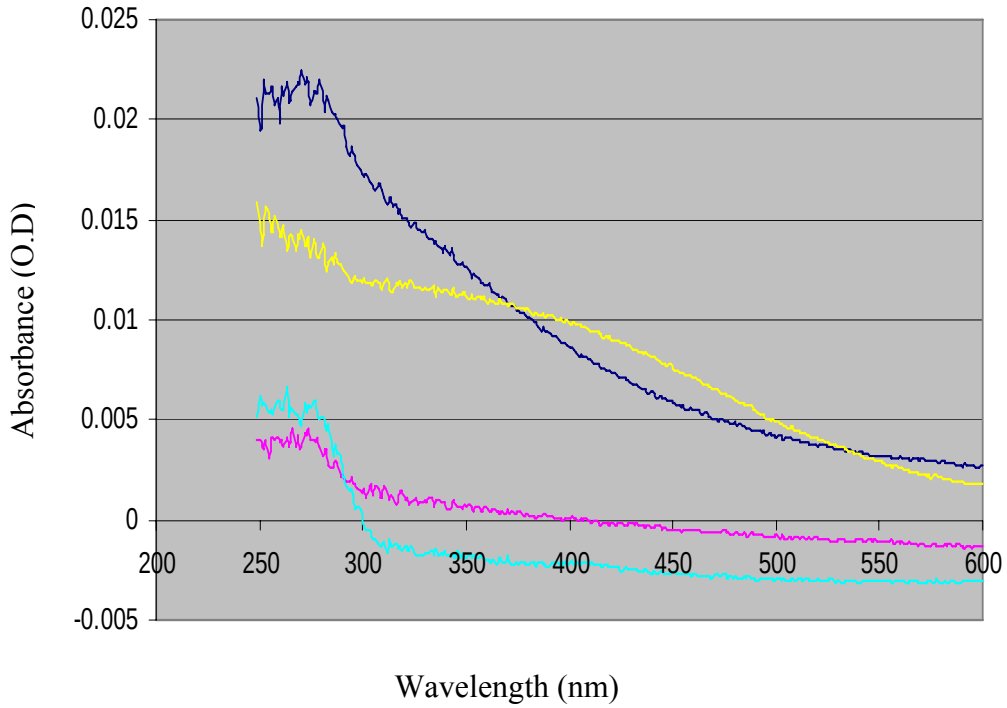


Figure 4.9 Representative plots of absorbance spectrum of unwashed Bacillus spores studied for QE measurements. Samples: *B. globigii* (i) (Blue), *B. globigii* (ii**)** (red), *B. globigii* (iii) (yellow) and *B. cereus* (light blue). Scanning wavelength was from 200-600 nm.

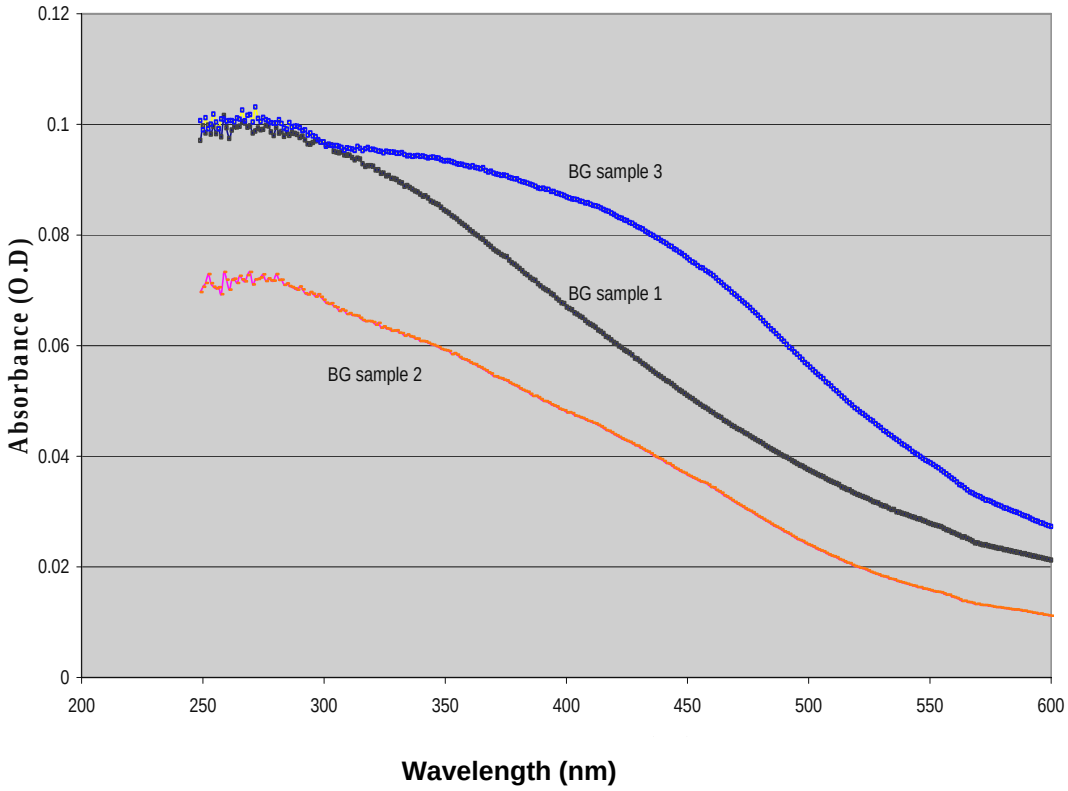


Figure 4.10 Representative plots of absorbance spectrum of washed spores of *B. globigii* samples (i), (ii) and (iii). Absorbance (O.D) versus wavelength (nm) is shown. Scanning wavelength was from 200-600 nm.

4.2.2.1 Quantum efficiency of unwashed and washed *B. globigii* (i) spores

We measured QE of unwashed and washed *B. globigii* (i) spores in three different excitation wavelengths. In Figure 4.11 we show representative fluorescence spectrum of unwashed and washed redried spores at excitation wavelengths 280, 360 and 400 nm. The emission ranges were 290-570, 370-700 and 410-700 nm. Similar plots were obtained with other *B. globigii* spores (Figures not shown). The difference in peak was due to the difference in the number or the concentration of samples investigated.

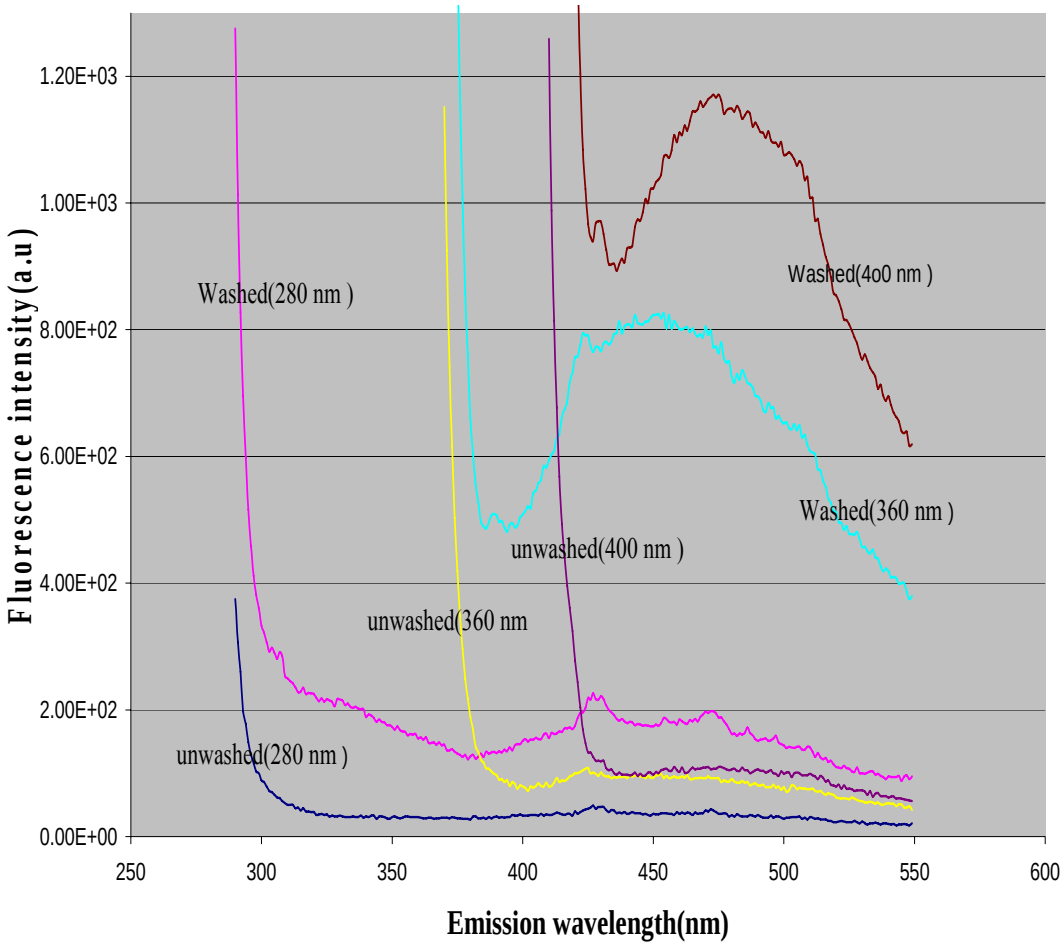


Figure 4.11 Representative fluorescence spectrum of unwashed and washed *B. globigii* (i) spores, measured at excitation wavelengths 280, 360 and 400 nm. The emission ranges were 290-570, 370-700 and 410-700 nm.

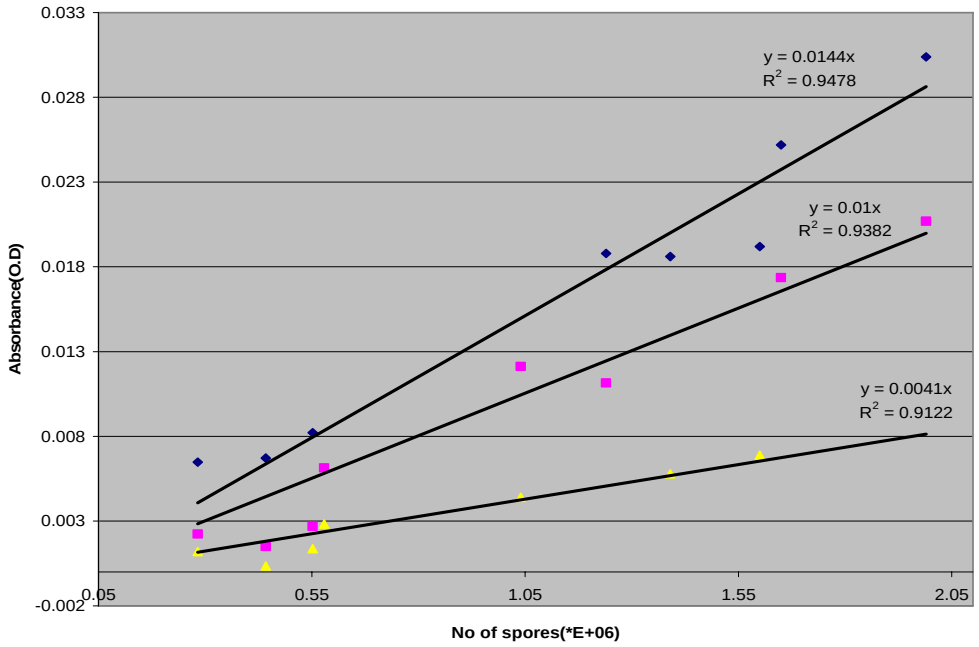
As explained in the text, the absorbance and the corrected integrated fluorescence intensity of various preparations of *B. globigii* (i) spores (unwashed) were calculated and is tabulated as shown in Table 4-4.

Table 4-4 The optical measurements of unwashed *B. globigii* (i) spores

Trial No	**No of Spores (C)**	**Absorbance(O.D) at excitation wavelength**			**Corrected integrated fluorescence intensity (counts) at excitation wavelength**		
		280 nm	**360 nm**	**400 nm**	**280 nm**	**360 nm**	**400 nm**
1	2.83E+05	0.0065	0.0022	0.0012	7.73E+03	2.13E+04	2.28E+04
2	5.52E+05	0.0082	0.0027	0.0014	1.27E+04	3.94E+04	2.58E+04
3	5.79E+05	0.0178	0.0061	0.0028	2.22E+04	6.14E+04	5.17E+04
4	1.39E+06	0.0283	0.0089	0.0058	3.08E+04	9.99E+04	8.72E+04
5	1.60E+06	0.0192	0.0109	0.0069	2.76E+04	1.01E+05	9.91E+04
6	1.04E+06	0.0201	0.0121	0.0044	2.15E+04	5.30E+04	4.67E+04
7	4.42E+05	0.0067	0.0015	0.0004	4.87E+03	9.76E+03	9.77E+03
8	1.99E+06	0.0304	0.0207	0.0157	2.84E+04	1.88E+05	2.50E+05
9	1.65E+06	0.0252	0.0174	0.0133	2.93E+04	1.64E+05	1.92E+05
10	1.24E+06	0.0188	0.0112	0.0082	2.60E+04	1.25E+05	1.31E+05

Based on the optical data shown in Table 4-4, we plotted the absorbance (O.D) versus the number of spores for each sample measurements as shown in Figure 4.12(A). In Figure 4.12(A) the absorbance versus number of spores of *B. globigii* (i) spores (unwashed) at 280 360, and 400 nm, is displayed. The absorption cross section for *B. globigii* (i) is obtained from the slope of the plots as 14.4E-09, 10E-9 and 4.1E-9 mm^2/spore at 280, 360, and 400 nm excitation respectively. The integrated fluorescence intensity (counts) of *B. globigii* (i) spores is shown versus absorption (O.D) in Figure 4.12(B). The line is linear least squares fit. The slopes are proportional to the QE of *B. globigii* (i) spores (unwashed) at the selected excitation wavelengths.

A



B

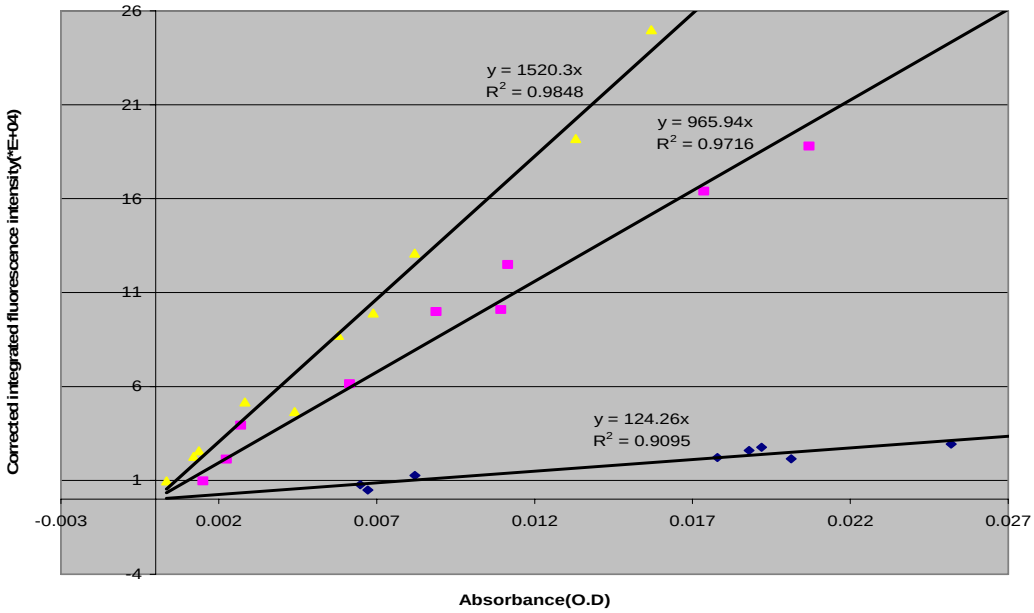


Figure 4.12(A-B**)** The QE of unwashed *B. globigii* (i) spores at 280, 360 and 400 nm excitation wavelengths. In 4.12(A) the absorbance (O.D) as function of number spores is displayed and in 4.12(B) the integrated fluorescence intensity (counts) of *B. globigii* (i) spores is shown versus absorption (O.D). The lines are linear least squares fit represented by the corresponding colours blue, red and yellow respectively. The slopes are proportional to the QE of *B. globigii* (i).

We also measure the absorbance and fluorescence of *B. globigii* (i) after being washed and redried. Below in Table 4-5 we show the optical measurement of washed and redried spores at excitation wavelength 280,360 and 400 nm.

Table 4-5 The optical measurements of washed and redried spores of *B. globigii* (i)

Trial No	**Absorbance (O.D) at excitation wavelength**			**Corrected integrated fluorescence intensity (counts) at excitation wavelength**		
	280 nm	**360 nm**	**400 nm**	**280 nm**	**360 nm**	**400 nm**
1	0.0775	0.0598	0.0458	3.89E+04	1.27E+05	1.51E+05
2	0.0871	0.0631	0.0469	3.88E+04	1.01E+05	1.08E+05
3	0.0466	0.0349	0.0263	1.73E+04	5.65E+04	6.57E+04
4	0.0345	0.0283	0.0201	1.56E+04	4.90E+04	5.37E+04
5	0.0623	0.0475	0.0357	2.74E+04	7.07E+04	8.40E+04
6	0.0990	0.0768	0.0603	4.21E+04	1.38E+05	1.72E+05
7	0.0701	0.0384	0.0281	3.23E+04	7.98E+04	8.69E+04
8	0.1022	0.1012	0.0801	6.80E+04	1.95E+05	2.15E+05

In Figure 4.13 the integrated fluorescence intensity (counts) of washed and redried spores of *B. globigii* (i) is shown versus absorption (O.D) at excitation wavelengths 280, 360 and 400 nm. The line is linear least squares fit. The slopes are proportional to the QE of *B. globigii* (i) spores at the selected excitation wavelengths.

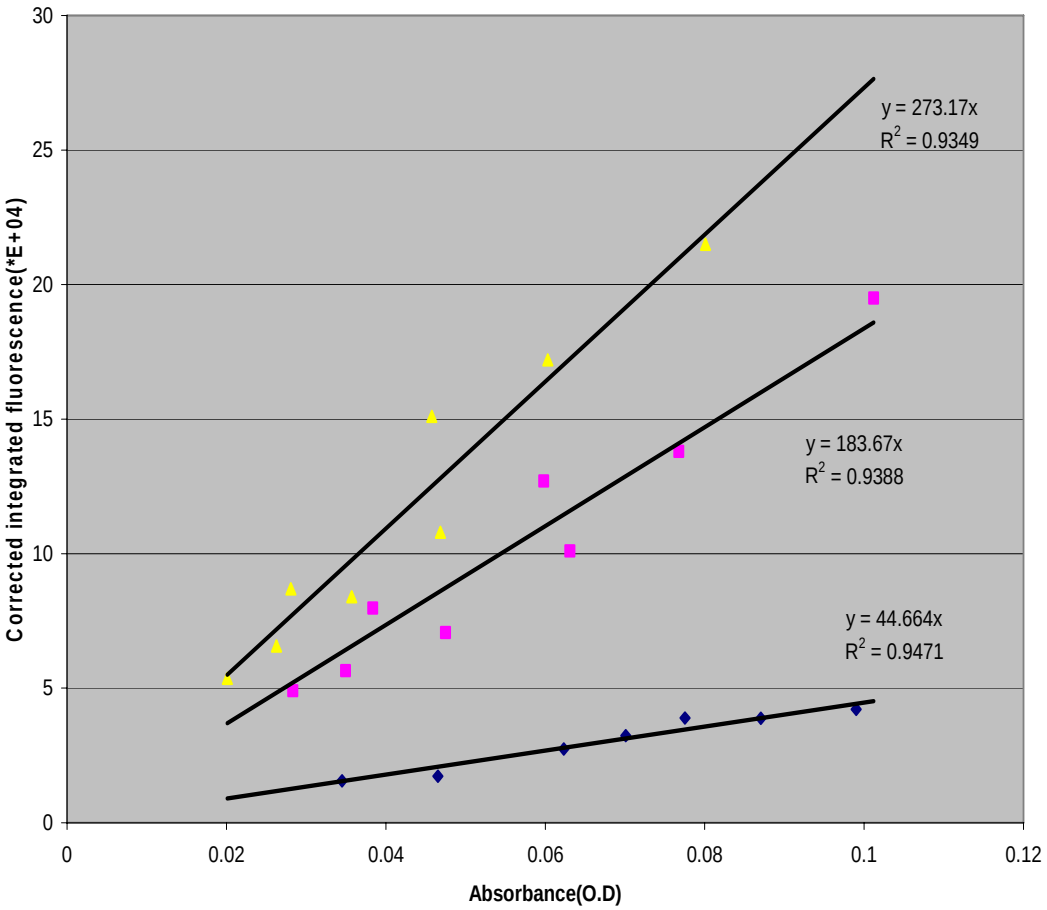


Figure 4.13 The QE of washed *B. globigii* (i) spores at excitation wavelengths 280, 360 and 400 nm. The integrated fluorescence intensity (counts) of *B. globigii* (i) spores is shown versus absorption (O.D). The line is linear least squares fit and is represented by colours: blue, red and yellow. The slopes are proportional to the QE of *B. globigii* (i) spores.

Below in Figure 4.14 we show the QE comparison plot at excitation wavelength 400 nm for a washed and unwashed sample.

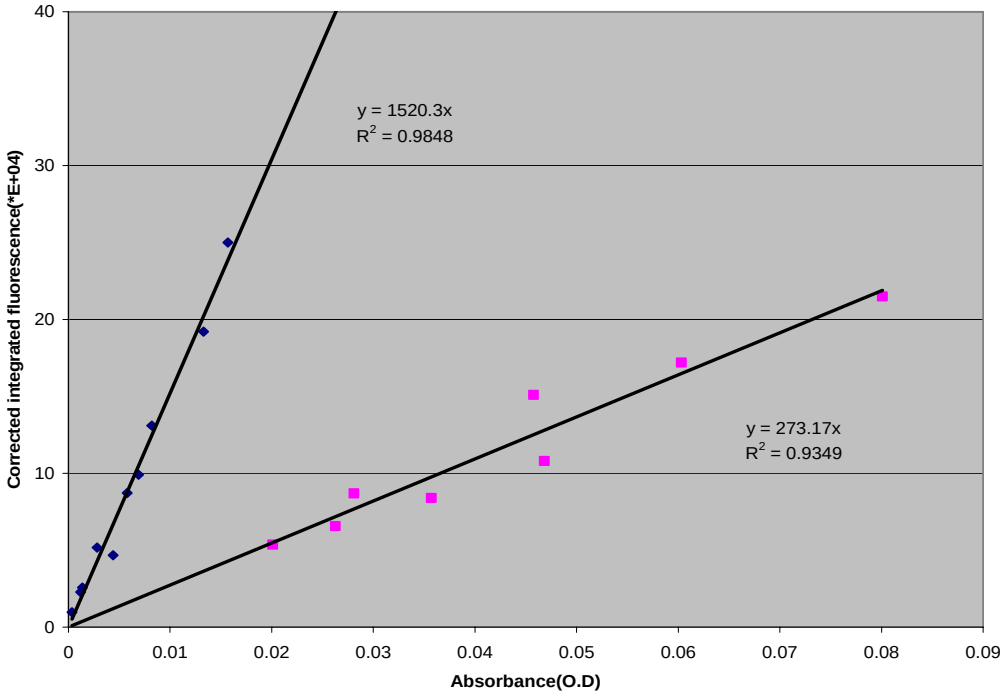


Figure 4.14 Comparison of fluorescence between washed (red) and unwashed (blue) *B. globigii* (i) spores at 400 nm excitation wavelength. The corrected integrated fluorescence intensity (counts) of *B. globigii* (i) spores is shown versus absorption (O.D). The line is linear least squares fit. The slopes are proportional to the QE of *B. globigii* (i) spores at 400 nm.

4.2.2.2 Quantum efficiency of unwashed and washed *B. globigii* (ii) Spores

As explained in the previous section, here also we measured QE of an outside made *B. globigii* (ii) spores. In the following, we present the corresponding optical data measurements for unwashed and washed *B. globigii* (ii) spores. In Table 4-6 we show the optical measurement of unwashed *B. globigii* (ii) spores made at excitation wavelengths 280, 360 and 400 nm.

Table 4-6 Measured absorbance (O.D) and integrated fluorescence intensity (counts) of unwashed *B. globigii* (ii) spores at 280, 360, 400 nm excitation wavelengths

Trial No	No of spores/mm^2	Absorbance (O.D) at excitation wavelength			Corrected integrated fluorescence intensity(counts) at excitation wavelength		
		280 nm	360 nm	400 nm	280 nm	360 nm	400 nm
1	1.22E+06	0.0123	0.0046	0.0025	1.03E+04	1.12E+04	9.71E+03
2	1.82E+06	0.0209	0.0101	0.0077	2.75E+04	5.73E+04	5.89E+04
3	2.96E+06	0.0212	0.0078	0.0058	1.26E+04	4.32E+04	4.15E+04
4	1.96E+06	0.0241	0.0111	0.0079	2.83E+04	6.09E+04	4.67E+04
5	1.99E+06	0.0423	0.0099	0.0067	4.57E+04	4.49E+04	4.44E+04
6	2.08E+06	0.0207	0.0082	0.0063	2.06E+04	3.26E+04	3.02E+04
7	2.38E+06	0.0246	0.0118	0.0087	3.29E+04	6.80E+04	6.08E+04
8	2.75E+06	0.0542	0.0132	0.0100	6.15E+04	1.38E+05	1.33E+05
9	4.07E+06	0.0407	0.0154	0.0097	4.62E+04	8.65E+04	7.76E+04
10	3.02E+06	0.0302	0.0108	0.0073	9.57E+04	2.01E+05	1.74E+05

In Figure 4.15 (Top) we show plots of the absorbance versus number of spores of *B. globigii* (ii) spores (unwashed). The absorption cross sections of spores are obtained from the slope of the plots as 10E-09, 4.1E-09 and 2.60E-09 mm^2/spore at 280, 360 and 400 nm respectively. The integrated fluorescence intensity (counts) of unwashed *B. globigii* (ii) spores is shown versus absorption (O.D) in Figure 4.15 (Bottom). The line is linear least squares fit. The slopes are proportional to the QE of *B. globigii* (ii) spores (unwashed).

In Table 4-7 we show the optical measurement of washed and redried spores of *B. globigii* (ii) made at excitation wavelengths 280, 360 and 400 nm. The corrected integrated fluorescence intensity (counts) of the spores is shown versus absorption (O.D) in Figure 4.16. The line is linear least squares fit. The slopes are proportional to the QE of *B. globigii* (ii) spores (washed).

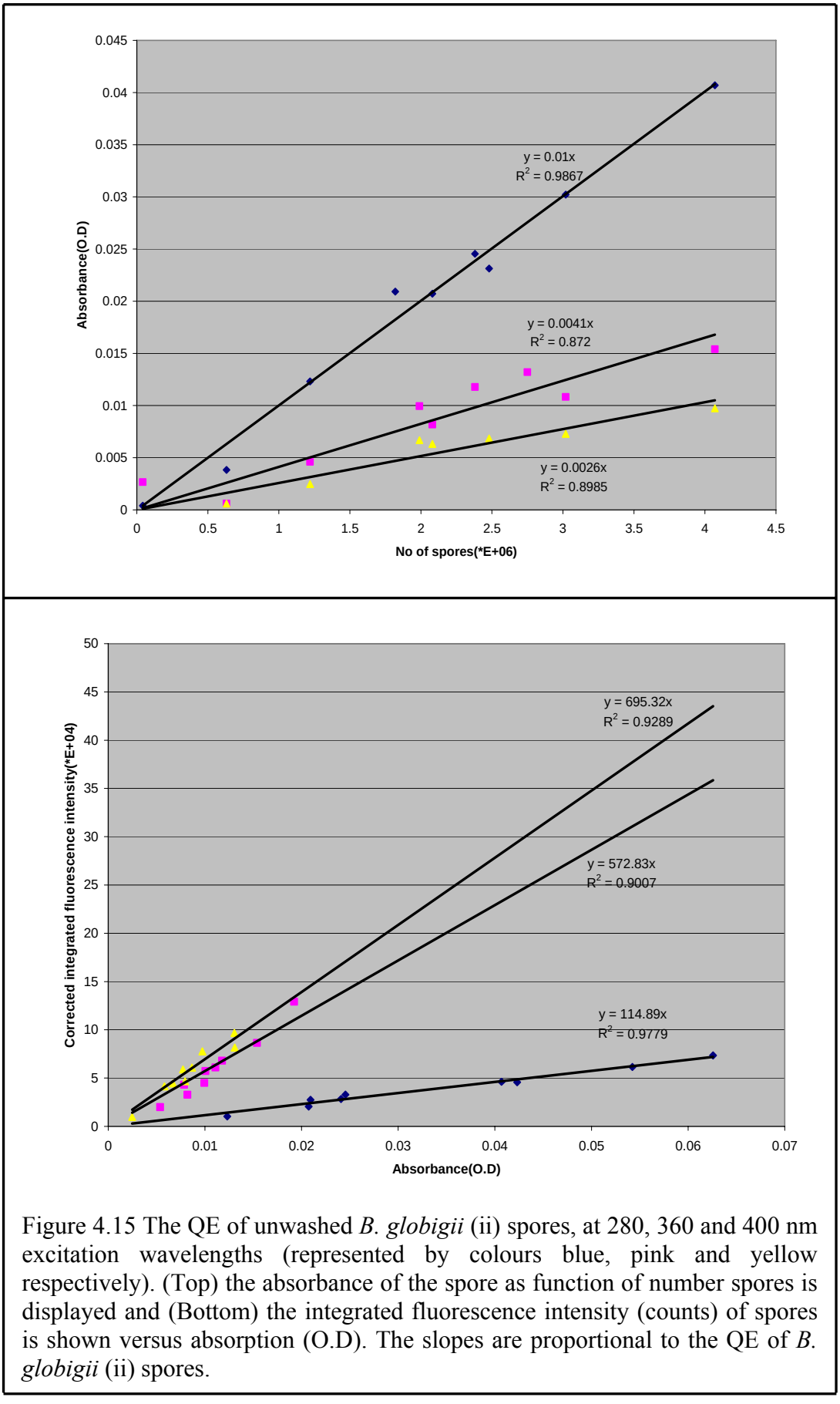


Figure 4.15 The QE of unwashed *B. globigii* (ii) spores, at 280, 360 and 400 nm excitation wavelengths (represented by colours blue, pink and yellow respectively). (Top) the absorbance of the spore as function of number spores is displayed and (Bottom) the integrated fluorescence intensity (counts) of spores is shown versus absorption (O.D). The slopes are proportional to the QE of *B. globigii* (ii) spores.

Table 4-7 The optical measurement of washed *B. globigii* (ii) spores at excitation wavelengths 280, 360 and 400 nm

Trial no	Absorbance (O.D) at excitation wavelength			Corrected integrated fluorescence intensity(counts) at excitation wavelength		
	280 nm	360 nm	400 nm	280 nm	360 nm	400 nm
1	0.0618	0.0459	0.0368	4.22E+04	1.04E+05	1.07E+05
2	0.0623	0.0321	0.0265	3.45E+04	6.57E+04	6.65E+04
3	0.0745	0.0580	0.0472	5.85E+04	1.34E+05	1.21E+05
4	0.0912	0.0612	0.0534	6.77E+04	1.39E+05	1.36E+05
5	0.0989	0.0412	0.0350	5.99E+04	8.61E+04	7.31E+04
6	0.0823	0.0543	0.0472	5.45E+04	1.27E+05	1.17E+05
7	0.0412	0.0312	0.0218	2.42E+04	6.11E+04	5.55E+04
8	0.0312	0.0268	0.0190	2.70E+04	3.41E+04	2.91E+04

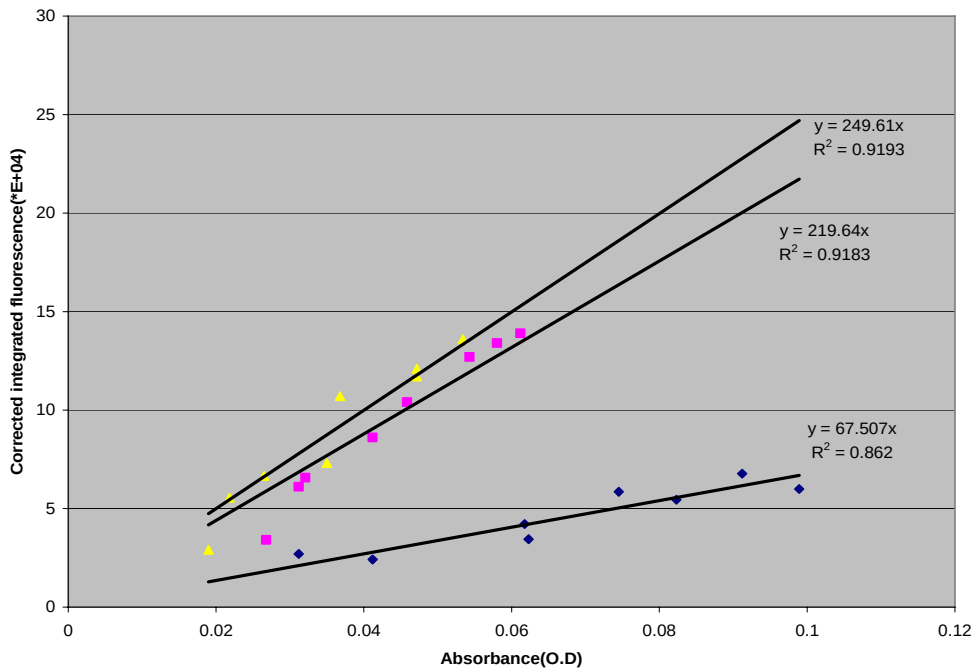


Figure 4.16 Fluorescence versus absorbance plots of washed and redried spores of *B. globigii* (ii) at 280, 360 and 400 nm excitation wavelengths (represented by colours blue, red and yellow respectively). The integrated fluorescence intensity (counts) of *B. globigii* (ii) spores is shown versus absorption (O.D). The line is linear least squares fit. The slopes are proportional to the QE of *B. globigii* (ii) spores.

4.2.2.3 Quantum efficiency of unwashed and washed *B. globigii* (iii) Spores

As explained in the text we have repeated the QE measurements of another outside made sample of *B. globigii*: the *B. globigii* (iii) spores. Table 4-8 is the optical measurements of unwashed spores at excitation wavelengths 280, 360 and 400 nm.

Table 4-8 The optical measurements of *B. globigii* (iii) spores (unwashed) at excitation wavelengths 280, 360 and 400 nm

Trial No	**No of spores**	**Absorbance(O.D) at excitation wavelength**			**Corrected integrated fluorescence intensity (counts) at excitation wavelength**		
		280 nm	**360 nm**	**400 nm**	**280 nm**	**360 nm**	**400 nm**
1	1.01E+06	0.0124	0.0098	0.0087	2.43E+04	9.95E+04	4.41E+04
2	2.26E+06	0.0297	0.0196	0.0191	5.23E+04	7.66E+04	9.98E+04
3	1.51E+06	0.0224	0.0222	0.0145	4.05E+04	1.05E+05	7.67E+04
4	1.02E+06	0.0159	0.0169	0.0109	1.93E+04	9.30E+04	5.35E+04
6	1.15E+06	0.0190	0.0131	0.0105	2.34E+04	9.64E+03	3.49E+04
6	9.16E+05	0.0155	0.0135	0.0054	2.08E+04	2.50E+04	2.08E+04
7	9.33E+05	0.0131	0.0077	0.0081	2.89E+04	2.28E+04	3.87E+04
8	1.57E+06	0.0211	0.0106	0.0149	3.88E+04	5.26E+04	6.73E+04
9	1.83E+06	0.0299	0.0094	0.0161	4.78E+04	4.43E+04	6.56E+04

In Figure 4.17 (Top) we show the plots of the absorbance (O.D) versus number of spores of unwashed *B. globigii* (iii) spores. The absorption cross section is obtained from the slopes as 15.9E-09, 10.2E-09 and 9.00E-09 mm^2/spore at excitation wavelengths 280 360, and 400 nm respectively. The integrated fluorescence intensity (counts) of the spores (unwashed) is shown versus absorption (O.D) in Figure 4.17 (Bottom). The line is linear least squares fit. The slopes are proportional to the QE of *B. globigii* (iii) spores (unwashed).

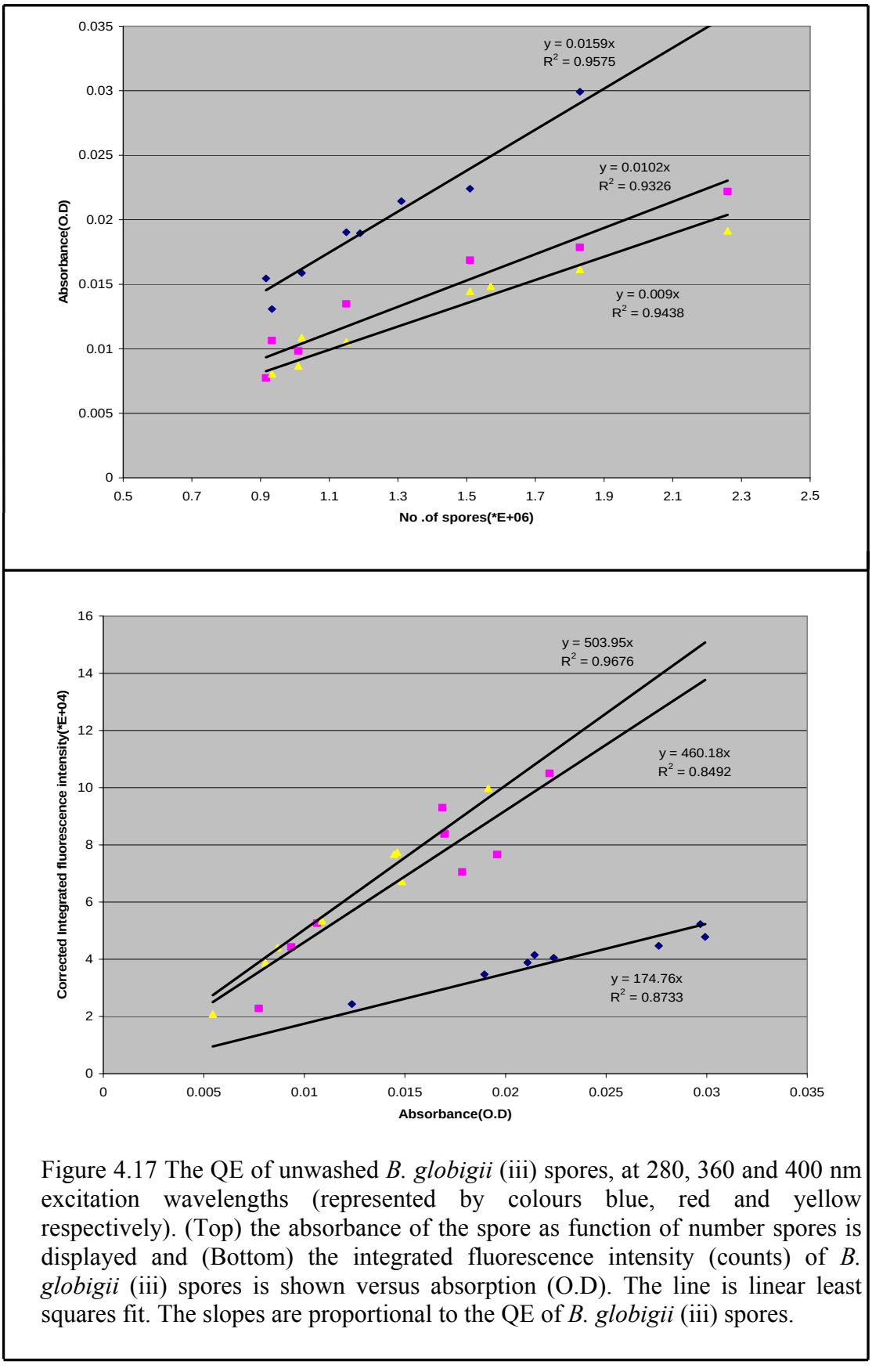


Figure 4.17 The QE of unwashed *B. globigii* (iii) spores, at 280, 360 and 400 nm excitation wavelengths (represented by colours blue, red and yellow respectively). (Top) the absorbance of the spore as function of number spores is displayed and (Bottom) the integrated fluorescence intensity (counts) of *B. globigii* (iii) spores is shown versus absorption (O.D). The line is linear least squares fit. The slopes are proportional to the QE of *B. globigii* (iii) spores.

We also measured the QE of washed and redried spores of *B. globigii* (iii) as function of its absorbance. In Table 4-9 we show the optical measurement of washed and redried spores at excitation wavelengths 280, 360 and 400 nm.

Table 4-9 The optical measurements of washed and redried spores of *B. globigii* (iii) at excitation wavelength 280, 360 and 400 nm

Trial No	**Absorbance(O.D) at excitation wavelength**			**Corrected integrated fluorescence intensity (counts) at excitation wavelength**		
	280 nm	**360 nm**	**400 nm**	**280 nm**	**360 nm**	**400 nm**
1	0.0456	0.0432	0.0625	4.18E+04	3.39E+04	4.83E+04
2	0.0657	0.1123	0.0848	3.80E+04	1.10E+05	1.10E+05
3	0.0670	0.0790	0.0936	5.43E+04	5.58E+04	9.25E+04
4	0.1206	0.1067	0.0520	9.94E+04	9.58E+04	4.77E+04
5	0.0978	0.0671	0.0154	7.02E+04	5.86E+04	1.52E+04
6	0.1210	0.0342	0.0243	9.47E+04	1.77E+04	1.94E+04
7	0.1153	0.1070	0.0960	7.60E+04	1.04E+05	9.79E+04
8	0.1122	0.0909	0.0810	8.15E+04	1.07E+05	9.50E+04

In Figure 4.18 the integrated fluorescence intensity (counts) of *B. globigii* (iii) spores (washed*)* is shown versus absorption (O.D). The line is linear least squares fit. The slopes are proportional to the QE of *B. globigii* (iii) spores (washed).

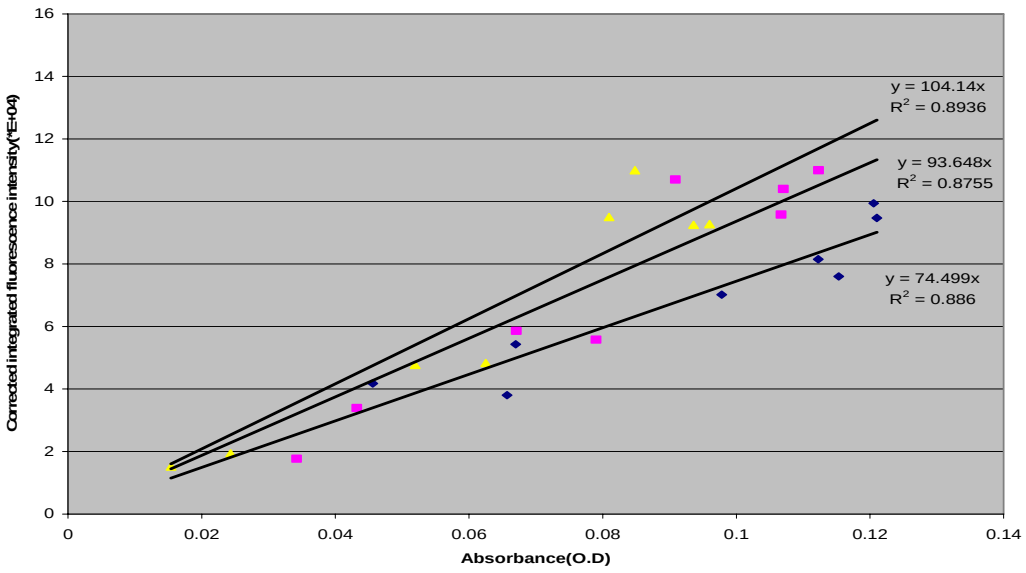


Figure 4.18 Fluorescence versus absorbance plots of washed *B. globigii* (iii) spores, at excitation wavelengths 280, 360 and 400 nm (represented by colours blue, red and yellow respectively). The integrated fluorescence intensity (counts) of *B. globigii* (iii) spores is shown versus absorption (O.D). The line is linear least squares fit. The slopes are proportional to the QE of *B. globigii* (iii) spores**.**

In Figure 4.19 we show a comparative plot for the integrated fluorescence intensity (counts) of *B. globigii* (iii) spores (unwashed and washed) versus absorption (O.D) at excitation 280 nm. It is evident from the figure that after washing and redried the samples, the QE is reduced considerably.

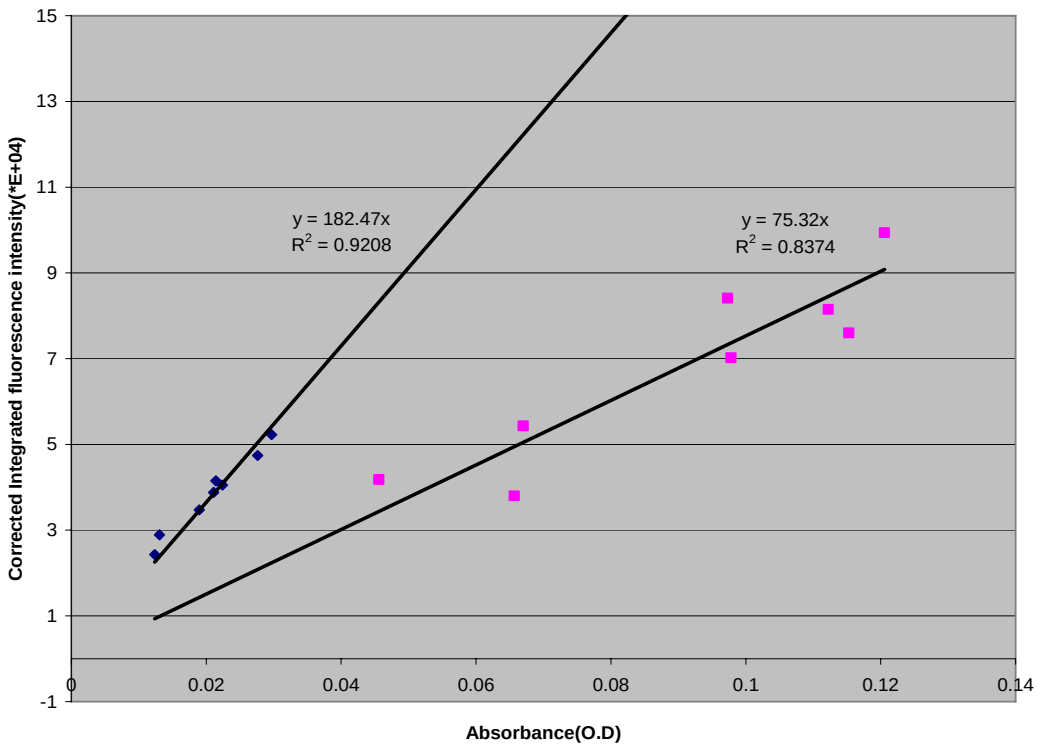


Figure 4.19 Comparison of fluorescence of unwashed (blue) and washed (red) spores of *B. globigii* (iii) spores. The corrected integrated fluorescence intensity (counts) versus absorption (O.D) is shown for excitation wavelength 280 nm. The line is linear least squares fit. The slopes are proportional to the QE of *B. globigii* (iii) spores.

4.2.2.4 Quantum efficiency of unwashed *B. cereus* spores

In this section we show the results of the QE measurements of *B. cereus* spores. For this experiment only unwashed samples were used since the supplied samples were in their pure form. In Figure 4.20 we show the representative fluorescence spectrum of unwashed spores at 280 and 300 nm excitation wavelengths. The emission wavelength range was from 290-550 and 310-590 nm respectively.

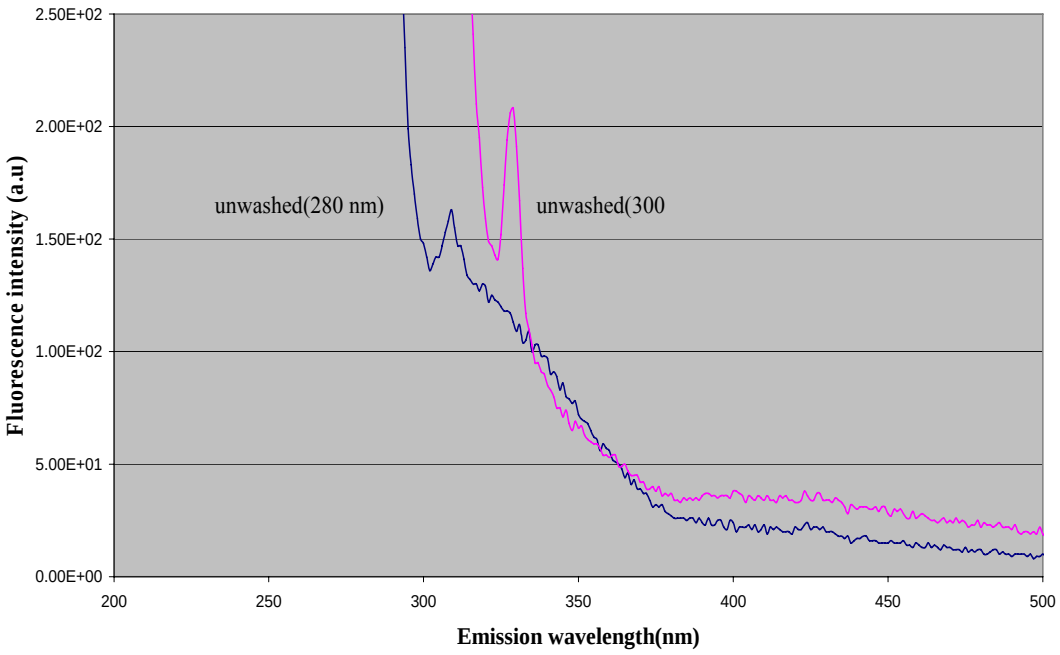


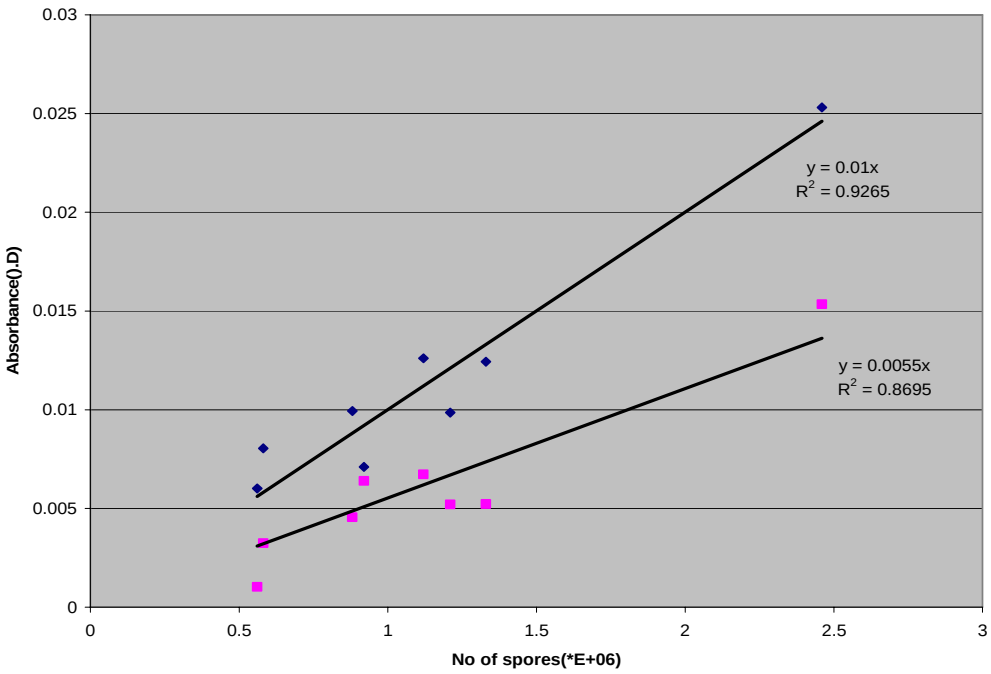
Figure 4.20 Representative fluorescence spectrum of unwashed *B. cereus* spores at excitation wavelengths 280 and 300 nm. The emission wavelength ranges were from 290-550 and 310-590 nm.

As explained in the previous sections, here too we measured the QE of *B. cereus* as a function of number of spores. In Table 4-10 we present the optical measurement of *B. cereus* for calculating QE. Notice that in the case of *B. cereus* we only calculated the QE at excitation wavelengths 280 and 300 nm.

Table 4-10 Measured absorbance and integrated fluorescence intensity (counts) of unwashed *B. cereus* spores at excitation wavelengths 280 and 300 nm

Trial No	**No of spores**	**Absorbance (O.D) at excitation wavelength**		**Corrected integrated fluorescence intensity (counts) at excitation wavelength**	
		280 nm	**300 nm**	**280 nm**	**300 nm**
1	5.61E+05	0.0060	0.0010	6.59E+03	1.07E+04
2	1.33E+06	0.0124	0.0052	2.13E+04	2.63E+04
3	1.21E+06	0.0099	0.0052	2.54E+04	3.53E+04
4	1.12E+06	0.0126	0.0067	1.31E+04	1.55E+04
5	5.81E+05	0.0081	0.0032	1.21E+04	1.55E+04
6	8.81E+05	0.0099	0.0046	9.74E+03	1.37E+04
7	2.46E+06	0.0253	0.0153	4.42E+04	6.77E+04
8	7.92E+05	0.0071	0.0064		

In Figure 4.21 (Top) we show the plots of the absorbance (O.D) versus number of spores of unwashed *B. cereus* spores at excitation wavelengths 280 and 300 nm. The absorption cross section is obtained from the slope of the plots as 10E-09 and 5.5E-09 mm^2/spore respectively. In Figure 4.21 (Bottom) the corrected integrated fluorescence versus absorbance plots for spores at 280 and 300 nm is shown. We could not find the QE at 360 and 400 nm since the fluorescence is very weak at these wavelengths. There was a surprising finding in the case of *B. cereus* spores. It produced a well defined fluorescence spectrum with high fluorescence intensity at excitation wavelengths near 280 nm and the QE at this wavelength is found to be about 0.40, a factor of 4 higher compared to other simulants.



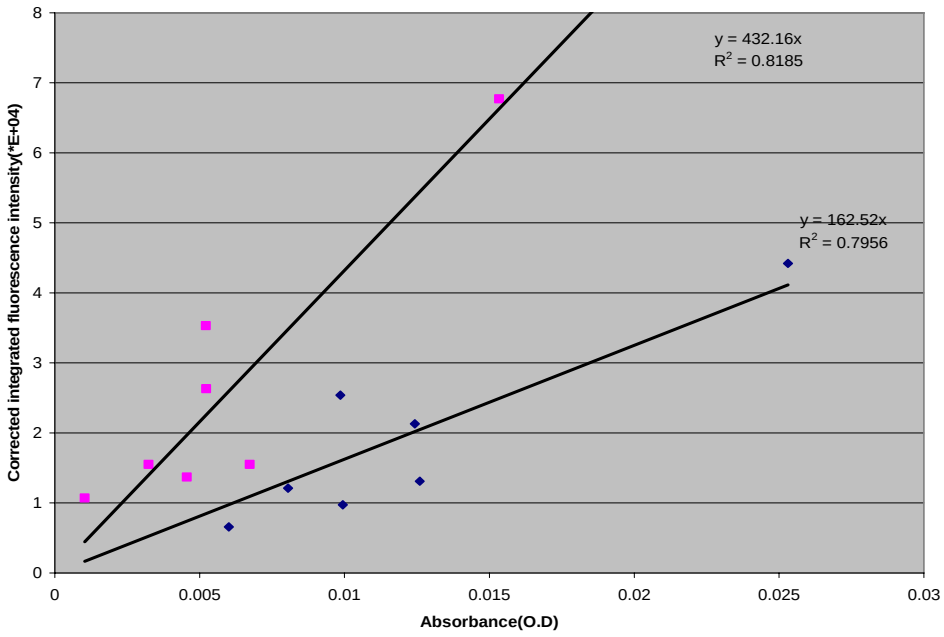


Figure 4.21 The QE of unwashed *B. cereus* spores. (Top) The absorbance of the spore as function of number spores. (Bottom) the corrected integrated fluorescence intensity of *B. cereus* is shown versus absorption (O.D). The line is linear least squares fit represented by colours blue and red for excitation 280 and 300 nm respectively. The slopes are proportional to the QE of *B. cereus*.

In Table 4-11 we show the summary of the optical property measurements of the dried *Bacillus* spores investigated. Results of both unwashed and washed samples measurements are shown. The absorption cross sections of unwashed samples were calculated from the slope of the plots of absorbance versus the number of spores. The QE of the samples were calculated relative to the known QE value of anthracene in ethanol.

Table 4-11 shows the summary of optical property measurements of the *Bacillus* samples investigated.

Samples	**Absorption cross section (mm^2/spore) at excitation wavelength**			**Quantum efficiency at excitation wavelength**		
	280 nm	**360 nm**	**400 nm**	**280 nm**	**360 nm**	**400 nm**
B. globigii (i) unwashed	14.4E-09	10E-09	4.10E-09	0.12±0.04	0.47± 0.14	0.72 ±0.14
B. globigii (i) washed				0.04 ± 0.01	0.09± 0.03	0.13± 0.04
B. globigii (ii) Unwashed	10E-09	4.1E-09	2.60E-09	0.11 ±0.03	0.28± 0.08	0.33± 0.10
B. globigii (ii) washed				0.06 ± 0.02	0.11± 0.03	0.12± 0.04
B. globigii (iii) Unwashed	15.9E-09	10.2E-09	9.00E-09	0.17± 0.03	0.23± 0.07	0.24±0.07
B. globigii (iii) washed				0.07±0.02	0.05±0.02	0.05±0.02
B. cereus unwashed	10E-09	5.5E-09 at (300 nm)		0.41±0.12	0.23±0.07 at (300 nm)	

4.2.3 Discussions on the Quantum Efficiency Measurements of *Bacillus* Spores

The QE measurements of samples are complicated by its highly absorbing and scattering properties [66]. The QE of a biological sample is highly dependent on temperature. All of our fluorescence measurements were done at room temperature. A decrease in the fluorescence QE with temperature has been reported for the indole chromophore as a consequence of temperature-dependence of no radiative processes [144, 145]. As explained in the text, we measured fluorescence QE of the selected samples with respect to a standard solution of anthracene in ethanol and we give a brief discussion below.

It is evident from the Table 4-11 that different *Bacillus* spores behave differently with the same excitation range in terms of its optical behaviour and that may be a useful characterisation for detection and identification of *Bacillus* samples in an urgent need of treatment. Also we noticed that the QE of sample lowered as we washed the sample. This may be due to water going into the coat surface and making anatomical changes that hinder or lower the fluorescence nature of the sample.

The standard solution of anthracene in ethanol produced well defined absorption spectrum in the range of 300-400 nm excitations light. For instance, see the Figure 4.4 for an absorption spectrum of anthracene in ethanol at a 2.21E-04 mol/litre concentration. The main absorption peak produced at excitation wavelength 355.52 nm. The absorbance was 1.6153 O.D.

The fluorescence spectrum for instance, at 360 nm excitation was mirror symmetry of absorption spectrum with emission maximum at 402 nm (Figure 4.5). As expected for a simple molecule like anthracene, it is evident from Figures 4.4 and 4.5 a mirror symmetry nature of absorbance and fluorescence spectra. Ethanol itself did not produce much fluorescence where anthracene absorbs, except the Raman peak near 250 nm excitation wavelength (Figure 4.6(A-B)). We used the absorbance of anthracene in ethanol at five different excitation wavelengths, 320, 330, 340, 350 and 360 nm where it absorbed more (see Table 4-1). The corrected fluorescence intensities were calculated at these excitation wavelengths (see Table 4-2). Corrected integrated fluorescence of anthracene in ethanol as functions of absorbance (O.D) at 320, 330, 340, 350 and 360 nm was plotted (see the Figure 4.7). Each line is linear

least squares fit. Each slope is proportional to the known the QE of anthracene in ethanol at the excitation wavelength selected [138].

To check the machine response to the various excitation wavelengths selected, a response graph was plotted with the measured fluorescence (counts) versus absorbance (O.D) against excitation wavelengths ranging from 320 to 360 nm. It was interesting to notice that the response graph (Figure 4.8) was nearly linear, suggesting that fluorescence intensities are independent of excitation wavelength of a simple molecule like anthracene. Therefore, we chose the slope value of the plot (Figure 4.8) at 320 nm excitation wavelength for calculating QE of spores at 280 nm and the slope value at 360 nm for calculating QE at 360 and 400 nm. Since below 320 and above 360 nm excitation, the response graph (Figure 4.8) no longer yields a linear relationship between the measured fluorescence versus absorbance against excitation wavelength. If we choose the slope of the response plot to calculate the slope value for 280 and 400 nm, we will come up with severe anomalies with QE values for spores. We also noticed that there was not much variation to be observed if we use the slope value of 320 and 360 nm for calculating QE at 280 and 400 nm.

The QE of *Bacillus* samples studied show unique properties that the main contribution to QE is due to the tryptophan absorbance near 290 nm excitation. Both washed and unwashed samples showed more or less the same QE at 280 nm, suggesting that fluorescence QE is independent of the excitation wavelengths in the range 280–290 nm, whereas the QE value varies as the excitation increases. Our suggestions were supported by the other studies by Saavedra et al. [146]. In a structural study of the proteins of the *Bacillus stearothermophilus* based on intrinsic tryptophan fluorescence, they reported that the fluorescence QE were independent of the excitation wavelength in the range 280-310 nm.

The QE measurements of *Bacillus* spores investigated produced very interesting results. As seen from the Table 4-11, the measured QE values of the different sources of *B. globigii* samples at excitation wavelengths 280, 360 and 400 nm showed differences in both the unwashed and washed redried spores. Unwashed *B. globigii* samples at excitation wavelengths 280 nm showed comparable values, even though *B. globigii* (iii) showed a factor of 1.5 differences between other two sources. Notice that even when we used different sources of unwashed *B. globigii*,

the average QE at 280 nm is 0.13±0.03, a value comparable to the QE measurement of tryptophan and supporting our suggestion that main contribution to QE of *Bacillus* spores at 280 nm is due to tryptophan. The QE values of the same samples in washed redried condition showed more or less the same values at excitation 280 nm (0.06±0.02), but surprisingly this value is a factor of 2.2, lower than the values due to unwashed measurement. We suspect that the lowering of QE values at 280 nm in washed redried condition is due to the loss of some of tryptophan residue present at the outer coat layer of samples due to washing. There were other suggestions as well, like perhaps the unwashed sample may contain some growth medium that enhanced the fluorescence intensity in the unwashed condition, thereby showing an increase in QE.

Looking at the QE values for unwashed spores at 360 nm excitation, *B. globigii* (ii) *and* (iii) showed comparable values (0.28 and 0.23), but *B. globigii* (i) showed a factor of 1.8 higher at this excitation (0.47). In the case of washed redried samples at this wavelength, there isn't much correlation between QE values for *B. globigii* samples: 0.09, 0.11 and 0.05 for *B. globigii* (i), (ii) and (iii) respectively. The values for *B. globigii* (i) and (ii) are comparable but for sample (iii), it is a factor of 2 lower than the other samples. Comparing the QE at 360 nm for washed and unwashed measurements, we can see that the three sources of *B. globigii* yield different values; factors 5.2, 2.6 and 5 lower for washed *B. globigii* (i), (ii) and (iii) respectively.

When comparing QE of unwashed samples between 280 and 360 nm, it was interesting to notice that *B. globigii* (i) showed a factor of 4 higher at 360 nm, whereas *B. globigii* (ii) and (iii) showed a factor 2.6 and 1.3 higher at this excitation wavelength. We can also compare QE of washed samples at this wavelength with values at 280 nm. As seen from the Table 4-11, at 360 nm the QE value for all *B. globigii* samples showed a factor of 2 higher that the values at 280 nm, indicating the fact that the same source of the fluorescence components is present in the *B. globigii* samples causing an increase of QE values at 360 nm excitation.

Again, it is noticeable that at 400 nm, *B. globigii* samples showed some anomalous results both at unwashed and washed redried measurements. Between the QE measurements at 280 and 400 nm, *B. globigii* (i) showed a factor of 6 higher at 400 nm for unwashed and a factor of 3 higher for washed. Looking at 400 nm alone

between unwashed and washed, this sample showed an unexpected difference in QE of the order of a factor 5.5 lower for washed redried measurements.

For *B. globigii* (ii), between 280 and 400 nm excitation wavelengths, the QE value is a factor 3 higher at 400 nm for unwashed and a factor 1.8 higher for washed samples and the QE value at 400 between unwashed and washed showed a factor 2.8 lower for washed samples. Again, for *B. globigii* (iii), QE values for excitation between 280 and 400 nm, showed only a factor 1.4 higher at 400 nm for unwashed and a factor of 1.4 higher at 280 nm. This lowering of QE at 400 nm for this sample was not expected. We also noticed in this case, that the QE value at 400 nm was lowered by a factor 4.8 for washed samples.

The QE measurement for a non *B. globigii* sample (eg *B. cereus*) showed surprising results. As seen from the Table 4-11, at 280 and 300 nm excitation wavelengths, it yielded a value of 0.41 and 0.23 respectively. This was not expected compared to *B. globigii* samples measurements.

All the differences in QE are within our conservative error except the QE value for *B. globigii* (i) at 400 nm. We believe that a substantial portion of the QE at 280 nm is due to a large absorption cross section at this excitation, mainly due to tryptophan, which is thought to be primarily responsible for the fluorescence of *Bacillus* spores. It was interesting to notice that the absorption cross section of all samples measured at 280 nm excitation has almost the same value in the order of 10^{-9} mm^2 per spore and so qualitatively indicating that the same amount of tryptophan may be present in all *Bacillus* spores.

Regarding the detection limit in terms of absolute number of spores, for example, from Table 4-12 for the optical measurements of unwashed *B. globigii* (i) spores , that the number of spores selected range was from 2.83 E+05 – 1.24 E + 06 in a 2 mm diameter surface. In this case the absolute number of spores for the detection limit was 10,000 well agreeing with the number of bacteria that would cause diseases. Since the diameter of the light source that we used had a width of 1.3 mm for a slit width of 4 nm, we assumed that the entire incident light was well within the exposed area of the 2 mm surface. We counted the number of spores assuming that the spores were distributed uniformly over the area.

To reduce the number of spores in the area, we could have used a smaller exposed area for the sample, for example, 1 mm diameter exposure. But the problem here would have been that as the area was smaller, we would need a finer light beam to meet with the spores. One solution for this problem is to reduce the slit width from 4nm to lower value, such as 1nm or 2 nm, however this would lead to decreases the fluorescence intensity from the spores.

We have reported earlier that the room temperature multi-wavelength autofluorescence signature of wet bacterial endospores is considerably different from that of the dry endospores [141]. Small, but indistinct, differences are observed between the dry and the re-dried spores. The fluorescence spectra support the model of reversible water migration between inner spore compartments and the environment. If this is true, we should expect not a significant change in QE due to washing and redried measurements. But our measurements with washed redried samples showed substantial differences in the QE value of the order of a factor 5 lower than the unwashed measurements.

Marco et al. [147] reported that the dimension of individual *B. globigii* (i) spores decrease reversibly by 12 % in response to a change in the environment from fully hydrated to air-dried state establishing the fact that the dormant spore is a dynamic physical structure. Besides the size change causing a lower in QE, there may be other factors that come into play. Due to washing, the emission intensity will change since after being washed and redried, the spores will lose their homogeneous shape and be in a state of aggregation. Steven et al. reported that the total fluorescence for aggregates bacteria increases with particle size [65], supporting the suggestions of Westphal et al. [148] and Macro et al. [147] for characteristic size changes of samples due to hydration. But in our case, we noticed a lower QE (that is related to fluorescence and absorbance) value after we re-dried the sample, even though it was in aggregation state in which more fluorescence would be expected. We suspect one of reasons for the lower QE value for the samples, after being washed and re-dried, is due to the decrease in size and so producing less fluorescence intensity. At this stage no further explanation can be given in this section. More work need to be done to understand the environmental changes to the *Bacillus* spores fully.

4.3 Detection of *Bacillus* Spores from Different Substrates/Dust/Pollen and Effects of Washing on Fluorescence

In this section, we are presenting the various results obtained in the study of fluorescence measurements of *Bacillus* simulants from different environments and preparation conditions. For convenience this section is divided into several subsections including separate sections for PCA and cluster analysis results for each investigation.

4.3.1 Detection of Spores from Different Substrates

In this section we show the results of a comparative study of detection of samples investigated in two different substrates (plastic tape and quartz slide). In Figure 4.22(A-B) we show 2D fingerprints of two various substrates investigated for the collection and fluorescence measurements. The selection of the optimal substrate depends on the environment where the spore was noticed. Figure 4.22(A) is the fluorescence profile of the plastic tape substrate and (B) for quartz substrate. For comparison, fluorescence fingerprints of spores coated on these substrates are shown in Figure 4.22(C-D) as a representative measurement of *B. globigii* (i). The intensity is equally divided by 15 contour lines. The greatest intensity is blue and the least intensity is red. The intensity scale in Figure 4.22(A) and in Figure 4.22(C) is the same and the intensity scale in Figure 4.22(B) and in Figure 4.22(D) is the same. Similar results were obtained with other samples studied with the same tape and quartz slide (Figures not shown).

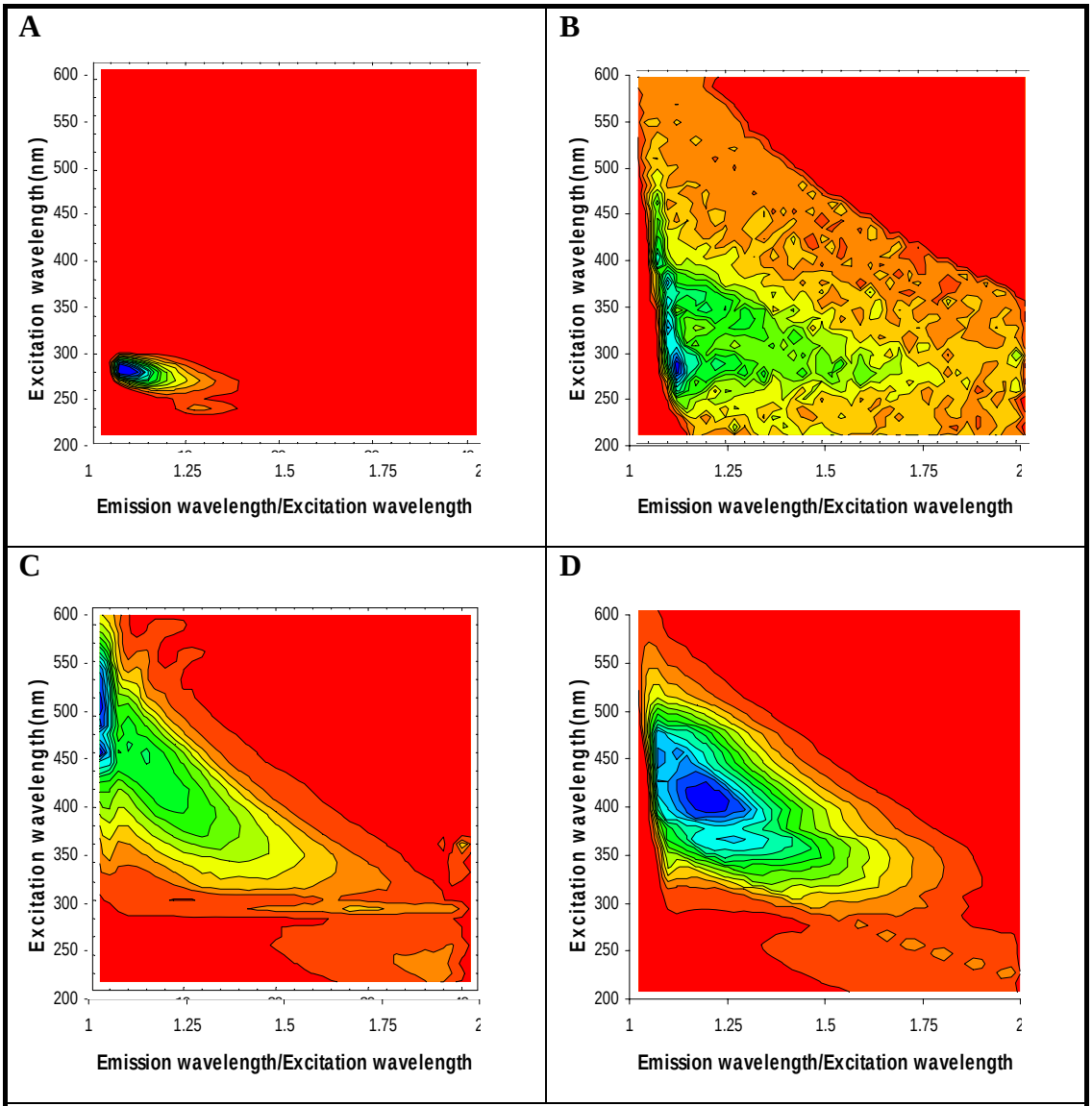


Figure 4.22 Fluorescence fingerprints of substrates and spores coated on substrates using excitation wavelengths from 200-600 nm. In Figure (A-B) fingerprints from tape and quartz slide is displayed. In Figure (C-D) fluorescence fingerprints of dry spores of *B. globigii* (i) on tape and quartz slide are displayed for comparison. Details of measurements are explained in the text. In all figures, the contours represent equal changes in the fluorescence intensity, and the peak intensity is normalised to 1.

4.3.2 Detection and Identification of Dry Spores from Dusts Background on Tape Substrate

In this section we show the results of our investigation with the spore forming *Bacillus* samples and the non spore samples like corn smut and ovalbumin. In sections 4.3.2.1 and 4.3.2.2 we will show the fluorescence measurements and the PCA analysis of fluorescence data respectively.

4.3.2.1 Spores Fluorescence

As explained in section 3.3.2 we measure the fluorescence signals from three *Bacillus* samples mixed with three different dust sources. In Figure 4.23(A-D) we show the 2D fluorescence fingerprints that is measured for dust A and the three different bacterial spores (*B. globigii* (i), *B. cereus* and *B. popilliae*) on dust A for excitation 200-600 nm. Similar results were obtained with two of other dust samples B and C with the same bacterial spores used with dust A (Figures not shown).

We also measured the fluorescence signals from three other samples without added dusts. We measure the fluorescence signals from a sample of *B. globigii* spores prepared from an outside source: *B. globigii* (ii). We also measure corn smut and ovalbumin. These fluorescence signals are shown in Figure 4.24(A-C).

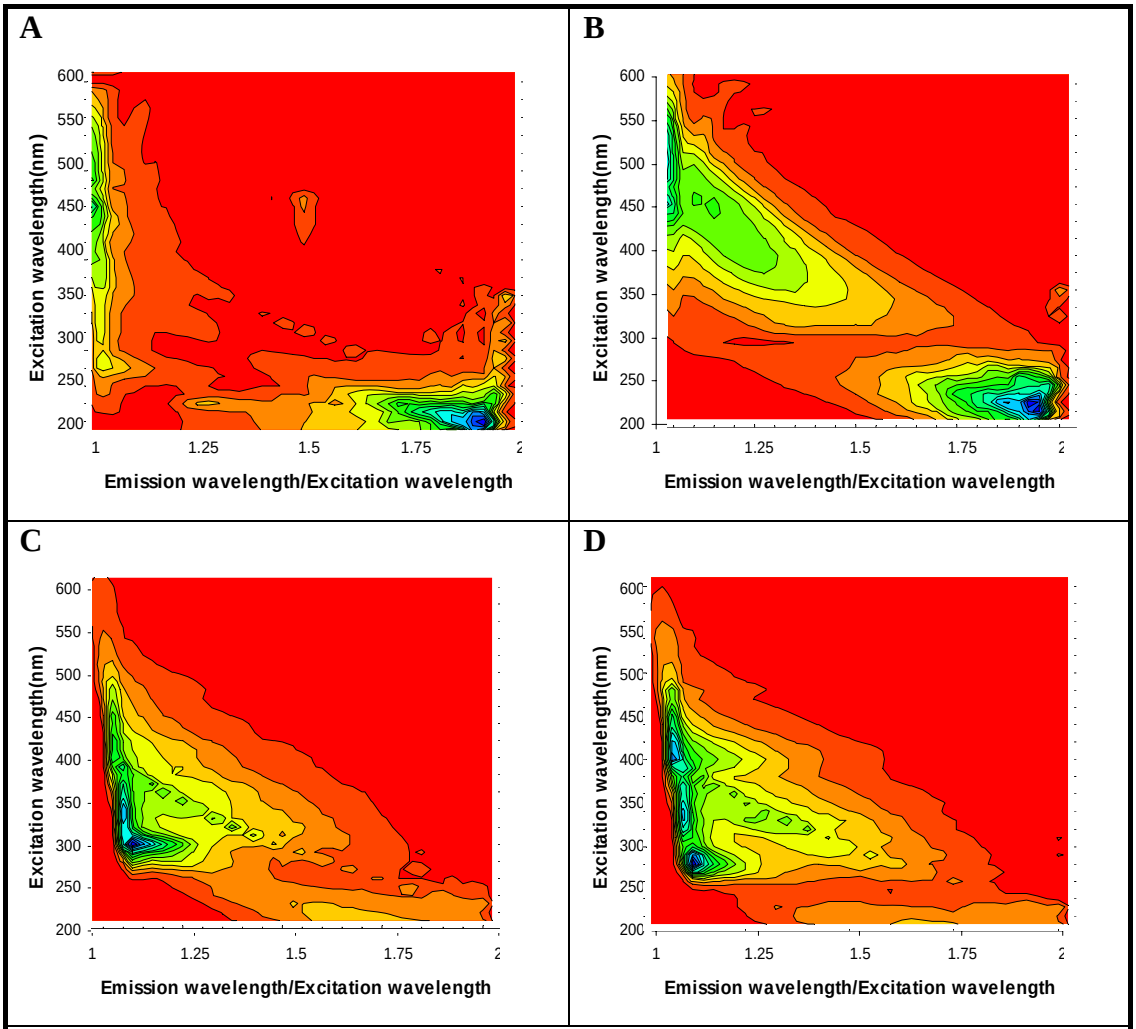


Figure 4.23 Fluorescence fingerprints of (A) dust A; (B) *B. globigii* (i) and dust A; (C) *B. cereus* and dust A; (D) *B. popilliae* and dust A using excitation wavelengths from 200-600 nm is displayed. The fluorescence signal is divided into 15 equally spaced contour lines. Blue represents the maximum intensity.

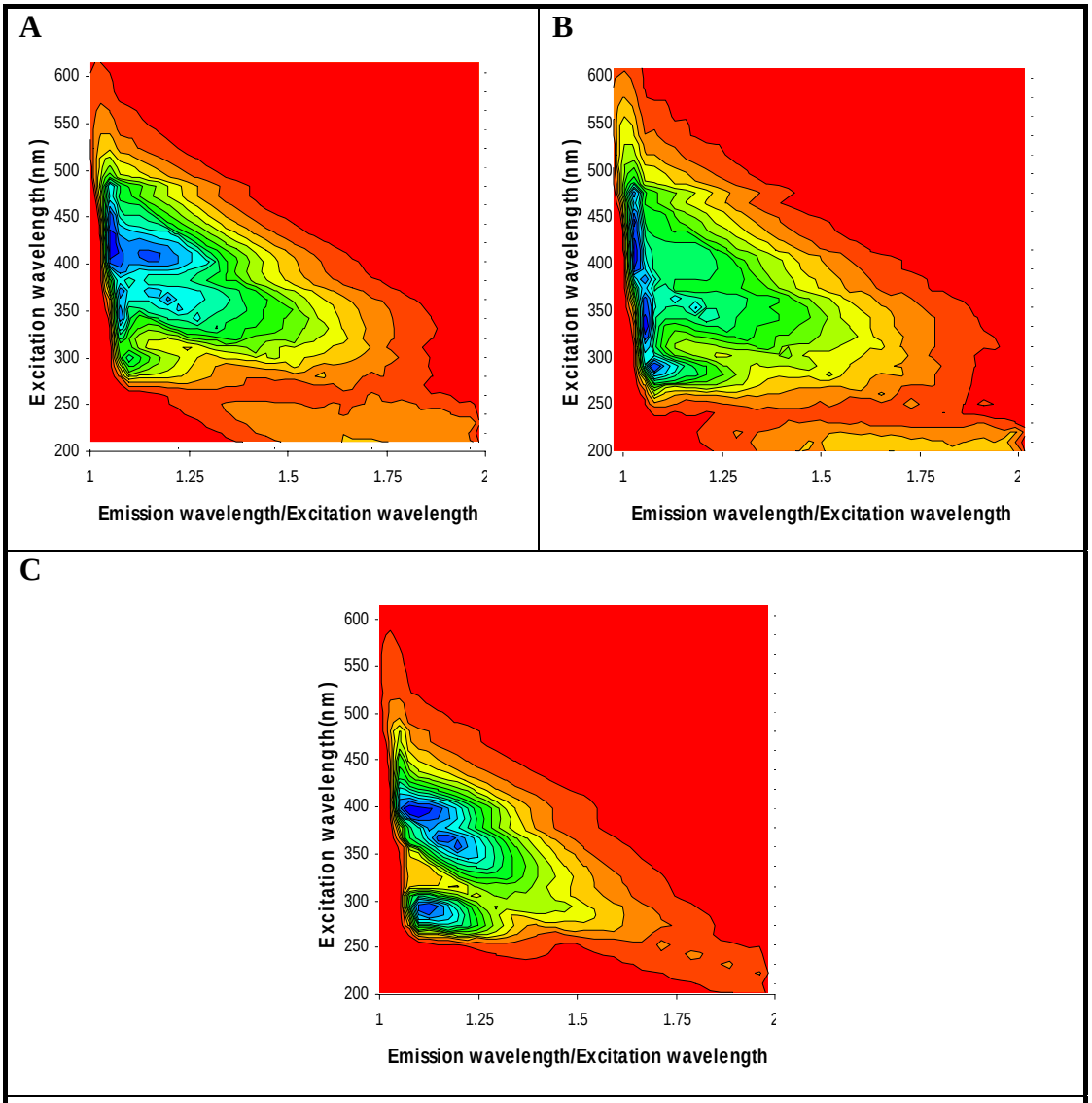


Figure 4.24 Fluorescence fingerprints of (A) *B. globigii* (ii); (B) corn smut and (C) ovalbumin using excitation wavelengths from 200-600 nm is displayed. The fluorescence signal is divided into 15 equally spaced contour lines. Blue represents the maximum intensity.

4.3.2.2 PCA and Cluster Analysis

In this section, we present PCA analysis procedures and results in detail. To visualise and assess the discrimination potentials of fluorescence spectroscopy, PCA was applied to the spectrum of the data set which provides a lower dimensional representation of a data set in a coordinate system that preserves the maximum amount of data set's variance [78]. PCA and PAM cluster classification were carried out using R software package (GNU project, Version 2, Boston, MA, http://www.gnu.org).

PCA for the discrimination of spores from different dust background, the same fluorescence data that we have measured (on tape substrate) for reproducibility study, from three dusts background A, B and C and fluorescence data from nine variables (three samples of *B. globigii* (i): we called as *B. globigii*(i) -1, *B. globigii* (i) -2, *B. globigii* (i) -3; three samples of *B. cereus*: called as *B. cereus* -1, *B. cereus* -2, *B. cereus* -3; three samples of *B. popilliae*: called as *B. popilliae* -1*, B. popilliae* -2 and *B. popilliae -3*) were used.

Also we used the fluorescence data from the outside made source, the *B. globigii* (ii). To check the validity of PCA analysis in discriminating spore forming samples from non spore samples, we used the fluorescence data from the measurements of corn smut and ovalbumin.

As explained in the introduction to data analysis, the first step in PCA analysis is make a data matrix for the fluorescence data obtained from the samples under study. We wrote a computer program in R language to read the data files from above mentioned fluorescence data (see the R program in Appendix 7.5). The program will make a data matrix from the fluorescence data according to the wavelength selection. In the following section we will only show the procedure for PCA analysis for the *Bacillus* samples with the added dusts.

We decided to limit the data matrix to five different excitation wavelengths at 280, 310, 340, 370 and 400 nm. We selected these wavelengths because these show reasonable fluorescence signals from bacteria and bacterial spores and wanted to arbitrarily limit the excitation to five wavelengths. Since the measured fluorescence signals contains noise mainly due to scattering from background, we modified the program to select data points from the emission range 330-430, 360-

460, 390- 490, 420-520 and 450-550 nm for the selected excitation wavelengths(100 nm wavelength range). The fluorometer (F-4500) we used measured emission in every 0.2 nm. Therefore in this selected wavelength range we have a total of 500 data point for each excitation wavelength for each spectrum. The computer program was designed in such way that it should concotanate the data points from the selected excitation wavelength and makes a data matrix for the fluorescence data. In this case, since we have nine different samples and co-adding five excitation spectral data points for each sample, the data matrix for nine samples (with added dust) will contain 9x2500 data points.

The next step in PCA analysis was to make a correlation/covariance matrix of the data matrix that we made for the nine variables/samples. PCA can be conducted with either of this matrix, the choice depending on the data being used. The correlation matrix is used in preference to the covariance matrix when the variables have very different absolute values. In choosing to use the correlation matrix over the covariance matrix, the variability of a variable can be exaggerated. However, by taking correlations, the degree of variation from the mean of the variable is considered, not the absolute variation (represented by the standard deviation). The effects of this and other standardizing technique are discussed by Cao et al [149].

In this study, the correlation matrix (see for example Table 4-12) is used. The entries in the correlation matrix (C) are the correlation between each of the variables (linear Pearson's correlations [150] are used here) i.e.

$$C = \begin{pmatrix} 1 & r_{12} & \ldots.. & r_{1p} \\ r_{21} & 1 & \ldots.. & r_{2p} \\ . & & . & \\ . & & . & \\ r_{p1} & r_{p2} & \ldots. & 1 \end{pmatrix} \qquad (4.2)$$

where $r_{ij} = c_{ij} / s_i s_j$

c_{ij} = covariance between variables i and j.

s_i, s_j = standard deviation of the variables i and j.

Note that if the covariance matrix is used the entries are the covariance between the variables with the variances along the diagonals.

In Tables 4-12, 4-13 and 4-14, the PCA output from R package for analysis for nine variables (samples) described above is shown. The correlation matrix (Table 4-12) for the PCA indicates the degree of dependency between the variables. A high correlation, close to +1 or -1, indicates a high degree of interdependency between the variables. As described above, the PCs are independent linear combination of variables. The PC scores (a_{ij} in Table 4-13) give the constants for each of the variables forming the PCs. The PCs scores (a_{ij}) are derived from eigenvectors of the correlation matrix and are scaled such that $(a_{i1})^2 + (a_{i2})^2 + \ldots\ldots + (a_{ij})^2 = 1$ for all i.

The variance of any PC is given by the associated eigenvalue of the correlation matrix as shown in the Table 4-14, for example, the variance of first principal component PC1 is 5.342 for PCA analysis that described above. This means, among nine variables/samples PC1 gives the largest variance due to the *B. globigii* (i) samples along PC1 axis. The PCs are ordered so that PC1 explains as large a proportion of overall variance as possible with successive PCs explaining decreasing amount of variance while remaining independent of previous PCs.

Table 4-12 Correlation matrix for the PCA analysis of three samples each of *B. globigii* (i), *B. cereus* and *B. popilliae* from dust A, B and C backgrounds.

Variables	***B. globigii*(i)-1**	***B. globigii*(i)-2**	***B. globigii*(i)-3**	***B. cereus*-1**	***B. cereus*-2**	***B. cereus* -3**	***B. popilliae*-1**	***B. popilliae* -2**	***B. popilliae* -3**
***B. globigii* (i)-1**	1.00								
***B. globigii* (i)-2**	0.99	1.00							
***B. globigii* (i)-3**	0.99	0.99	1.00						
***B. cereus*-1**	0.07	0.01	-0.02	1.00					
***B. cereus*-2**	0.08	0.02	-0.01	0.99	1.00				
***B. cereus*-3**	-0.01	-0.04	-0.07	0.68	0.73	1.00			
***B. popilliae*-1**	-0.49	-0.56	-0.56	0.63	0.63	0.60	1.00		
***B. popilliae*-2**	-0.34	-0.38	-0.42	0.88	0.09	0.69	0.86	1.00	
***B. popilliae*-3**	-0.35	-0.37	-0.42	0.87	0.88	0.69	0.80	0.99	1.00

Table 4-13 Principal Components Scores (a_{ij})

Variables	**PC1**	**PC2**	**PC3**	**PC4**	**PC5**	**PC6**	**PC7**	**PC8**	**PC9**
***B. globigii*(i)-1**	0.21	-0.51	0.01	0.21	-0.15	0.02	0.51	-0.02	0.61
***B. globigii*(i)-2**	0.22	-0.50	-0.01	0.05	0.38	0.10	-0.71	-0.11	0.18
***B. globigii*(i)-3**	0.24	-0.49	-0.00	0.19	0.12	-0.05	0.27	0.15	-0.74
***B. cereus*-1**	-0.35	-0.30	0.32	-0.13	-0.50	0.57	-0.16	0.24	-0.08
***B. cereus*-2**	-0.36	-0.31	0.21	-0.15	-0.29	-0.76	-0.14	-0.18	-0.03
***B. cereus*-3**	-0.31	-0.23	-0.88	-0.28	-0.06	0.07	0.03	-0.01	-0.02
***B. popilliae*-1**	-0.49	0.07	-0.17	0.87	-0.05	-0.06	-0.16	0.16	0.02
***B. popilliae*-2**	-0.42	-0.08	0.17	0.07	0.34	0.27	0.23	-0.73	-0.09
***B. popilliae*-3**	-0.42	-0.08	0.18	-0.19	0.60	-0.08	0.18	0.57	0.16

Table 4-14 Eigenvalue

Principal Components	**PC1**	**PC2**	**PC3**	**PC4**	**PC5**	**PC6**	**PC7**	**PC8**	**PC9**
Eigenvalue	5.342	2.943	0.418	0.218	0.063	0.011	0.003	0.002	0.001

Note: 1

Variance (PC1) = 5.342 or 59.4 %

Variance (PC2) = 2.943 or 32.7 %

Combined PC1 and PC2 explain 92.1 % of the overall variance and will satisfy the criteria [136] explained in the text. Similar PCA analyses were done for the other samples without added dusts, studied in this section (data not shown) to check the stability of cluster analysis using PCA as preprocessing techniques.

Note: 2

As explained earlier, we can express PC1 as the linear combination of original variables (*B. globigii* (i)-1, *B. globigii* (i)-2, …, *B. popilliae*-3) by using the PCs scores from Table 4-14 as follows

PC1 = 0.21 x *B. globigii* (i)-1 + 0.22 x *B. globigii* (i)-2 +…..+ ……. +… + (-0.42 x *B. popilliae*-3)

Similarly,

PC2 = (-0.51) x *B. globigii* (i)-1 + (-0.50) x *B. globigii* (i)-2 + ……. + …. + (-0.08 x *B. popilliae*-3)

where the coefficients for the original variables are the loading or scores factors.

Also, note from the correlation matrix Table 4-12 that similar samples have almost the same correlation coefficients. Here, for example, the *B. globigii* (i) samples possess the highest correlation coefficients and *B. cereus*, *B. popilliae* are negatively correlated with the *B. globigii* (i) samples and their contributions to PC1

are very negligible. Therefore we can assume that most variance along PC1 was due to *B. globigii* (i) samples. PC2 give the second highest variance among the samples and as shown above, the contribution due to these components can be compared. As seen in the Table 4-13, the loading factors for *B. globigii* (i) and *B. cereus* actually determine the contribution of variance along PC2.

One of the nice features of PCA is that nearly every result can be represented graphically. The score along PCs reveal the relationship among the samples. This information is displaced in a score plot, also called eigenvector projection [151]. Similar samples group together in clusters (classes) in the score plot. Also, since the PCs are orthogonal, the Euclidian distance between samples can be used to measure similarity quantitatively. Interpretation of a single PC and the features of PC model is possible through the connection to the original variables.

A loading plot displays both the importance of each variable to the interpretation of a PC and relationship among variables in that PC. The coordinates of a variable in a loading plot are its loadings along the orthogonal and normalised PCs. First, a variable contribution to a PC is directly proportional to the squared loading. Thus, the distance of a variable to the origin along a PC is a quantitative measure of the importance of that variable in the PC. A variable near the origin carries little or no information in the PC, while a large distance from the origin (high loading) means that the variable is important in the interpretation of PC. Second, the mutual location of the variable reflects the coherences among them. Variables grouped together hold the same information in the PC. Also, variables located on the same side of the origin are positively correlated, while variables located on opposite sides are negatively correlated. The larger the separation of two negatively correlated variables, the stronger is the negative correlation. These statements can be generalised to the interpretation of the PC model.

In Figure 4.25 we show the scree plot for choosing the number of PCs as explained in the text. From the scree plot it was evident that PC1 and PC2 (two PCs) is sufficient to explain the overall variance and will support the criteria[136] cited in the text.

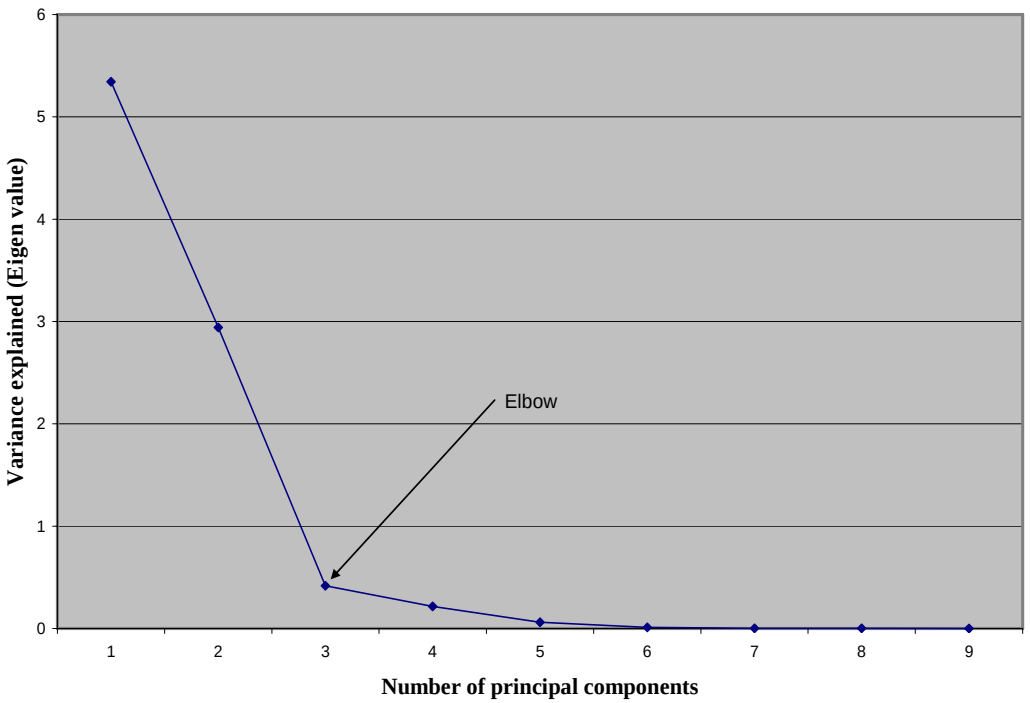


Figure 4.25 Scree plot for PCA analysis as explained in the text. The first two PCs would be analysed using this method

In Figure 4.26(A) [152] we show the two-dimensional PCA score plot (PC1 versus PC2) of three samples of *B. globigii* (i), *B. cereus*, and *B. popilliae* for variable loadings. The PCs score plots give the relative positions and grouping of the variables (samples). There appear to be three grouping of the variables. We also considered smoothing the data. We collected the fluorescence signal every 0.2 nm. The monochromator slits are set at 5 nm. We used a 20-node spline to smooth the 500-point emission spectrum from each excitation wavelength, as this would give smoothing that approximately equalled the bandwidth of the monochromators. The smoothing created almost no change to the PCA.

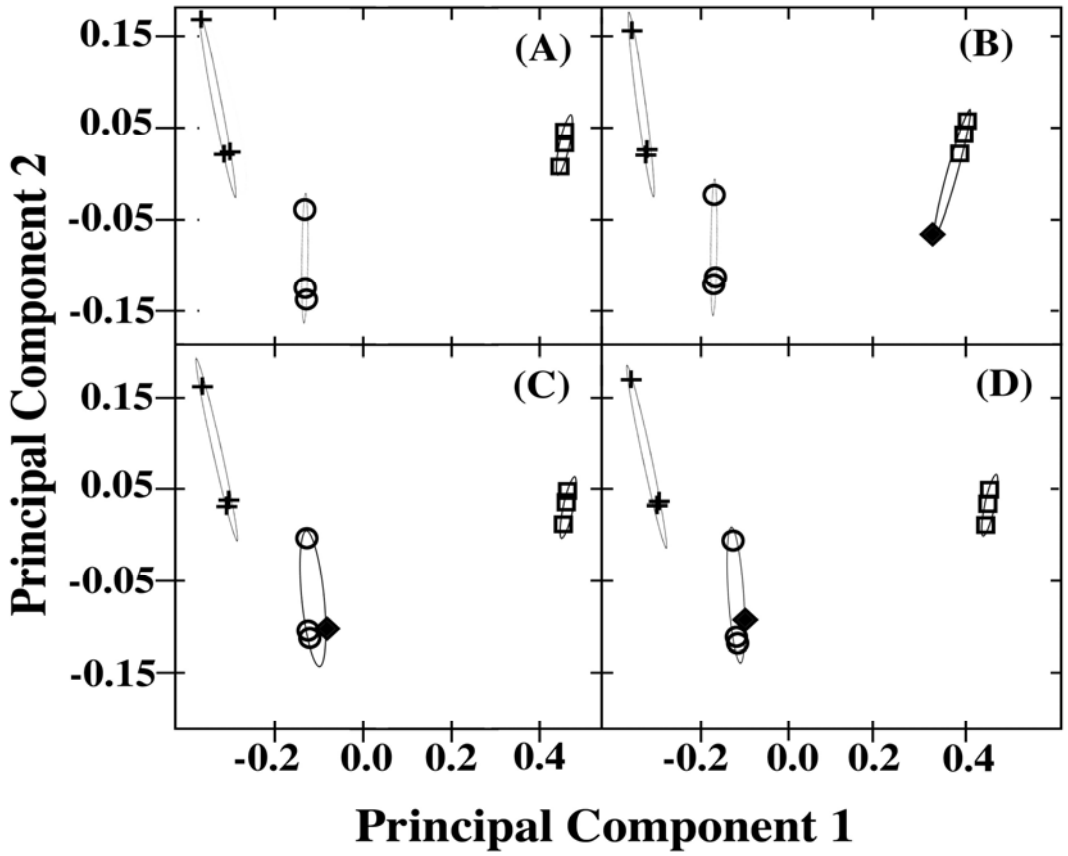


Figure 4.26 (A) PCA plot of three samples of *B. globigii* (i) (open squares), *B. cereus* (open circles) *and B. popilliae* (crosses) (9 measurements). Each sample is mixed with dust as explained in the text. PC1 versus PC2 is plotted. Plots of the PCA analysis of same three samples of *B. globigii* (i), *B. cereus*, *B. popilliae* (9 measurements) with the added sample of (B) *B. globigii* (ii) (solid diamond), (C) corn smut (solid diamond) and (D) ovalbumin (solid diamond) [152].

In Figure 4.26(B) we show the plot from the PCA analysis of the same 9 measurements of *B. globigii* (i), *B. cereus*, and *B. popilliae* with the *B. globigii* (ii) prepared from an outside source. We show the plot from PCA analysis of the same 9 measurements with corn smut added in (C), and ovalbumin added in (D). For plotting

the PCs we use the first two PCs determined from the excitation wavelengths of 280, 310, 340, 370 and 400 nm.

Ideally, we would like to employ an analysis that would group or cluster the *B. globigii* (ii) with our other *B. globigii* (i) measurements and have the corn smut and ovalbumin not cluster with any of our three bacterial spore measurements. The silhouettes of the clusters shown in Figure 4.26 are given in Figure 4.27 [152]. Each shaded bar is associated with an object in Figure. 4.26. The corresponding symbols are given. The number at the end of each cluster is the average S_i for that cluster. In Figure 4.27(B), the *B. globigii* (ii) clusters well with the other *B. globigii* (i) samples. In Figure 4.27(C) and (D), the ovalbumin and corn smut cluster with the *B. cereus*. This is an incorrect clustering.

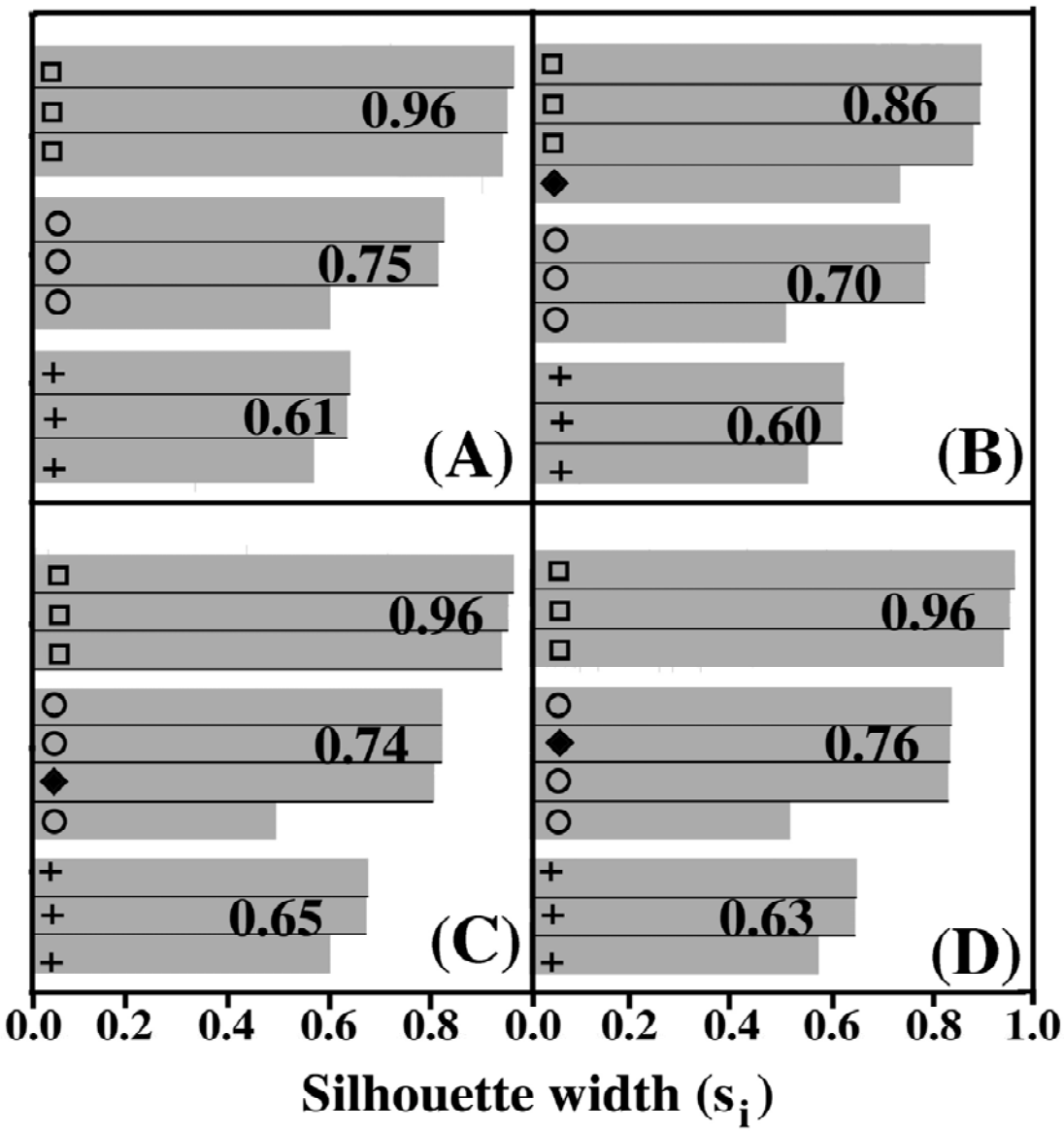


Figure 4.27 Silhouette plot of the strength of clustering using PAM with three samples of *B. globigii* (i) (open squares), *B. cereus* (open circles), *B. popilliae* (crosses) (9 measurements) as shown in Figure 4-26. Silhouette plot of the strength of clustering using PAM with the same three samples of *B. globigii* (i), *B. cereus*, *B. popilliae* with added sample of (B) *B. globigii* (ii) (solid diamond), (C) corn smut (solid diamond) and (D) ovalbumin (solid diamond) [152].

4.3.3 Detection and Identification of *Bacillus* Spores from Background Pollen like Pig weed

4.3.3.1 Spore Fluorescence

We show in Figure 4.28(A-D) the fluorescence fingerprints that are measured for *B. cereus*, *B. globigii* (iii), pig weed and a mixture of pig weed and *B. globigii* (iii). Fluorescence for the samples are measured using excitation wavelengths from 200-600 nm on a quartz substrate.

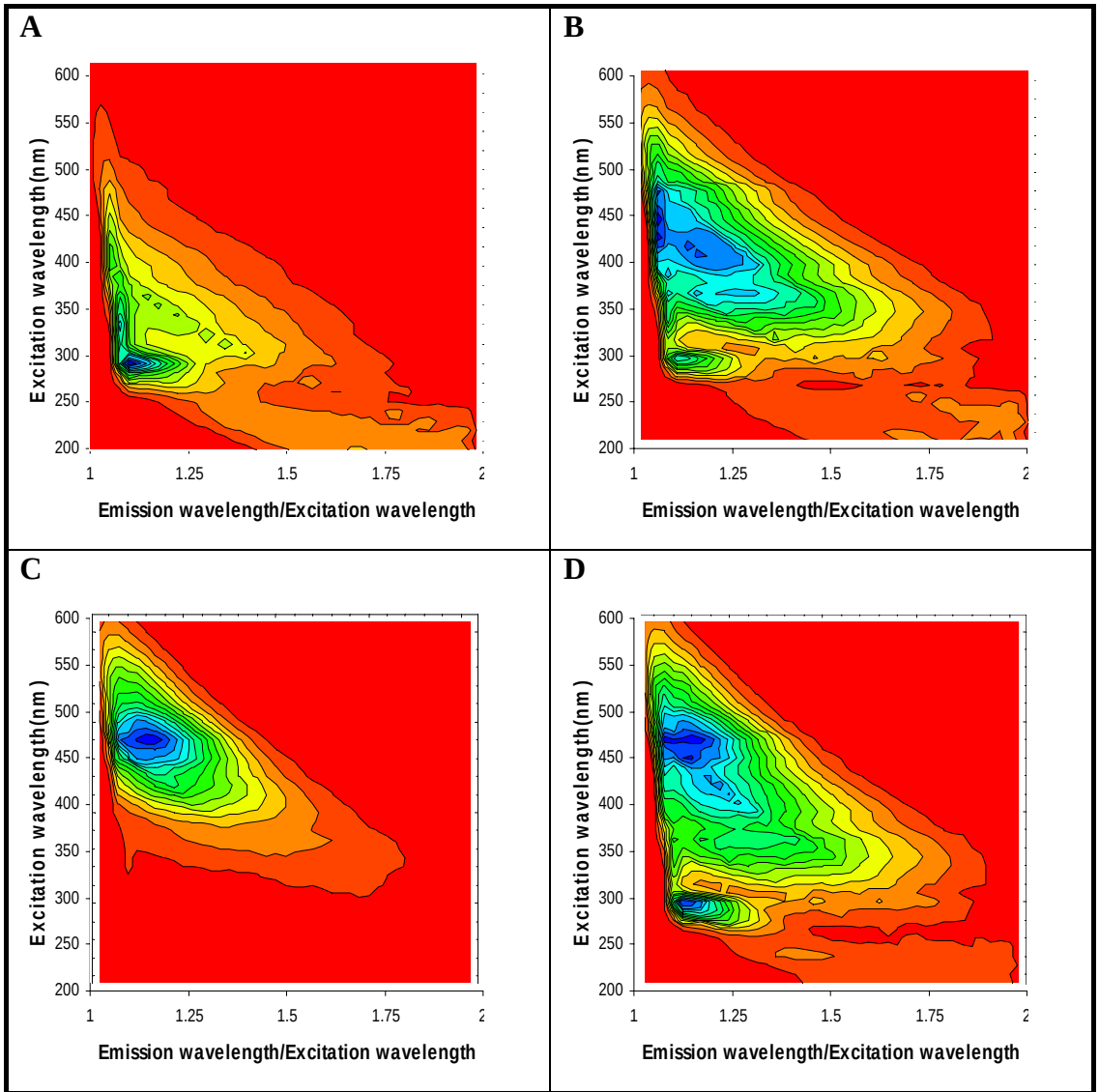


Figure 4.28 Fluorescence fingerprints of (A) *B. cereus*, (B) *B. globigii* (iii), (C) pig weed and (D) mixed samples of pig weed and *B. globigii* (iii) using excitation wavelengths from 200-600 nm is displayed. The fluorescence signal is divided into 15 equally spaced contour lines. Blue represents the maximum intensity.

4.3.3.2 PCA and Cluster Analysis

This is an extension of the previous analysis for detection and identification *Bacillus* simulants. Here, we were interested to look at the strength of PCA in discriminating *Bacillus* spores from a mixture of pollen as well as other *Bacillus* species. The main differences in this case were the selection of variables/samples and excitation wavelength used.

The data matrix for this analysis was taken from the fluorescence data (measured on a quartz substrate) of three measurements each of *B. cereus*, *B. globigii* (iii), pig weed (pollen grain) and mixed samples of pig weed and *B. globigii* (iii). We have 12 samples/variables for PCA. We noticed that the pig weed sample that we used was producing very high fluorescence intensities in the emission range 500-600 nm with excitation ranges 400-500 nm (see Figure 4.28A) and therefore can be discriminate the simulant from pollen. We followed the same procedure for PCA analysis for this section.

As before, we decided to limit the data matrix to five different excitation wavelengths at 280, 320, 360, 400 and 440 nm. We selected these wavelengths because these show reasonable fluorescence signals from bacterial spores and wanted to arbitrarily limit the excitation to five wavelengths as we done earlier. We selected data points from the emission range 310-410, 350-450, 390- 490,430-530 and 470-570 nm for the selected excitation wavelengths (100 nm emission data length). The fluorometer (SLM 8000C) we used was measured emission in every 1 nm. Therefore in this selected wavelength, as explained in the previous case, we have a total of 500 data point for each excitation wavelength for each spectrum and the entire data matrix consisting of 12 samples will have 12x500 data points.

PCA analysis was done as explained in the previous section. The correlation matrix for this analysis is given in the Table 4-15. *B. cereus*-1, *B. cereus*-2, *B. cereus*-3; pig weed-1, pig weed-2, pig weed-3; *B. globigii* (iii)-1, *B. globigii* (iii)-1, *B. globigii* (iii)-1; pig weed-1 and *B. globigii* (iii)-1, pig weed-2 and *B. globigii* (iii)-2, pig weed-3 and *B. globigii* (iii)-3 represent the three different measurement of

each of *B. cereus* spores, pig weed, *B. globigii* (iii) spores and a mixture of pig weed with *B. globigii* (iii) respectively. In Table 4-15 the eigenvalue is given.

Table 4-15 Correlation matrix for PCA analysis of *Bacillus* spores from pollen background

Variables	***B. cereus*-1**	***B. cereus*-2**	***B. cereus*-3**	**Pig weed-1**	**Pig weed-2**	**Pig weed-3**	***B.globigii*(iii)-1**	***B.globigii*(iii)-2**	***B.globigii*(iii)-3**	**Pig weed-1 and *B.globigii*(iii)-1**	**Pig weed-2 and *B.globigii*(iii)-2**	**Pig weed-3 and *B.globigii*(iii)-3**
***B.cereus*-1**	1.00											
***B.cereus*-2**	0.98	1.00										
***B.cereus*-3**	0.88	0.94	1.00									
Pig weed-1	-0.61	-0.68	-0.76	1.00								
Pig weed-2	-0.60	-0.69	-0.77	0.99	1.00							
Pig weed-3	-0.59	-0.67	-0.74	0.99	0.99	1.00						
B.globigii* (iii)-*1	0.30	0.31	0.38	0.21	0.17	0.24	1.00					
B.globigii* (iii)-*2	0.56	0.57	0.61	-0.05	-0.09	-0.02	0.95	1.00				
B.globigii* (iii)-*3	0.37	0.37	0.42	0.18	0.14	0.20	0.99	0.97	1.00			
Pig weed-1 and *B.globigii* (iii)-1	0.02	0.03	0.01	0.49	0.46	0.50	0.77	0.66	0.77	1.00		
Pig weed-1 and *B.globigii* (iii)-2	-0.02	-0.04	-0.15	0.62	0.60	0.63	0.64	0.52	0.65	0.95	1.00	
Pig weed-1 and *B.globigii* (iii)-3	-0.17	-0.18	-0.23	0.71	0.68	0.71	0.66	0.50	0.65	0.96	0.97	1.00

Table 4-16 Eigenvalue

PC	**PC1**	**PC2**	**PC3**	**PC4**	**PC5**	**PC6**	**PC7**	**PC8**	**PC9**	**PC10**	**PC11**	**PC12**
Eva	5.99	5.01	0.54	0.39	0.05	0.004	0.01	0.004	0.002	0.001	0.001	0.001

Note: Variance (PC1) = 5.99 or 49.9 %;
Variance (PC2) = 5.01 or 41.75 %
Variance (PC3) = 0.536 or 4.46 %

Combined PC1 and PC2 explain 91.65 % of the overall variance and again will satisfy the criteria[136] explained in the text and therefore we chose the first two PC for plotting the score plots.

In Figure 4.29 we show the two-dimensional PCA plot (PC1 versus PC2) of the four samples of *B. globigii* (iii), *B. cereus*, pig weed and *B. globigii* (iii) mixed with pig weed, with each three different preparations. A total of 12 feature points can be seen with the corresponding symbols used. We collected the fluorescence signal every 1 nm. The monochromator slits are set at 4 nm. We used a 20-node spline to smooth the 100-point emission spectrum from each excitation wavelength, as this would give smoothing that approximately equalled the bandwidth of the monochromators. Each sample is measured as a dry spore on quartz slide as explained in the text.

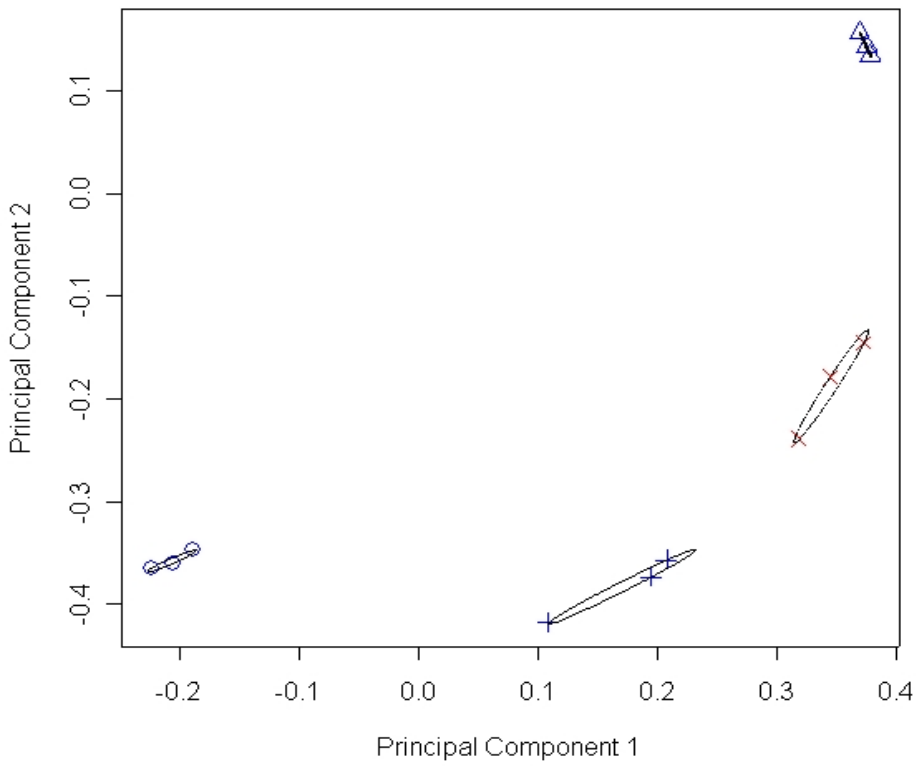


Figure 4.29 PCA plot of four samples of *B. cereus* (open circles), pig weed (triangles), *B. globigii* (iii) (plus) and mixture of *B. globigii* (iii) and pig weed (cross) (12 measurements). PCA was done with five excitation wavelengths 280, 320, 360, 400 and 440 nm and the corresponding emission wavelengths used from 310-410, 350-450, 390-490, 430-530 and 470-570 respectively. PC1 versus PC2 is plotted.

We also considered using three PCs in plotting the score plot for this analysis. As seen from Table 4-16, the third components produce a variation of 4.46 % to the overall variation of data sets. In Figure 4.30 we show three components plot of PCA analysis with PC1, PC2 and PC3. It is interesting to notice that the effect of addition of the third components to draw PCA plots hardly yields any new information of data set except for the fact that it simply adds more noise data.

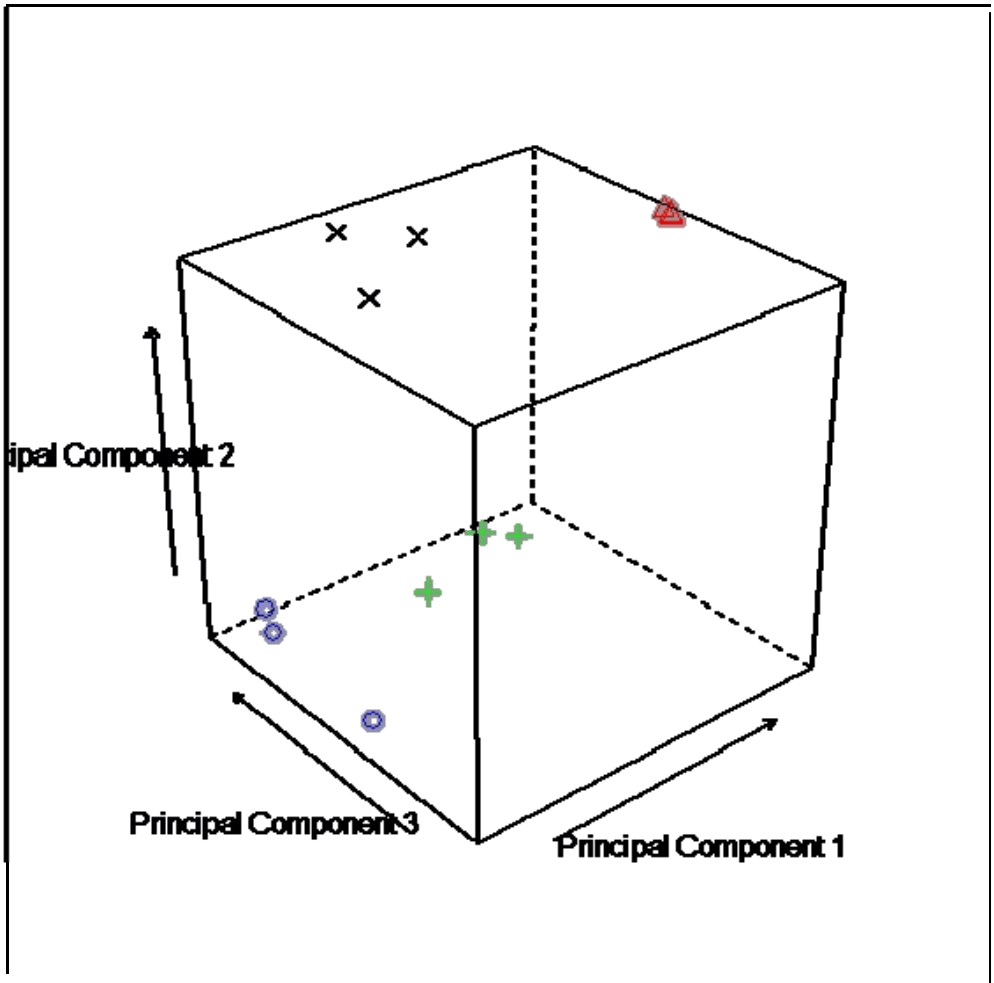


Figure 4.30 PCs 1, 2 and 3 for the 12 fluorescence measurements made with four different samples each with three different preparations. The circles' are *B. cereus*. The triangles' are pig weed. The pluses are the *B. globigii* (iii) and the crosses' are mixed sample of *B. globigii* (iii) and pig weed.

In Figure 4.31 we show the hierarchical and silhouette plots of results shown in Figure 4.29 as the similarity relationships between the 2D fluorescence spectra of samples, prepared by using unweighted pair group method with arithmetic average. In cluster dendrogram and PAM plots, the numbers 1, 2, 3 …. etc will represent the samples represented by the corresponding symbols in the scatter plot. Each shaded bar on the silhouette is associated with an object in Figure 4.29. The corresponding symbols are given. The number at the end of each cluster is the average S_i for that cluster.

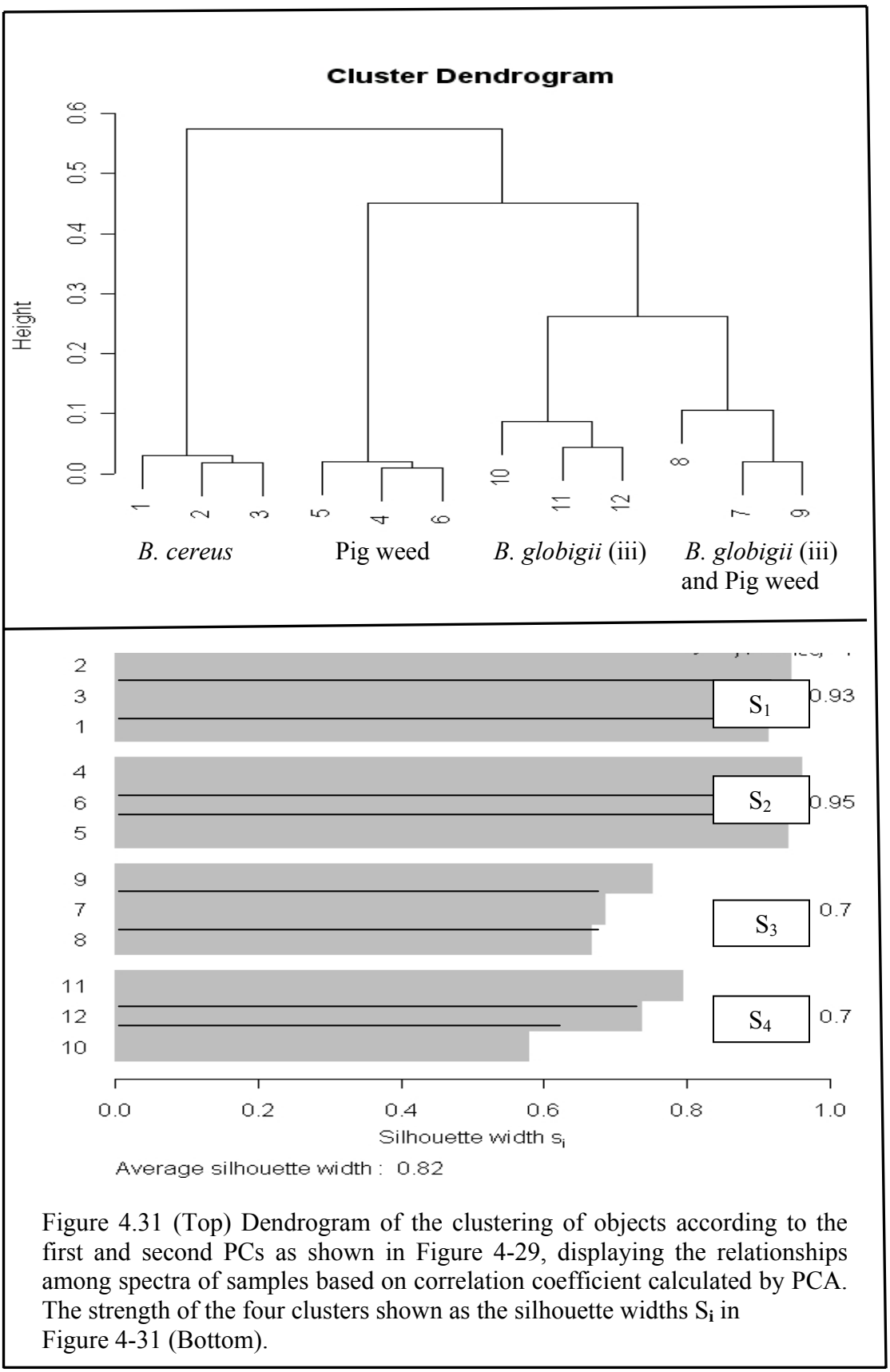


Figure 4.31 (Top) Dendrogram of the clustering of objects according to the first and second PCs as shown in Figure 4-29, displaying the relationships among spectra of samples based on correlation coefficient calculated by PCA. The strength of the four clusters shown as the silhouette widths S_i in Figure 4-31 (Bottom).

4.3.4 Effects of Washing on Identification of *Bacillus* Spores

In this section we will show the results obtained in this research for the study of the effects of washing on spores fluorescence and PCA application for the discrimination of *Bacillus* spores from non *Bacillus* simulants. In sections 4.2.4.1 and 4.2.4.2 we show the results obtained using the application of 2D fluorescence fingerprints and PCA analysis of the data obtained.

4.3.4.1 Spore Fluorescence

In Figure 4.32 we show a series of fluorescence fingerprints for the non *B. globigii* samples investigated on a quartz substrate. The three separate measurements of each sample are similar to the plots shown. Figure 4.32(A-C) is the fluorescence fingerprints of *B. cereus*, *B. popilliae* and *B. thuringiensis* measured on quartz substrate using excitation wavelengths from 200-600 nm. Details of the measurements are given in the text.

We checked for changes in the fluorescence spectrum before and after washing the spores. In Figure 4.33(A-F) we show the fingerprints for the fluorescence measured from the three different preparations of *B. globigii* (i, ii, and iii) before and after being washed twice. The excitation is from 200-600 nm. Details of the measurements are given in the text.

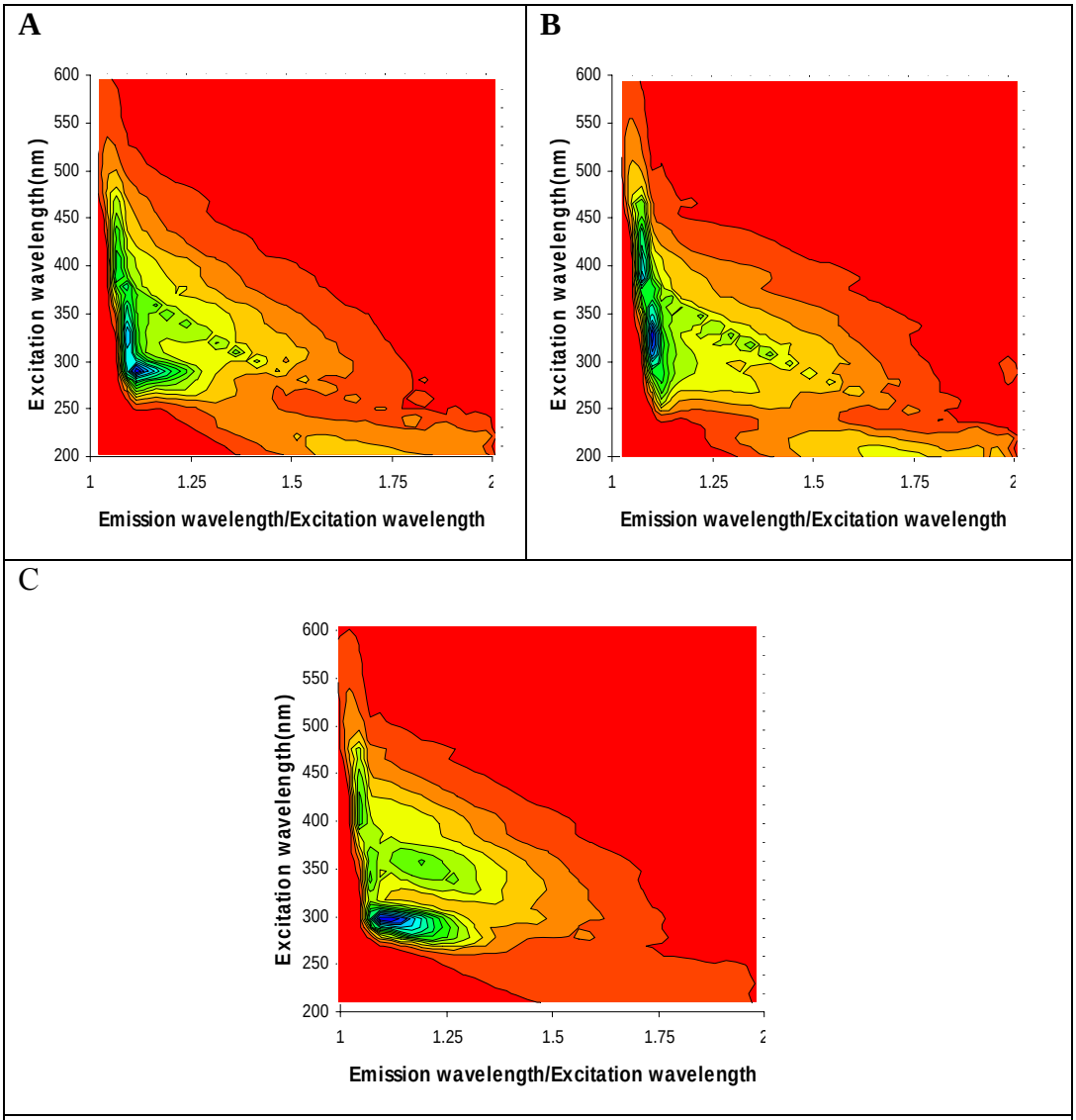


Figure 4.32 Fluorescence fingerprints of (A) *B. cereus*; (B) *B. popilliae* and (C) *B. thuringiensis* (all unwashed) using excitation wavelengths from 200-600 nm are displayed. The fluorescence signal is divided into 15 equally spaced contour lines. Blue represents the maximum intensity.

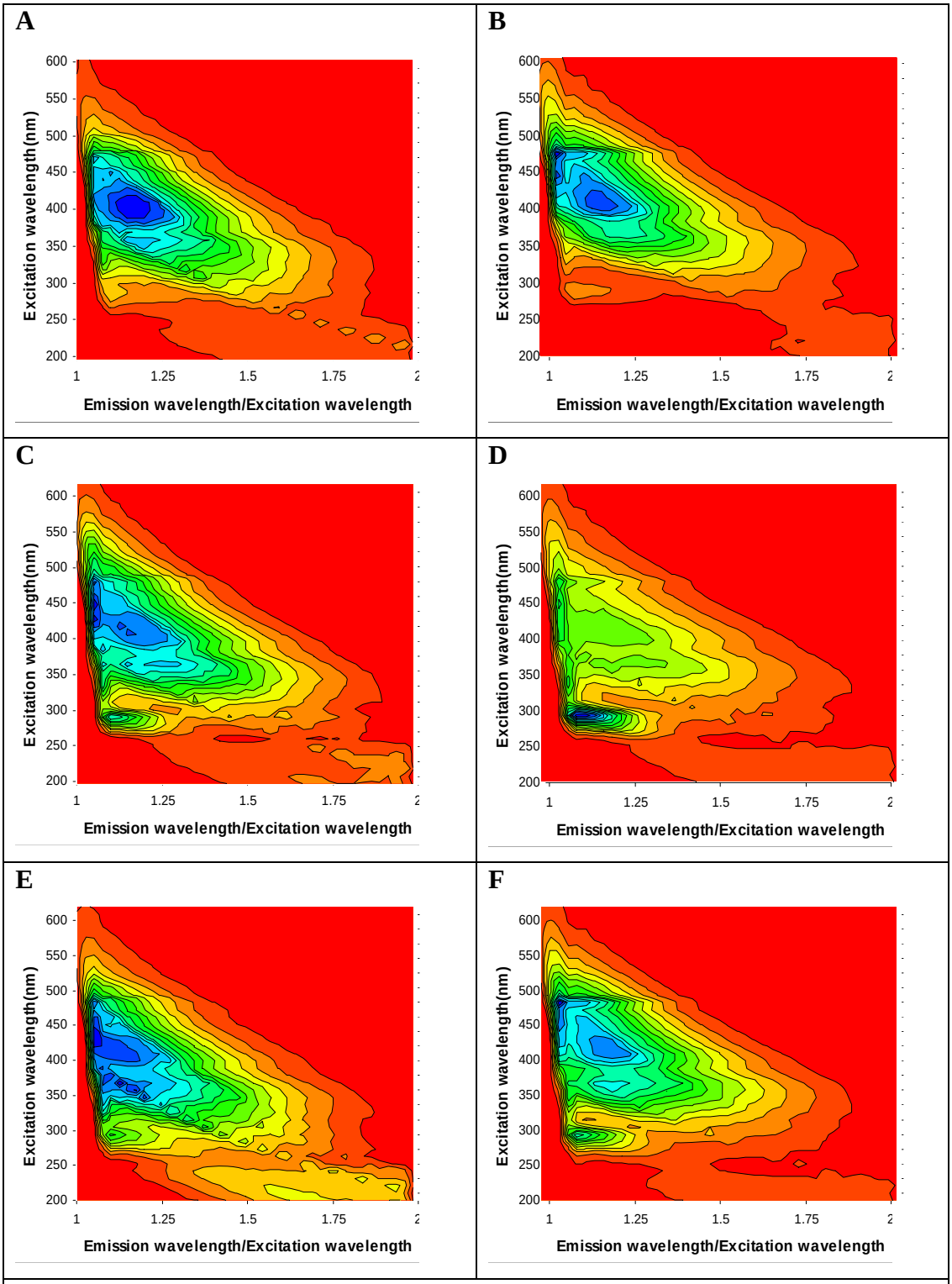


Figure 4.33 Fluorescence fingerprints from dry spores of (A) *B. globigii* (i) before washing, (B) *B. globigii* (i) after washing, (C) *B. globigii* (ii) before washing, (D) *B. globigii* (ii) after washing, (E) *B. globigii* (iii) before washing, and (F) *B. globigii* (iii) after washing. The excitation wavelengths from 200-600 nm were used.

4.3.4.2 PCA and Cluster Analysis

For the study of the effects of washing on spore's fluorescence, PCA analysis was done again as explained in the previous sections. Fluorescence data for this analysis were obtained from the measurements of 27 samples as explained in the experimental section. A data matrix for PCA analysis was made from the fluorescence data. Procedures were the same as before and the difference here was also for the selection of excitation wavelength. As we have done in the previous analyses, five spectra each from each sample were selected at 280, 310, 340, 370 and 400 nm excitation wavelength. Since the measured fluorescence signals contains noise mainly due to scattering from background we selected data points from the emission range 310-410, 340-440, 370-470, 400-500 and 430-530 nm for the selected excitation wavelengths (100 nm emission spectral length). As mentioned before the fluorometer we used was measured emission in every 1 nm. Therefore in this selected wavelength as explained in the previous case, we have a total of 500 data points for each excitation wavelength for each spectrum and thereby for the entire data matrix of 27 samples will have 27x500 data points. Here, we only show the eigen vector calculated for the analysis. In Table 4-17 the eigenvalue associated with eigenvectors for 27 variables is given.

Table 4-17 Eigenvalue

PC	**PC1**	**PC2**	**PC3**	**PC4**	**PC5**	**PC6**	**PC7**	**PC8**	**PC9**
Eva	14.64	8.31	2.81	0.683	0.273	0.0978	0.0765	0.0358	0.0188
PC	**PC10**	**PC11**	**PC12**	**PC13**	**PC14**	**PC15**	**PC16**	**PC17**	**PC18**
Eva	0.0148	0.0094	0.0056	0.0052	0.0034	0.0029	0.0014	0.0013	0.0012
PC	**PC19**	**PC20**	**PC21**	**PC22**	**PC23**	**PC24**	**PC25**	**PC26**	**PC27**
Eva	0.0011	0.0010	0.00095	0.00074	0.00071	0.00058	0.00053	0.00048	0.00043

Note: Variance (PC1) = 14.64 or 54.22 %;
Variance (PC2) = 8.31 or 30.77 %
Variance (PC3) = 2.81 or 10.40 %

Combined PC1, PC2 and PC3 explain 95.39 % of the overall variance and will satisfy the criteria [136] as explained in the text.

In Figure 4.34 we show the scree plot for finding number of PC to retain. It is interesting to note that in this analysis of effect of washing on identification, it is worth using three PC rather than two as required by the criteria [136] to use the scree plot method to retain the number of PC.

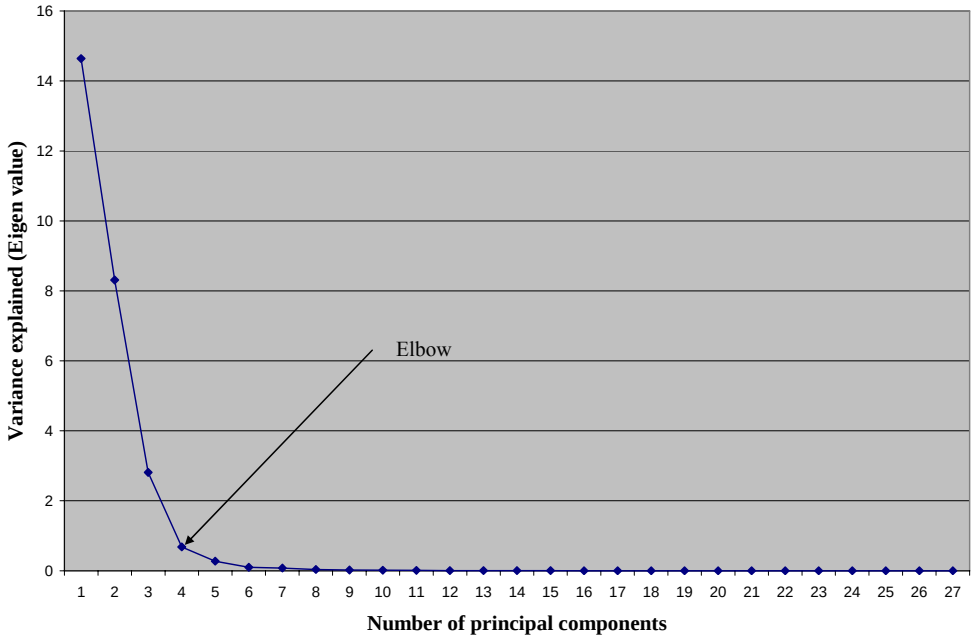


Figure 4.34 Scree plot for PCA analysis as explained in the text. The first three PCs would be analysed using this method

In Figure 4.35 we show the two-dimensional PCA plot (PC1 versus PC2) of six samples of *B. globigii* (i, ii, and iii) (washed and unwashed) and the samples of *B. cereus*, *B. popilliae,* and *B. thuringiensis.* Each sample was placed on three quartz slides and was, therefore, measured three times.

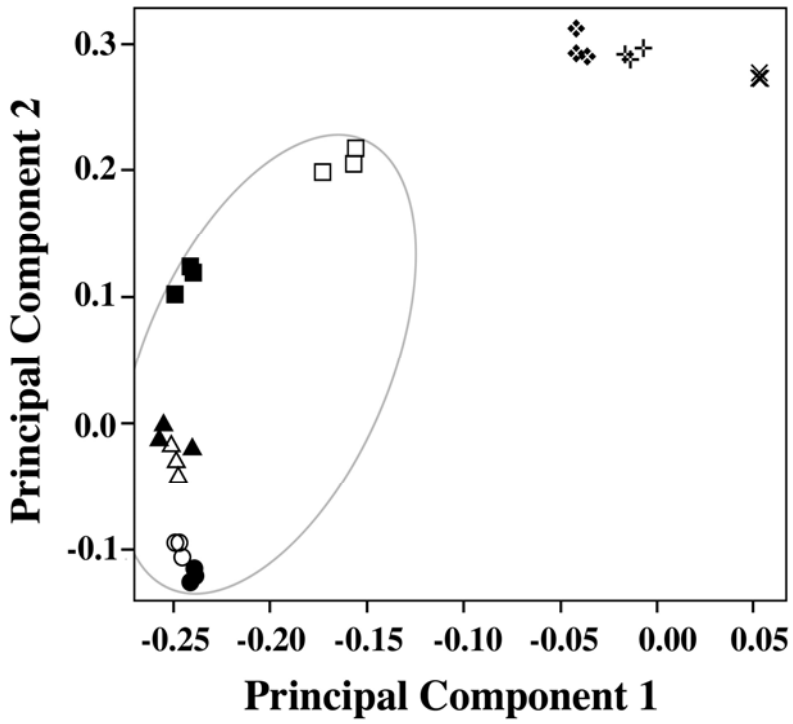


Figure 4.35 PC1 versus PC2 for the 27 fluorescence measurements made. The circles are *B. globigii* (i). The filled circles are before washing and the open circles are after washing. The triangles are *B. globigii* (ii). The filled squares are before washing and the open squares are after washing. The oval is drawn to demonstrate the clustering. The diamonds are the *B. cereus* samples, the x are the *B. popilliae* samples, and the + are the *B. thuringiensis* samples.

Two of the three preparations of *B. globigii* seem to show almost no change in PCs one and two when they are compared before and after washing. If we look at the first three PCs, as shown in Figure 4.36**,** we can see that significant changes along the third PC are seen before and after washing for these two preparations of the *B. globigii* spores. One can also see with the three PCs that the non *B. globigii* spores separate better. The use of only the first two PCs follows Kaiser's criteria [153].

Here, we can see that is might be advantageous to use additional PCs in the analysis as suggested by Yeung and Ruzzo [154].

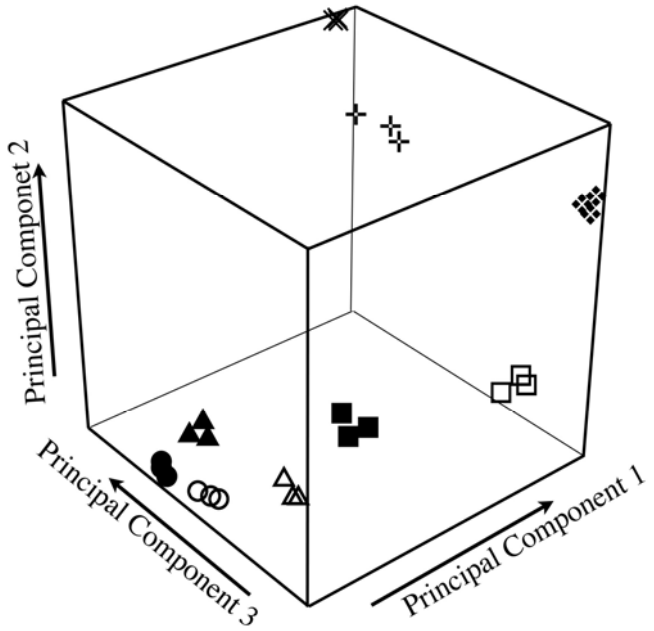


Figure 4.36 PCs 1, 2 and 3 for the 27 fluorescence measurements made. The circles are *B. globigii* (i). The filled circles are before washing and the open circles are after washing. The triangles are *B. globigii* (ii). The filled triangles are before washing and the open triangles are after washing. The squares are *B. globigii* (iii). The filled squares are before washing and the open squares are after washing. The diamonds are the *B. cereus* samples, the x are the *B. popilliae* samples, and the + are the *B. thuringiensis* samples.

It is evident from Figure 4.35 that the six *B. globigii* samples (circles, squares and triangles) clustered well as expected. There is a slight change with the *B. globigii* (iii) before and after washing (solid and open squares, respectively). We suspect that these spores contain more growth media and cell debris compared to other *B. globigii* samples. Still, all three *B. globigii* clearly discriminate with other *Bacillus* species: *B. cereus*, *B. popilliae* and the *B. thuringiensis* samples.

The cluster dendrogram and the silhouettes of the clusters are shown in Figure 4.37. In Figure 4.37(A) we show the hierarchical plots for the entire twenty-seven samples that were used for PCA. There are three clusters from the dendrogram shown in Figure 4.37(A). The cluster groups are shown with the brackets at the bottom of the figure. In Figure 4.37(B), we show the cluster strengths (as defined in Eq. 2.8). There are a total of 27 objects in the clustering scheme. The average silhouette width is 0.72. The top cluster contains 12 objects with an average silhouette width of 0.75. The middle cluster contains 6 objects with an average silhouette width of 0.55. The bottom cluster contains 9 objects with an average silhouette width of 0.80.

The clusters are all reasonable with respect to the strength of the cluster. It is important to note that all of the samples stayed with their appropriate clusters before and after washing.

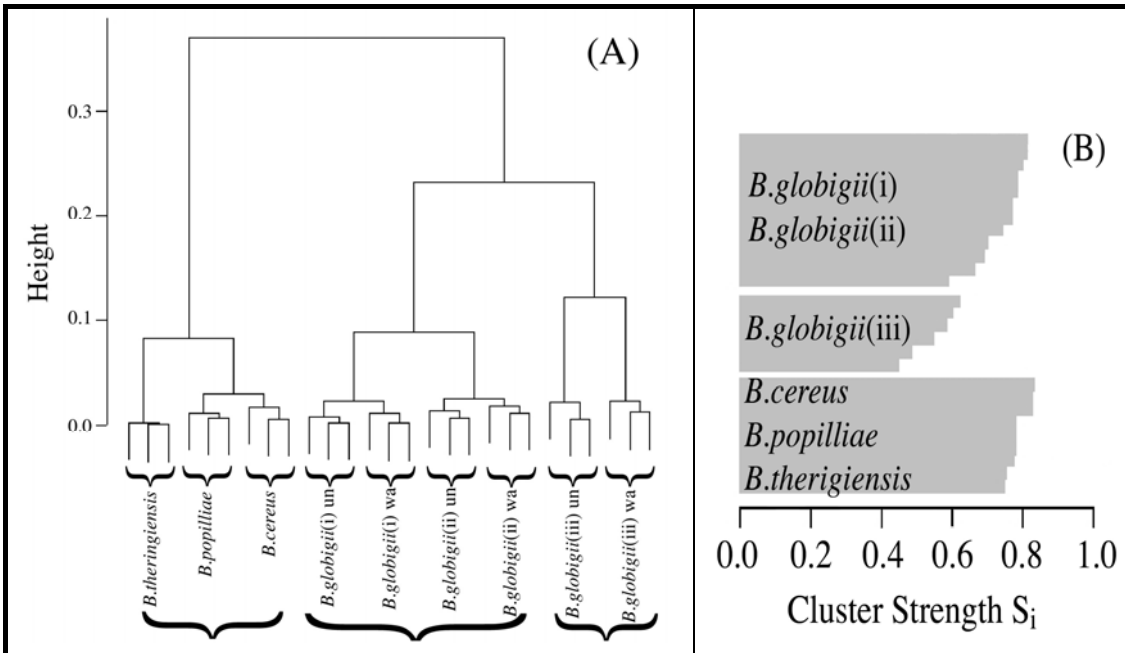


Figure 4.37 (A) Dendrogram of the clustering of objects according to the first and second PCs as shown in Figure 4.35 displaying the relationships among spectra of samples based on correlation coefficient calculated by PCA. Samples with "un" were measured unwashed and samples with "wa" were measured at being washed twice and redried. (B) The strength of the three clusters shown with the brackets in the bottom of Figure 4.37(A).

4.4 Autofluorescence of *Bacillus* spores and Role of Tryptophan: A Comparative study of Dry Spores and Spores in Aqueous and Ethanol Suspensions

In this section we will give the results of our study of tryptophan, its role for fluorescence signature of *Bacillus* spores (section 4.4.1) and the results we obtained for a comparative study of dry spores of *B. globigii* (i), spores suspension in aqueous and ethanol (section 4.4.2).

4.4.1 Tryptophan Absorption and Fluorescence

Absorption measurements were taken for dry, aqueous and ethanol suspension of tryptophan from scanning wavelength range 200 to 600 nm. Results are shown in Figure 4.38 (dry absorption spectrum is not included). The concentration of tryptophan in water was 9.058E-06 mol/litre and in ethanol is 7.834-06 mol/lit. The optical density at excitation wavelength 290 nm is 0.04536 and 0.06551 respectively. Ethanol peak is slightly red shifted (2.4 nm).

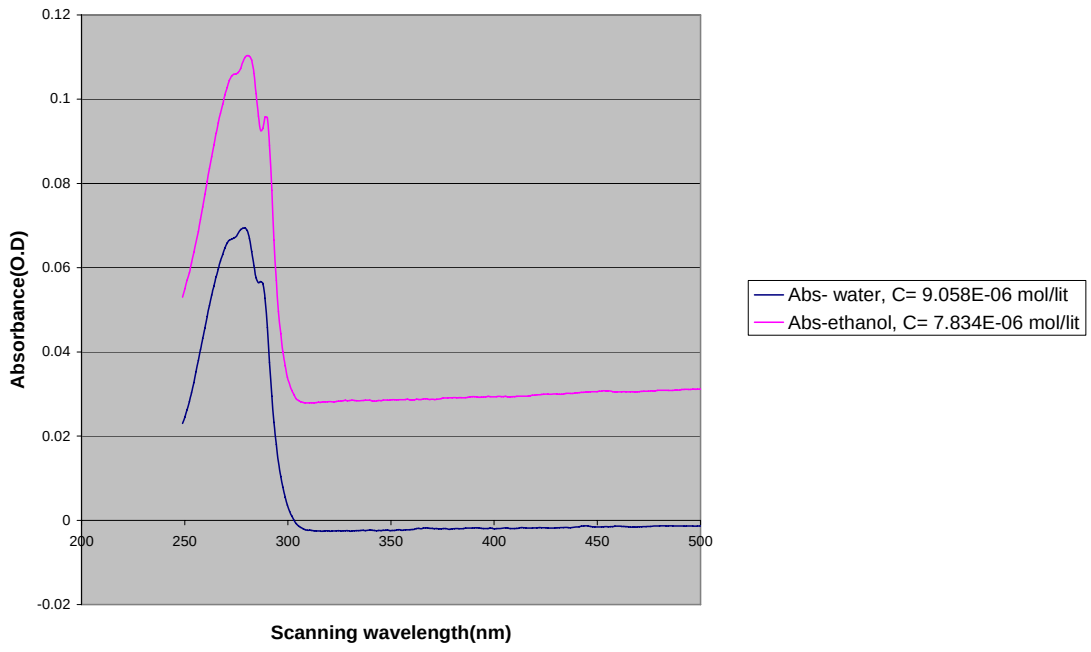


Figure 4.38 Sample absorption spectrum of tryptophan in aqueous (blue) and ethanol (red) suspension.

In Figure 4.39 we show sample fluorescence spectrum of tryptophan in dry, aqueous and ethanol suspension at 290 nm excitation wavelength with emission peaks at 339, 360, 348 nm respectively.

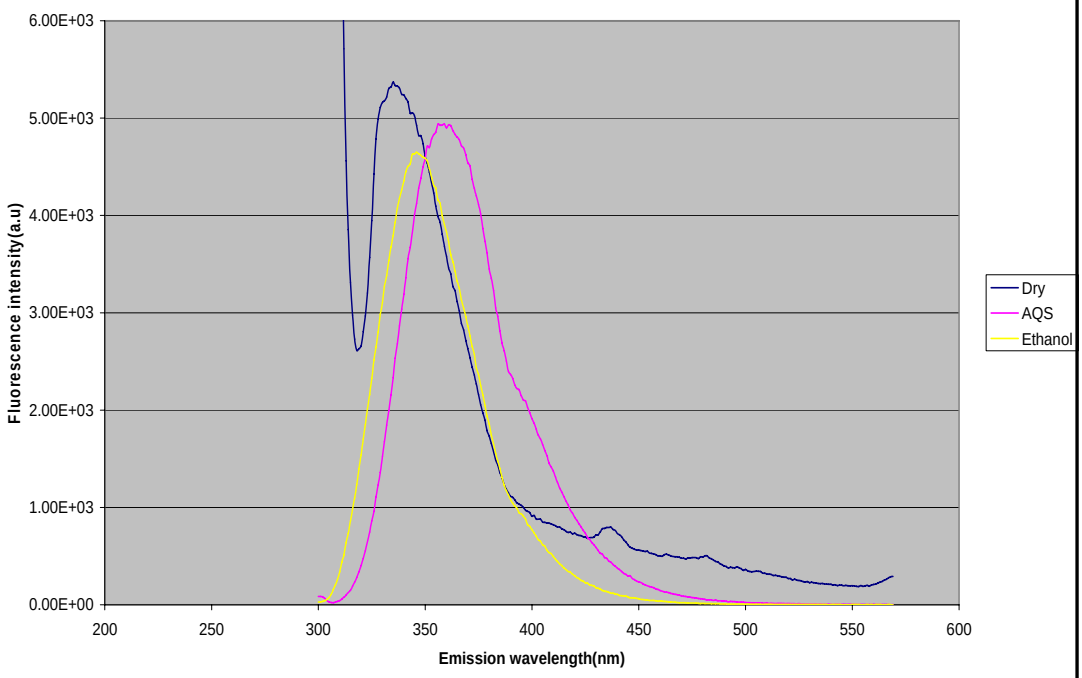


Figure 4.39 Sample fluorescence spectrum of tryptophan in dried form, in aqueous and ethanol solutions at 290 nm excitation wavelength with emission peaks at 339, 360, 348 nm respectively.

In Figure 4.40 we show the fluorescence fingerprints of tryptophan in dry (on filter paper), aqueous and ethanol suspensions. The excitation was stepped every 10 nm. The intensity is equally divided by 15 contour lines. The colours are added to help distinguish hills from valleys. The blue is the highest points and red is the lowest points. Note that for dry tryptophan the emission maximum occurs at 320 nm excitation, but for aqueous and ethanol suspension it seems to occur at a shorter excitation wavelength of the order 300 nm.

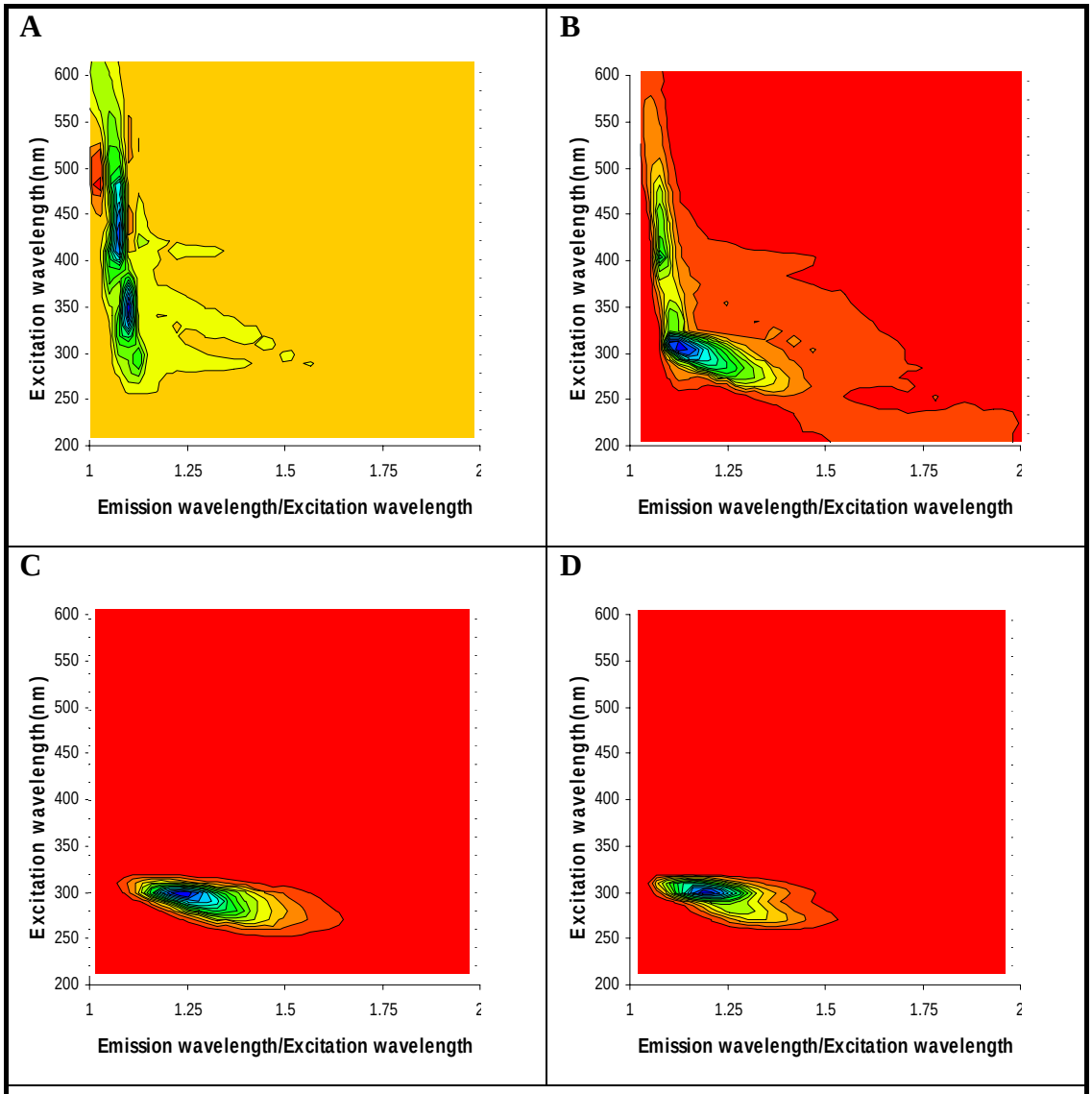


Figure 4.40 Fluorescence fingerprints of tryptophan in dried form, aqueous and ethanol solutions using excitation wavelengths from 200-600 nm. In Figure 4.40(A) Florescence fingerprint of selected filter paper is displayed and in Figure (B-C) the fluorescence fingerprints of tryptophan; (B) in dry, (C) in aqueous solution and in (D) in ethanol solution were displayed. The excitation was stepped every 10 nm. The intensity is equally divided by 15 contour lines. The greatest intensity is blue and the least intensity is red.

4.4.2 Autofluorescence of *Bacillus globigii* (i)

In an attempt to understand more about hydration effect on fluorescence of *Bacillus* spores, we repeated the fluorescence measurement with *B. globigii* samples. Measurements were done with dry spores on filer paper, spores in aqueous and ethanol suspensions, in fluorometer. The *B. globigii* samples showed distinct spectral signatures in dry and suspensions in aqueous and in ethanol. Only the results of fluorescence measurement of *B. globigii* (i) in dry spores in filter paper, spores suspensions in aqueous and ethanol is given below.

In Figure 4.41(A) we show the 2D fluorescence fingerprints for dried *B. globigii* (i) spores smeared on the filter paper. The spores essentially cover the filter paper, so we would expect little signal from the filter paper. In Figure 4.41(B-C) the 2D fluorescence fingerprints for *B. globigii* (i) in aqueous and ethanol suspensions are displayed. The maximum intensity of Figure 4.41(A-C) is normalized.

Note that the dried spores show an emission peak that is near 420 nm excitation and 500 nm emission wavelengths. There are changes in the fluorescence spectrum from the bacterial spore upon hydration. In the aqueous suspension of spores, the fluorescence profile is shifted to an excitation maximum at shorter wavelengths. The aqueous spores show an emission peak that is near 300-320 nm excitation and 400-420 nm emission wavelengths. In ethanol suspension the tryptophan peak is blue shifted about 10-15 nm. We also measured the distilled water that was used to suspend the *B. globigii* (i). There is no fluorescence seen. However, we do clearly see the familiar Raman peak of water (data not shown).

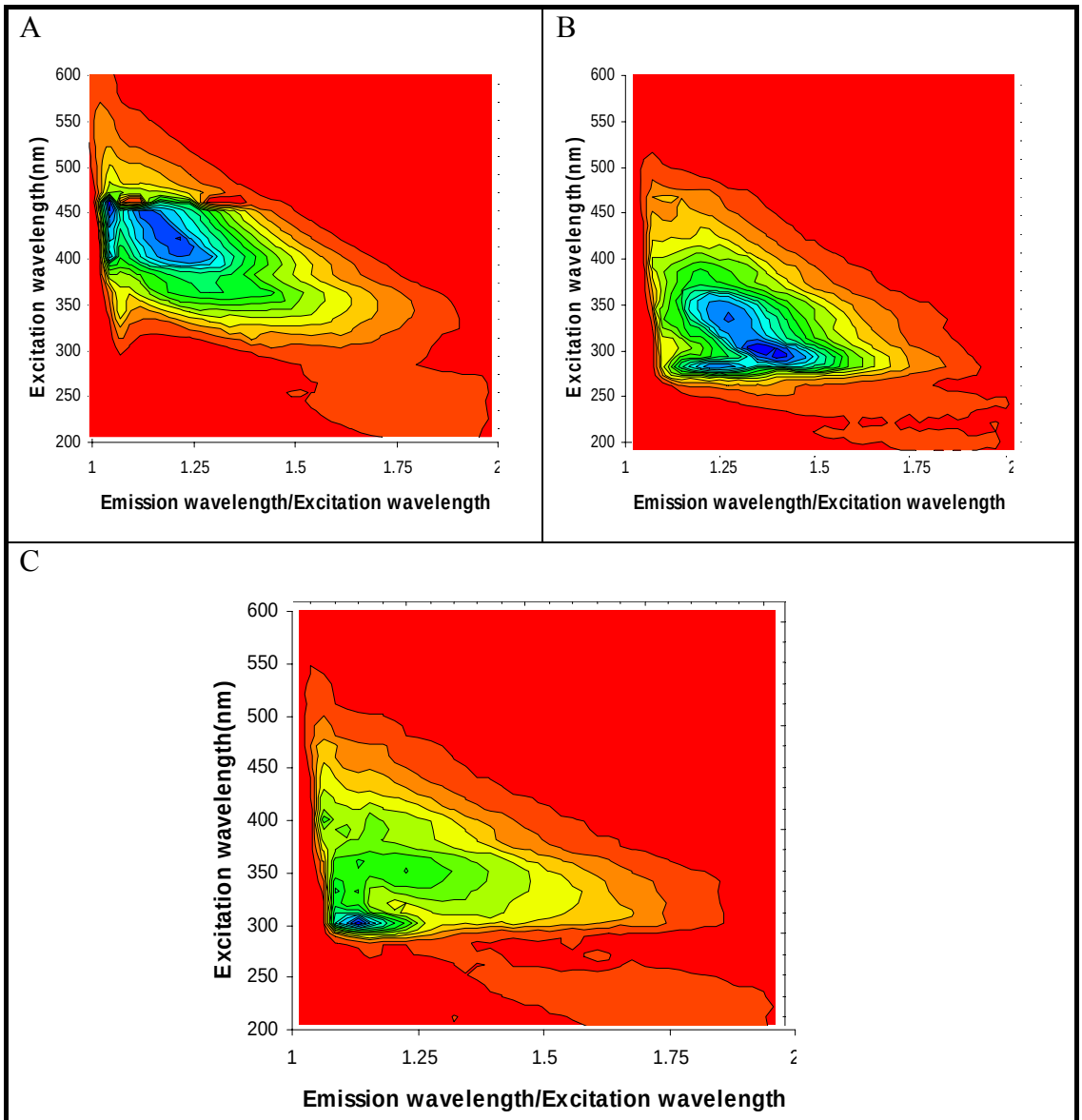


Figure 4.41 (A) Fluorescence fingerprints of dry *Bacillus* endospores (*B. globigii* (i)) on filter paper. In Figure 4.41(B-C) the fingerprints of *B. globigii* (i) endospores in aqueous and ethanol suspensions are displayed. The excitation was stepped every 10 nm from 200-600 nm. The intensity is equally divided by 15 contour lines. The greatest intensity is blue and the least intensity is red.

4.4.3 PCA and Cluster Analysis

A comparative PCA analysis for *B. globigii* (i) was done, as explained in the previous sections. The data matrix for this analysis was obtained from 9 fluorescence measurements of dry spores of *B. globigii* (i), spores in aqueous and ethanol suspensions (three measurements each) as explained in the experimental section.

Here also we restricted our selection of spectra to five different excitation wavelengths at 280, 310, 340, 370 and 400 nm respectively. The corresponding emission wavelength was selected from 310-410, 340-440, 370-470, 400-500 and 430-530 nm range (emission spectral range is 100 nm). The fluorometer we used measured emission in every 1 nm. Therefore in this selected wavelength as explained in the previous case, we have a total of 500 data points for each excitation wavelength for each spectrum and thereby for the entire data matrix for 9 samples will have 9x500 data points.

The correlation matrix for this analysis is given in the Table 4-18. The variables/samples are represent as *B. globigii* (i)-1 dry, *B. globigii* (i)-2 dry, *B. globigii* (i)-3; *B. globigii* (i)-1 aqueous, *B. globigii* (i)-2 aqueous, *B. globigii* (i)-3 aqueous, *B. globigii* (i)-1 ethanol, *B. globigii* (i)-2 ethanol and *B. globigii* (i)-3 ethanol for the three measurement of each of *B. globigii* dry spores, spores in aqueous and ethanol suspensions respectively. In Table 4-19 the eigenvalues are given.

Table 4-18 Correlation matrix for PCA analysis for *B. globigii* comparative study.

Variables	***B.globigii* (i)-1 dry**	***B.globigii* (i)-2 dry**	***B.globigii* (i)-3 dry**	***B.globigii* (i)-1 aqueous**	***B.globigii* (i)-2 aqueous**	***B.globigii* (i)-3 aqueous,**	***B.globigii* (i)-1 ethanol**	***B.globigii* (i)-2 ethanol**	***B.globigii* (i)-3 ethanol**
***B. globigii* (i)-1 dry**	1.00								
***B. globigii* (i)-2 dry**	0.98	1.00							
***B. globigii* (i)-3 dry**	0.99	0.99	1.00						
***B. globigii* (i)-1 aqueous**	-0.75	-0.71	-0.74	1.00					
***B. globigii* (i)-2 aqueous**	-0.77	-0.75	-0.76	0.96	1.00				
***B. globigii* (i)-3 aqueous**	-0.77	-0.74	-0.76	0.98	0.97	1.00			
***B. globigii* (i)-1 ethanol**	-0.55	-0.48	-0.52	0.62	0.66	0.64	1.00		
***B. globigii* (i)-2 ethanol**	-0.56	-0.47	-0.52	0.62	0.64	0.63	0.97	1.00	
***B. globigii* (i)-3 ethanol**	-0.58	-0.55	-0.57	0.75	0.83	0.78	0.87	0.84	1.00

Table 4-19 Eigenvalues

Principal Components	**PC1**	**PC2**	**PC3**	**PC4**	**PC5**	**PC6**	**PC7**	**PC8**	**PC9**
Eigenvalue	6.858	1.296	0.625	0.141	0.029	0.025	0.015	0.009	0.003

Note: Variance (PC1) = 6.858 or 76.2 %
Variance (PC2) = 1.296 or 14.4 %
Variance (PC3) = 0.625 or 6.93 %

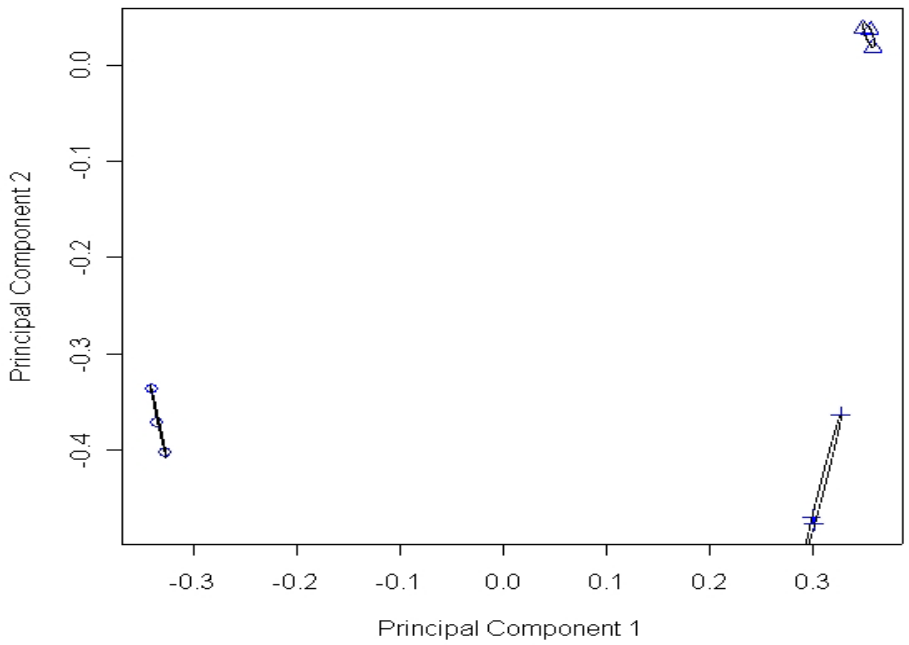
Combined PC1, PC2 will explain 90.6 % of the overall variance and satisfy the criteria [136] explained in the text.

In Figure 4.42 we show the PCA score plots of results of classification on replicate measurements (dry, aqueous and ethanol suspensions). PC1 and PC2 were used for plotting Figure 4.42(A) and for Figure 4.42(B), PC1, PC2 and PC3 were used.

Hierarchical clustering was applied to scrutinize spectroscopic similarities and spectral clustering. From silhouette plots the strength of clustering can be found by noting the average silhouette width. In Figure 4.43 we show the hierarchical as well as the silhouette plot due to PAM clustering for all three different preparations of *B. globigii* (i). There are 3 clusters from the dendrogram shown in Figure 4.43 (Top). The cluster groups are shown with the numbers at the bottom of the figure. In Figure 4.43(Bottom), we show the cluster strengths (as defined in Eq. 2.8). There are a total of 9 objects in the clustering scheme. The average silhouette width is 0.91. The top cluster contains 3 objects with an average silhouette width of 0.93. The middle cluster contains 3 objects with an average silhouette width of 0.97. The bottom cluster contains 3 objects with an average silhouette width of 0.82.

We have found that the dry spores, the spores' suspension in aqueous and ethanol are well separated. Slight variations can be seen from the clustering of suspended samples. We note that the suspensions probably contain a very different number of spores compared to the samples of spores smeared on filter paper.

(A)



(B)

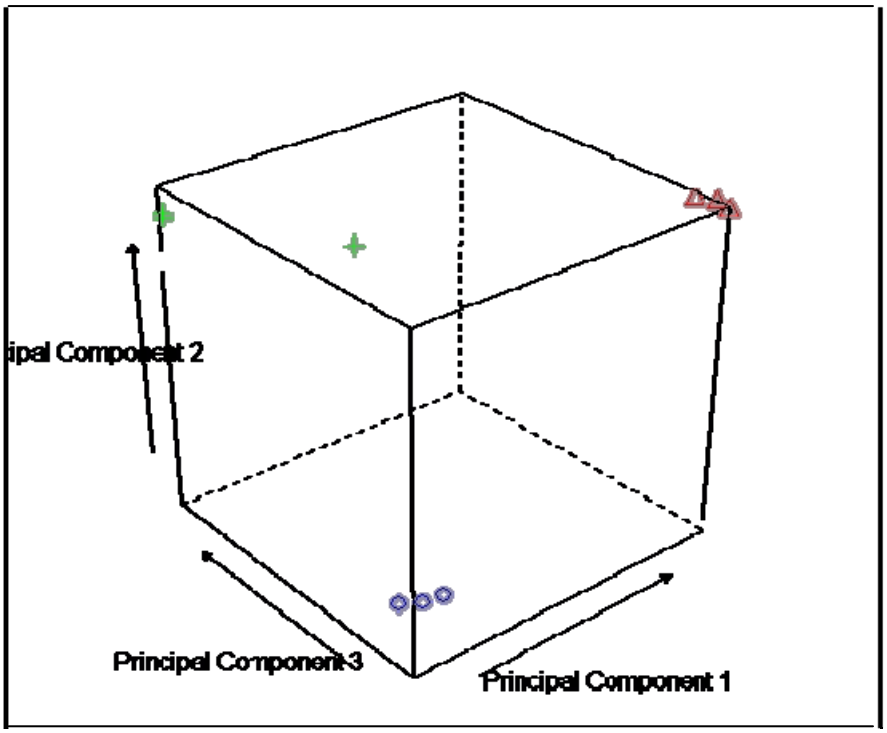


Figure 4.42 PCA plot of samples of *B. globigii* (i) spores in dried form (open circles), spores in aqueous (triangles) and in ethanol (plus) suspensions (9 measurements). In (A) PC1 versus PC2 is plotted and in (B) PC1, PC2 and PC3 are plotted

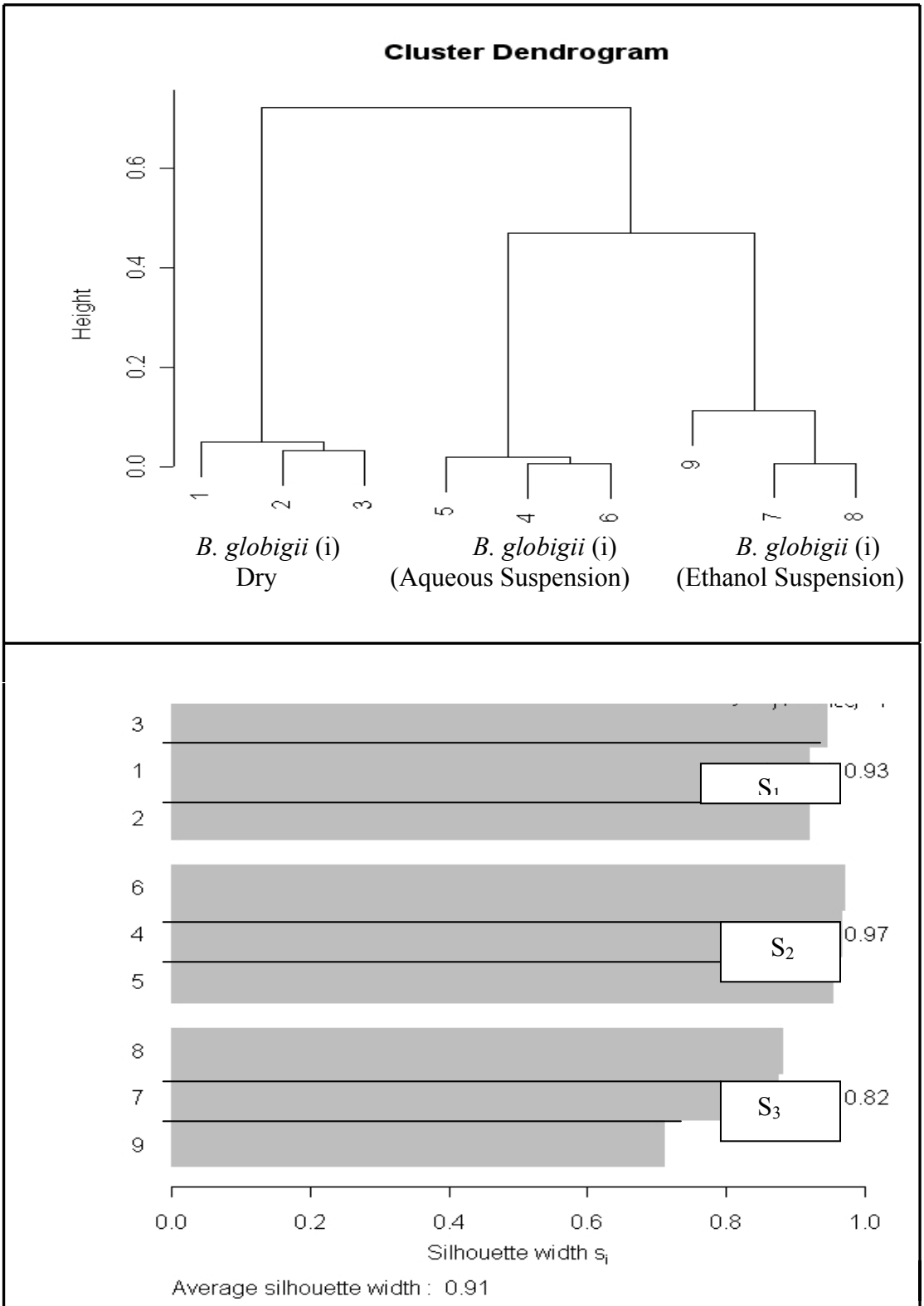


Figure 4.43 (Top) Dendrogram of the clustering of objects according to the first and second PCs are shown. (Bottom) The strength of the three clusters shown. The cluster strength is defined in the text.

4.5 Discussions on Spore Fluorescence and Data Analysis

In this section, a general discussion is given for the fluorescence measurements and data analysis that have been done in the previous four sections. In the following sections we will briefly discuss the most important features of the identification of *Bacillus* spores from different dust/pollen backgrounds and the effects of washing on identification and the comparative study of *Bacillus globigii* (i) using 2D fluorescence fingerprints and data analysis techniques.

The various fluorescence fingerprints of biological agent simulants investigated on different substrates and from different environments, yielded distinctive features. A few conclusions can be drawn from visual inspection of each fingerprint. Fluorescence spectroscopy does not significantly affect reproducibility or the accuracy of identification. Species-specific differences were noted in fluorescence fingerprints. Comparable fluorescence spectra had been found by Hill et al [55]. Our data confirmed the potential of fluorescence spectroscopy combined with analysis by PCA for a rapid, non destructive identification of *Bacillus* simulants [43, 56].

4.5.1 Spore Fluorescence

Bacterial spore contains a variety of intramolecular biomolecules associated with energy yielding reactions necessary for germination process. It retains numerous intrinsic fluorophores such as protein tryptophan, other amino acids and co-enzymes emitting photons following excitation in the UV region [155]. The *B. globigii* dry spores studied using fluorescence spectroscopy showed two peaks. One was a peak located at the excitation spectra peak near 290 nm with a peak near 330 nm for the emission spectra and the other was at 450 nm. Bronk and Reinisch [42] also found two peaks of fluorescence in the same relative wavelength regions, emission near 330 nm and 470 nm. The fluorescence emission spectra of the samples of other *Bacillus* samples like *B. cereus*, *B. popilliae* etc at excitation 280–290 nm are dominated by tryptophan emission peaks and confirm other studies [90]. However, the exact shape and size of this peak indicates the environment of the tryptophan. The emission spectra of the spores showed a maximum at 330 nm, suggesting a hydrophobic environment for its

tryptophan residues. The aqueous suspension of *B. globigii* spores shows fluorescence shift to 360 nm and in ethanol this the maximum was shifted to 340 nm, suggesting a rather more polar average location of the tryptophan [146]. These characteristic changes in the environment are the key to the species differentiation with fluorescence spectroscopy.

Any substrate can contribute fluorescence intensity to the measured intensity of the sample and this may hide or overlap features due to the sample alone. The substrates used for collection and identification of spores showed distinct spectral properties. For the tape collection, as in Figure 4.22(A), the plastic tape at excitation 280 nm produced a minimal amount of fluorescence and the peaks are away from the spore fluorescence peaks (fluorescence in the emission ranges around 300-308 nm). The fluorescence measurement of spores of *B. globigii* on the plastic tape showed no overlapping with the tape fluorescence as shown in Figure 4.22(C). *B. globigii* produces fluorescence at excitation wavelength 420 nm with emission maximum at 500 nm. Other *Bacillus* spores were also measured on plastic tape (data not shown). None of the spores produced fluorescence in the region where the tape fluoresces even though the tryptophan excites at 280-290 nm.

The plastic tape may not be useful as a substrate if we want to measure the fluorescence at excitation below 280 nm, since it contains many of the same elements present in the biological simulants [62]. Compared to the tape collection and fluorescence measurement, measurements on quartz slide seemed more useful, since it hardly produced any fluorescence (see the Figure 4.22(B)) and showed less scattering light leaking to the monochromator. *B. globigii* fluorescence on tape and quartz slide showed not much difference as seen in Figure 4.22(C) and 4.22(D). Similar results were obtained with other samples (data not shown). Quartz slides were also very useful for samples with a very small size.

In this study, using 2D spectroscopy we were able to correctly detect and discriminate various *Bacillus* spores (biological agent simulants) from a dust background. Our results were comparable with recent studies using LIBS methods for detection and classification of biological aerosols [62, 63]. It is evident from Figure 4.23(A) that the selected dust (A) is not producing much fluorescence. It does produce some fluorescence at excitation 280 nm which may be due to the plastic tape

substrate (see the Figure 4.22(A)) for tape fluorescence. It may produce some fluorescence due to tryptophan, since the collected dust was from home. Most home dust is mainly due to human skin that contains tryptophan. It is not clear how this affected our measurements. Similar results were obtained with two other dust samples fluorescence measurements (figures not given).

It was interesting to notice from Figure 4.23(B) that the spores' fluorescence of *B. globigii* (i) completely covered the dust A. It produced well defined fluorescence fingerprints and showed distinguished signatures for the samples investigated with the background of home dust samples. *B. globigii* spores produced a distinct fluorescence signature from the same dust A. Note that the dried spores show an emission peak that is near 420 nm excitation and 500 nm emission wavelengths (Figure 4.23(B). Similar results were obtained with other dust samples (figures not shown).

In Figure 4.23(C) we can observe the fluorescence signature of *B. cereus* with emission peaks around 330-340 at excitation wavelength 300 nm mainly due to tryptophan. The scattering due to dust and plastic tape substrate can still be seen and this can be minimized if we use more spores. It was interesting to note that the fluorescence fingerprint of *B. globigii* (i) showed no tryptophan fluorescence near 300 nm excitation as observed in *B. cereus*, thus *B. globigii* (i) is clearly identifiable and distinguishable from *B. cereus* even from the same background. This is an excellent finding for the discrimination of *Bacillus* spores using 2D fluorescence spectroscopy. At this stage no explanation can be given for the peaks at 420 nm excitation. More work needs to be done to get a satisfactory explanation for the cause of this peak.

In the case of *B. popilliae* (Figure 4.23(D)) from the same dust sample A, the fluorescence signature did not show much difference with the tape fluorescence. *B. popilliae* did not seem to produce much fluorescence even in the tryptophan fluorescence range unlike other *Bacillus* samples investigated in this part of the research. It produces less fluorescence at excitation 290 nm with emission peaks around 330 nm, but that may overlap with dust and tape fluorescence. It may be advisable to do the measurements in a quartz substrate.

The discriminatory nature of the fluorescence fingerprints of three samples *B. cereus*, *B. globigii* (i) and *B. popilliae* discussed above can be compared with other samples investigated in the same tape substrate without added dust samples. The fluorescence fingerprints of *B. globigii* spores prepared from an outside source, *B. globigii* (ii) showed almost the same fluorescence signature as that of *B. globigii* (i) measured with dust samples (Figure 4.24(A)). A noticeable difference is that *B. globigii* (ii) showed clear tryptophan fluorescence at 290 nm excitation with emission peaks near 330-340 nm in addition to the peaks at 420 nm, which is not as strong as that of *B. globigii* (i) at this excitation wavelength. The difference in fluorescence intensity at this region may be due to different concentrations of spores used in the two samples measurement.

Fluorescence measurements with non bacterial samples of corn smut and ovalbumin is another area of discussion. It seems a fluorescence fingerprint of corn smut is similar to *B. globigii* (ii) (Figure 4.24(B)), but it can be distinguished from the fluorescence signatures of *Bacillus* spores by careful observation of its various peak positions. It also produces fluorescence near 300 nm excitation with emission near 340 nm, but peaks at 420 nm excitation is not as strong as that of *B. globigii* samples. It is not clear that simply comparing with signatures of *Bacillus* spore is the best method to distinguish a non bacterial spore like corn smut from *Bacillus* spores.

An interesting finding of fluorescence fingerprints with ovalbumin was that it produced three distinct peaks, two of which are entirely at different excitation wavelengths and one exactly in the same position of tryptophan fluorescence, as that of *Bacillus* spores investigated in this research (Figure 4.24(C)). Ovalbumin in a dry state produced one peak at excitation 290 nm with emission peak near 330 nm which is similar to tryptophan peak, since ovalbumin is a protein that also contains tryptophan. The second peak produced at 370 nm excitation with emission near 440-450 nm. The third fluorescence peak occurred at 390 nm excitation with emission maximum at 440 nm. At this stage no explanation can be given for fluorescence peaks at 370 and 390 nm excitation wavelength for ovalbumin.

We are extending the discussions of spectroscopy results that we have done in the previous section, to the detection and identification of *Bacillus* spores from pollen background. We will limit this discussion to the fluorescence measurement of

B. cereus, *B. globigii* (iii), pig weed and a mixture of *B. globigii* (iii) with pig weed made on quartz substrate as it further reinforces the applicability of 2D fluorescence spectroscopy in discrimination of *Bacillus* spores from background fluorescence.

In the case of *B. cereus* sample used in this investigation (Figure 4.28(A)), no further discussion can be added except to mention the type of substrate used. It was measured on a quartz surface as explained in the text, and similar fluorescence fingerprints were obtained. More discussion will be given in the applicability of multiple samples in the PCA analysis of this section. Fluorescence of *B. globigii* (iii) was also measured on a quartz surface as explained in the text. As seen from the Figure 4.28(B), *B. globigii* (iii) produced almost the same fluorescence signature as that of *B. globigii* (ii) that we have discussed earlier.

As discussed in the previous section, dust samples used for the study of background fluorescence were collected from the home environment and were found not to produce much fluorescence (Figure 4.23(A)). Therefore it is quite natural to believe that we can detect and discriminate fluorescence of *Bacillus* spores from home dust. But this picture will be different if we choose to collect dust or pollen from the environment and we did with pig weed, a pollen grain.

Surprising results were obtained as seen from the Figure 4.28(C-D). In fluorescence fingerprint of pig weed (Figure 4.28(C)), we can see a well defined emission peak around 550 nm with excitation at 480 nm and this peak is well away from any of the other peaks due to the *Bacillus* spores investigated. When this pollen was mixed with one of the *Bacillus* spores: *B. globigii* (iii), we noticed with surprise, that the mixed fluorescence signature yielded a very good result (see the Figure 4.28(D)). The fluorescence signature shows a well discriminated peaks ranges and offered some sort of insight into detecting *Bacillus* spores mixtures/environment particles like pollen grain. Here, with mixed samples of *B. globigii* (iii) and pig weed, we noticed that we can detect and discriminate the fluorescence due to *Bacillus* spores from pollen grains. Comparing Figures 4.28(B) and 4.28(D), we can see that Figure 4.28(D) contains the same information as that of Figure 4.28(B), except that the peak at 420 nm due to *B. globigii* (iii) is not as strong as that in the Figure 4.28(B). This is only a preliminary investigation and more research needs to

be done, with more *Bacillus* spores with varying pollen grain samples to detect and discriminate *Bacillus* spores from mixtures.

We checked for changes in the fluorescence spectrum before and washing the spores. As explained in the text, we measured the fluorescence signals of six samples of *Bacillus* spores: *B. globigii* (i), *B. globigii* (ii), *B. globigii* (iii) (all washed and unwashed), *B. cereus*, *B. popilliae* and *B. thuringiensis* (unwashed) for the study of the effects of washing on the detection and identification of *Bacillus* spores. In the previous sections we have already discussed the fluorescence signatures of above samples (unwashed) except *B. thuringiensis*. Therefore, only a brief account will be given for the fluorescence properties and will be discussed in detail with PCA analysis.

Regarding *B. cereus* and *B. popilliae* fluorescence fingerprints as seen in Figure 4.32(A-B) no further discussions can be added. The fluorescence fingerprint of *B. thuringiensis* (Figure 4.32 C) showed some what similar characterises of *B. globigii* samples except the fluorescence peak at 420 nm excitation is not as strong as *B. globigii* samples whereas the peak due to excitation at 360 nm seemed more prominent. No explanation can be given at this stage.

Looking at washed and unwashed fluorescence fingerprints of *B. globigii* samples (Figure 4.33(A-F)), we noticed essentially no difference except a difference in fluorescence intensity due to washing. But this study showed that *B. globigii* spores can be distinguished from other *Bacillus* spores, even when different laboratories produce the spores.

In the following sections, a brief account of the hydration effects on spores' fluorescence is given in comparison with tryptophan fluorescence. As mentioned in the experimental section, we compared the spores' fluorescence in dry, spores suspensions in aqueous and ethanol.

Tryptophan in solution in water has two excitation optima, one about 225 nm and one about 280 nm, and emission maxima for either excitation occurring with peaks about 350 nm. When tryptophan is bound in a cell or spore the emission maxima are shifted to shorter wavelength and the peak is near 290 nm. The spectrophotometer and fluorometer that we used was not sensitive to measure the absorption and fluorescence at excitation at 225 nm.

As seen from Figures 4.38 and 4.39, our absorption measurements with aqueous and ethanol solution at excitation 290 nm, showed the ethanol peak slightly red shifted (2.4 nm) and for fluorescence measurement with ethanol, the ethanol peak is blue shifted (by 10 nm). As seen from Figure 4.40(B-D) the fluorescence for dried, aqueous and ethanol solutions of tryptophan showed very distinct profiles. As expected, the dry form of tryptophan peak emission was at 330 with excitation 290 nm on filter paper. This filter showed two distinct peaks at excitation wavelengths 340 and 420 nm with emission peak at 380 and 460 nm respectively (see the Figure 4.40(A)) and did not overlap with the tryptophan peak. Emission peak due to aqueous solution of tryptophan is at 360 nm with the same excitation at 290 nm whereas emission peak due to ethanol was blue shifted to 350 nm, (see Figure 4.40(C-D)). This was not expected.

We initially thought that tryptophan in ethanol behaves just like dry spores and expected no shift/difference in fluorescence peaks. But the results show that like water molecules move into tryptophan environment, ethanol molecules also moved into tryptophan environment. This finding leads us to check the role of tryptophan in spores' fluorescence and did as a special case of fluorescence of *B. globigii* (i) in dry spores, spores suspension in aqueous and ethanol suspension and the result is given below.

This section, looks especially at the fluorescence of dry spores of *B. globigii* (i), spores in aqueous and ethanol suspension and noticed that *B. globigii* (i), show a distinct difference in the fluorescence profiles. This result was not expected. As seen from Figure 4.41(A) dried spores of *B. globigii* (i) on filter paper show a fluorescence peak that is near 420 nm excitation and 500 nm emission wavelength. Also note from Figures 4.41(A) and 4.41(B), that the fluorescence profile is shifted to an excitation maximum at shorter wavelengths in the aqueous suspension. The aqueous spores show an emission peak that is near 320 nm excitation and 420 nm emission wavelength.

In ethanol suspension (see Figure 4.41(C)) no such shifting can be seen except that the tryptophan peak blue shifted about 10 nm. This is an excellent finding and confirms our earlier findings with tryptophan fluorescence (recall the results due to tryptophan fluorescence in ethanol suspension). Also note that the peak that we have

seen in aqueous spores near 420 nm excitation is not strong in ethanol suspension. This does not mean that ethanol environment will not have any influence on spore fluorescence. Since the emission peak is shifted to the blue region, we can assume that ethanol molecules move into the spore cortex as well as the spores' interior. Unfortunately no literature can be quoted to support this assumption.

The spore is thought to be inert and in a resting phase. Only recently has it be measured that *Bacillus* spores increase in size when hydrated [147, 148]. The measurement of the size increase was fit to two diffusion rates. The two rates are interpreted as water not only penetrating into the spore coat and cortex, but also to be penetrating into the spore core. If the rates for hydration are as fast as reported in reference [148], then our samples were completed hydrated with the two hour equilibration time.

We initially thought that the change in fluorescence between the wet (aqueous) and the dry spore might be due to phosphorescence. The wet spore would contain enough water to allow the phosphorescence to be quenched. However, the dry spore would have a strong phosphorescence signal that would mix with the fluorescence. If this were the case, then the fluorescence profile of the dry spore minus the fluorescence profile of the wet spore should leave just the phosphorescence of the dry spore. As such we plotted the difference spectra (figures not shown), surprisingly none of the difference spectra resembled the phosphorescence measured [141]. We also measured the phosphorescence from the wet spores (data not shown). When the spores were wet, the phosphorescence was, as expected, essentially zero. Phosphorescence, alone, cannot explain the difference in the fluorescence signals observed between the dry and the wet spores. There are other changes that must also be occurring. The changes in the fluorescence spectra are very large, indicating important changes within the spore.

There have been other indications that the interior of the spore changes when the spore is hydrated. Neihof et al.[156] stated that the interior of the spore is not maintained as anhydrous by observing similar curves for water absorption and desorption as a function of the relative humidity with intact and crushed *Bacillus subtilis* spores. Marshall and Murrell [157] used deuterated and tritiated water with spores and were able to conclude that at least 97% of the water in spores is able to

exchange with the environment. Furthermore, they found that this exchange is almost complete in 2–3 min at 0°C.

Still, the nature of this water in the interior of the spore is not fully understood. While these measurements do not provide an absolute answer to role of the water and changes in the water, they do indicate that the water in the spore does have important interactions with the proteins in the spore. The changes in the fluorescence spectrum between wet and dry spores also mean that investigators trying to formulate protocols to detect and identify bacterial spores need to be concerned about the hydration of the sample. Many measurements are made with the spores in suspension as a safe method of handling spores. However, these measurements might not be appropriate to characterize the fluorescence from dry spores. Conversely, one might be able to take advantage of changes in the fluorescence signal upon hydrating a sample to help to discriminate among different bacteria or to separate the signal of background dirt and dust from the contaminant spore.

As a future direction, for identification and detection of pathogens from other biological agents, measurements of fluorescence lifetimes may be useful. The time range for fluorescence processes is much shorter than microseconds. A frequency domain fluorescence study of calcium binding metalloproteinase from *Staphylococcus aureus* has shown that this two-tryptophan contain protein exhibits a double exponential fluorescence decay [158]. At 10^0 C in 0.05 M Tris-HCl buffer (pH 9.0) containing 10 mM $CaCl_2$, fluorescence lifetimes of 1.2 and 5.1 ns are observed. Steady state and fluorescence domain solute quenching studies are consistent with the assignment with the assignment of the two lifetimes to the two tryptophan residues. The tryptophan residue characterised by a shorter lifetime has a maximum of fluorescence at about 317 nm and the second one exhibit a maximum of its emission at 350 nm.

From this study and others [159], the bacterial fluorescence lifetime is assumed to be very shorter (much shorter than 1 micro second). Macromolecules like pathogen spores can exist in more than a single conformation. The intensity decay could reveal two decay times and thus the presence of more than one conformation state. Thus by measuring and comparing the fluorescence lifetimes, the pathogens can be discriminate from other biological agents found in the background.

To conclude this section of discussions on EEM spectroscopy applications of dry spores, we could say that, in general, the contours of measured fluorescence signals showed the expected peak related to tryptophan at a fluorescence wavelength of 330 nm. An exceptional case is the fluorescence of dry spores of *B. globigii* (i), which showed no tryptophan peak, but showed the peaks when the spores were in aqueous and in ethanol suspensions. Fluorescence spectra of all simulants studied is reproducible. We have also showed that a *B. globigii* spores can be distinguished from other *Bacillus* spores, even when different laboratories produce the spores. Fluorescence QE measured for spores are reproducible to within a factor of 2 to 5.

4.5.2 Data Analysis

Fluorescence spectra of *Bacillus* spores contain information derived simultaneously from all chemicals (metabolites and cell components) that are present in the spores. The advantage of fluorescence spectroscopy compared with other biochemical methods of identification is that this information is retrieved rapidly from a single test. The need to assign the respective fluorescence spectra to particular metabolites is circumvented by analyzing variations in the signal pattern rather than the identification of individual compounds. The PCA was specifically useful for the analysis of fluorescence spectra where individual spectra contain many more data points (attributes) than the number of spectra (sample size). The advantages of PCA for attribute reduction used in fluorescence spectroscopy was supported by other studies for analysis of NMR, FTIR, and Raman spectra [160-163].

If the spectra of the biologics have repeatable differences, the data points from each sample should cluster together, showing successful discrimination of a sample-specific fluorescence signature. The most discriminatory spectral regions identified by PCA point to particular metabolites that contributed to a successful discrimination. These spectral regions (metabolites) are consistent for all spores within the particular species. Thus, fluorescence spectroscopy/PCA also provides a means of rapidly screening for stable phenotypic markers (metabolite profiles) that can be used for classification. The robustness and reliability of the fluorescence spectroscopy/PCA analysis were proven by testing against additional, independent samples that were not included in the development of the initial classifiers.

Classifiers were considered robust when the accuracy and crispness of identification using an independent validation set approached those obtained on the training set.

PCA is a useful tool for analysing spectral properties biological samples and has been able to provide an excellent data reduction technique [62, 63]. However, there are some weaknesses with the approach used in this research for detection and identification. PCA produces an artefact known as the horseshoe effect (similar to the arch effect), in which the second axis is curved and twisted relative to first, and does not represent a true secondary gradient (see for example, the Figure 4.29).

It is not clear from this research whether PCA is able to distinguish between spore forming and non spore samples. PCA does not cope well with small data sets. As seen in this research, the results tend to be affected by outlying samples. For instance, in Figures 4.27(C- D), the ovalbumin and corn smut, the non bacterial samples cluster with the *B. cereus*. This is an incorrect clustering.

We selected only 5 excitation wavelengths for the PCA. We chose excitation wavelengths where the fluorescence from the spores is large. We also chose a 100 nm range of emission wavelengths to sample the typical 50 to 100 nm wide fluorescence peaks. Since our dry samples reflect stray light into the detection system, we started our emission wavelength collection at 50 nm longer than the excitation wavelength.

In the case of detection from three different dust backgrounds, even with the small number of samples, it is clear that both PCA and PAM clustering separate and distinguish among the three different bacterial spores of *B. globigii* (i), *B.cereus*, and *B. popilliae* (Figures 4.26A and 4.27A). Even when one of the spores (*B. globigii* (ii)) is produced from a different source, it is still appropriately classified (Figures 4.26B and 4.27B). When samples are used that are not bacterial spores (for examples, ovalbumin and corn smut), the PCA and PAM cluster analysis attempted to cluster these samples with the *B. cereus* (Figures 4.26(C-D) and 4.27(C-D)). Since PAM must be given the number of clusters to use, PAM would apportion all the samples into one of three clusters. If a sample really did not belong in a particular cluster, then this would be indicated in the silhouette plot, with the strength (S_i) considerably lower than the other strengths. Each sample is measured as a dry spore on tape with a background of domestic dust from three different locations (living

room, bathroom and bedroom). Even though these samples contain a mixture of materials, one can see by the grouping of the data that the spore fluorescence signal is consistent and therefore dominates in the analysis.

In the case of detection spores from pollen background, it is again clear that both PCA and PAM clustering separate and distinguish two of the bacterial spores (*B. cereus* and *B. globigii* (iii)) among the four different samples. Notice that the non *B. globigii* sample, *B. cereus* was clearly clustering separate from *B. globigii* (iii) as well as non bacterial samples, pig weed and the combination of *B. globigii* (iii) and pig weed (see Figure 4.31). Also notice that from the mixed samples of *B. globigii* (iii) and pig weed, *B. globigii* (iii) clustering separate and distinguish from *B. cereus* and pig weed.

When we use more variable/samples for PCA and clustering, we will face some misclassification as we observed in PCA analysis of effects of washing on fluorescence data. We have 27 variables from nine different preparations of six samples of which three samples belong to the same *Bacillus* species (*B. globigii*). Using visual inspection, one can see that *B. globigii* can be identified separate from the other samples from any of the three preparations (Figures 4.35 and 4.37). It is evident from Figure 4.35 that the six *B. globigii* samples (circles, squares and triangles) clustered well as expected. There is a slight change with the *B. globigii* (iii) before and after washing (solid and open squares, respectively). We suspect that these spores contain more growth media and cell debris compared to other *B. globigii* samples. Still, all three *B. globigii* and clearly discriminate with other *Bacillus* species: *B. cereus*, *B. popilliae* and the *B. thuringiensis* samples.

Also, washing and redrying the spores does not create major changes in the fluorescence and that is reinforced with PCA. It can be seen that fluorescence combined with PCA is a valuable tool in identifying bacterial spores. Notice from Figures 4.35 and 4.37 that different samples of *Bacillus* spores are correctly cluster together before and after the spores have been washed and redried and *B. globigii* spores can be distinguished from other *Bacillus* spores, even when different laboratories produce the spores.

In this research, PCA and cluster analysis played a crucial role in understanding more about the hydration effects of spores' fluorescence. In the spectroscopy section,

we have already discussed the changes in luminance between dried and aqueous spores based on fluorescence fingerprints. Here, we are reinforcing our discussion in a meaningful way.

From the Figure 4.43 it is evident that PCA and PAM can explain to some extend (at least qualitatively), the hydration effects of the tested sample. The study of spores shows clear differences in fluorescence between dried *B. globigii* spores, spores in aqueous and ethanol suspension. Our results also revealed a strong effect of hydration on the fluorescence spectra of the *Bacillus* endospores, implying water migration between inner compartments and the environment. The effect of ethanol suspension can not be compared with that of water suspension.

The results of PCA discussed above are stable if minor changes are made in the analysis. For instance, if six excitation wavelengths are used instead of five, the results are not significantly changed. However, if major changes are made, such as using only excitation wavelengths longer than 440 nm, the results are significantly affected. At these longer wavelengths there is very little fluorescence signal from the spores and signal noise create spurious results with the PCA. Likewise, if the emission wavelengths include values closer to the excitation wavelengths, then elastic scattering from the sample dominates the signal and the PCA analyses the amount of elastic scattering over the smaller fluorescence signal. While the PCA analysis is reasonably robust to measurement parameters, the results can be invalidated by a careless selection of measurement parameters.

Chapter 5

Conclusions

Spectrofluorometry has been used to investigate many living bacterial - spores [53-55, 92, 164]. The fluorescence technique has been considered previously for bacterial pathogen identification [39, 40, 42]. Fluorescence spectroscopy is a sensitive technique to identify fluorophores of bacterial spore and investigate the environment of fluorophores within the spores. In this technique, an excited wavelength of light induces fluorescence within the spore, which is received by a sensitive photodetector. The analysis of the emitted fluorescence as a function of the excited wavelength results in a unique spectral pattern. One important purpose of the fluorescence measurements made with different *Bacillus* simulants was to determine the difference in the fluorescence spectra, which could be used for potential identification.

Fluorescence spectroscopy/PCA also provides a means of rapidly screening for stable phenotypic markers (metabolite profiles) that can be used for classification. The robustness and reliability of the fluorescence spectroscopy/PCA analysis were proven by testing against additional, independent samples that were not included in the development of the initial classifiers. Our study shows that EEM fingerprinting and PCA analysis are viable techniques for detection and identification of *Bacillus* spores from different environments.

Fluorescence spectroscopy has been shown to be a potentially powerful microorganism fingerprinting technique for rapid identification of bacteria and fungi. The good news about the resonance fluorescence technique is that it is fast [40] and simple. The difficulty with fluorescence spectroscopy is that although it can tell the difference between dust and bacterial spores it cannot fully differentiate between spores and many other organic bioaerosols. However, despite the encouraging success of the above-mentioned studies, there is still interest in other approaches to, and tools for, the rapid identification of bacteria, bacterial spores and biological substances. To quote from a previous study [165] ,

"Current [fluorescence-based] prototypes are a large improvement over earlier stand-off systems, but they cannot yet consistently identify specific organisms because of the similarity of their emission spectra. Advanced signal processing techniques may improve identification".

We have measured the fluorescence emission spectra as a function of the excitation of the wavelength for several bacteria in dried and in suspension. The *in vitro* samples consisted of the bacteria spores at room temperature. The fluorometer measured the emission spectra with a 45/90 degree scattering geometry and collect light from a 1 cm quartz cuvet containing bacteria in dried and aqueous forms. The excitation wavelength was scanned from 200 to 600 nm in 10 nm steps. The data were smoothed with a sum of Gaussians, least squares fit to the measured data. The smoothed data were presented as a contour plot and stored as a square matrix.

We have measured the fluorescence signal as a function of the number of spores. The fluorescence QE of the samples under study were compared relative to a standard solution of a fluorophore such as anthracene. The fluorescence yield was from a measurement of absorption and fluorescence radiance. The absorption is a strong function of wavelength and a series of measurements of absorption and fluorescence were performed for different concentration/numbers then the measured fluorescence intensity / optical densities, implies the relationship. The signal to noise ratio will be small if we measure the fluorescence of the small number of spores. We can use these data to determine the limits of detection and the limits of identification for the algorithms.

The fluorescence QE measurements of samples studied, show unique properties that the main contribution to QE is due to the tryptophan absorbance near 290 nm excitation. Both washed and unwashed samples showed more or less the same QE at 280 nm, suggesting that fluorescence QE were independent of the excitation wavelength in the range 280–290 nm, whereas the QE value varies as the excitation increases.

The spectroscopy results of *Bacillus* spores from different dust/pollen backgrounds shows distinct discrimination between spores and other non *Bacillus* spores investigated in this research. All spectra measured were reproducible.

Samples which show less florescence can be measured on quartz slide and may be more reliable than in plastic tapes for minimising false positive detection. With dilute room temperature suspensions, we have found reproducible characteristics in the fluorescence spectra from several different species of bacteria.

We checked for changes in the fluorescence spectrum before and after washing the spores. The effect of washing on spores' fluorescence shows no problem with the discrimination features of fluorescence spectroscopy, but showed considerable difference between the values for QE between washed and unwashed samples. We are able to correctly cluster together different samples of *Bacillus* spores before and after the spores have been washed and redried. We have also showed that *B. globigii* spores can be distinguished from other *Bacillus* spores, even when different laboratories produce the spores.

The fluorescence emission spectra of the samples investigated at excitation 280–290 nm were found to be dominated by tryptophan emission. Despite these samples containing a large number of tyrosine residues, no shoulder was observed in the 305–315 nm regions. The emission spectra of the spores showed a maximum at 330 nm, suggesting a hydrophobic environment for its tryptophan residues. Spores in ethanol suspension, the maximum were shifted to 340 nm, suggesting a rather more polar average location of the tryptophan. In general, the excitation spectra peak near 280 nm with a peak near 330 nm for the emission spectra. The peak is dure primarily to tryptophan [90]. However, the exact shape and size of this peak indicate the environment of the tryptophan. These characteristic changes in the environment are the key to the species differentiation with fluorescence spectroscopy.

The autofluorescence of *Bacillus* spores changes with the hydration of the spores. The dried spores show a phosphorescence signal that is not present in the hydrated spore [141]. However, the difference between the dry and hydrated spore is more than the phosphorescence. These measurements support other studies that indicate the water is able to penetrate into the spore and affect the fluorescence signal form the spore interior.

We found that the emission spectra offer some preliminary information for identification of species, particularly for remote sensing where other means are not available. The emission spectra for bacteria do show some recognizable differences

between species. The emission spectra for several types of bacterial spores excited by 280 nm light appear quite similar to each other. Our experiments showed that the spectrum for dried spores appears rather different from that for spores, which have been suspended in water or in ethanol.

PCA is an effective and convenient way to reduce the high dimensional fluorescence data. The classification of samples using cluster analysis is easily and quickly evaluated from the PCs. Yet, there are obvious limitations. We selected only 5 excitation wavelengths for the PCA. We chose excitation wavelengths where the fluorescence from the spores is large. We also chose a 100 nm range of emission wavelengths to sample the typical 50 to 100 nm wide fluorescence peaks. Since our dry samples reflect stray light into the detection system, we started our emission wavelength collection at 50 nm longer than the excitation wavelength.

Even with the small number of samples, it is clear that both PCA and PAM clustering separate and distinguish among the three different bacterial spores. Even when one of the spores is produced from a different source, it is still appropriately classified. When samples are used that are not bacterial spores, the PCA and PAM cluster analysis attempted to cluster these samples with the *B. cereus*. Since PAM must be given the number of clusters to use, PAM would apportion all the samples into one of three clusters. If a sample really did not belong in a particular cluster, then this would be indicated in the silhouette plot, with the strength (S_i) considerably lower than the other strengths.

Notice that different samples of *Bacillus* spores are correctly cluster together before and after the spores have been washed and redried and *B. globigii* spores can be distinguished from other *Bacillus* spores, even when different laboratories produce the spores. It is evident that PCA and PAM can explain to some extend (at least qualitatively), the hydration effects of the tested sample. The study of spores shows clear differences in fluorescence between dried *B. globigii* spores, spores in aqueous and ethanol suspension

PCA is a powerful tool for the extraction of low dimensional information from heavily over determined high dimensional data. We do not claim that our clustering approach outperforms all the ones already suggested for the analysis of

microorganism expressions. Certainly, we can find situations where it is less appropriate. But we claim that it offers a number of advantages:

- Our method is partition technique as well as hierarchical and can easily be compared just by changing the size of the kernel used for checking the validity of different algorithms for finding the strength of clustering.
- It does not make any assumption concerning the shape, the size and the volume of clusters.
- The need to perform dimensionality reduction (like PCA) as a pre-processing step may be seen as a handicap. But in fact, we think that trying to perform clustering directly in a one hundred or even thousands- dimensional space is a very risky procedure. In addition, mapping onto a low- dimensional space offers many advantages in terms of visualization and qualitative interpretation of the data set.

In conclusion, we showed that the steady state UV-VIS induced primary emission spectra are likely to be useful for preliminary identification of microrganism. We have demonstrated a rapid procedure for detection and identification of *Bacillus* spores, using a multiparameter fluorescence techniques and pattern recognition of the fluorescence data by PCA. The results are in agreement with previous studies. Detection and identification were based on quantitative measurements of fluorescence from different backgrounds and the similarity/dissimilarity from PCA analysis of the same data rather than by subjective interpretation of biological and morphological characteristics. Fluorescence measurements combined with PCA suggest that it may be possible to readily discriminate bioaerosols from interferences.

Chapter 6

References

1. L. M. Bush, B. H. Abrams, A. Beall , and C. C. Johnson "Index case of fatal inhalational anthrax due to bioterrorism in the United States," New Engl. J. Med. **345**(22), 1607-1610 (2001).

2. W. K. Joklik, H. P. Willett, D. B. Amos, and C. M. Wilfert, "The classification and identification of bacteria," in *Zinsser microbiology*, W. K. J. e. al., ed. (Appleton and Lange, Norwalk, Conn, 1992), p. 13.

3. G. Gould and A. Hurst, *The Bacterial Spore* (Academic Press, New York, 1969).

4. P. Setlow, "Mechanisms which contribute to the long-term survival of spores of *Bacillus* species," J.Appl. Bacteriol **76**, 49S-60S (1994).

5. M. Paidhungat and P. Setlow, "Spore germination and outgrowth," in *Bacillus subtilis and Its Relatives: from Genes to Cells*, A. Hoch, R. Losick, and A. L. Shoneshein, eds. (J, Am, Soc, Microbiol, Washington DC, 2002), pp. 537 -548.

6. P. Setlow, in *Bacterial Stress Responses*, G. Storz and R. Hengee - Aronis, eds. (Am. Soc. Micorbiol, Washington, DC, 2000), pp. 217 -230.

7. A. Driks and P. Setlow, "Morphogenesis and properties of the bacterial spore," in *Prokaryotic Development*, Y. V. Brun and L. J. Shimkets, eds. (American Society for Microbiology, Washington D.C, 1999), pp. 191-218.

8. B. D. Church and H. Halvorson, " Dependence of the heat resistance of the bacterial endospores on their dipicolinic acid content," Nature **183**, 124-125 (1959).

9. G. W. Gould, "Germination," in *The Bacterial Spore,*, G. H. Gould and A. Hurst, eds. (Academic Press, New York, 1969), pp. 397-444.

10. T. A. Slieman and W. L. Nicholson, "Role of Dipicolinic Acid in Survival of Bacillus subtilis Spores Exposed to Artificial and Solar UV Radiation," Applied and Environmental Microbiology **67**(3), 1274-1279 (2001).

11. W. Volk and J. Brown, *Basic Microbiology* (Benjamin Cummings, New York, 1997).

12. M. P. Doyle, L. P. Beuchat, and T. J. Montville, in *Food Microbiology, Fundamentals and Frontiers*, M. P. Doyle, L. P. Beuchat, and T. J. Montville, eds. (American Society of Microbiology, Washington, DC, 1997).

13. R. I. Amann, W. Ludwig, and K. H. Schleifer, " Phylogenetic identification and in situ detection of individual microbial cells without cultivation.," in *Microbiol. Rev* (1995), pp. 143-169.

14. I. M. Head, J. R. Saunders, and R. W. Pickup, "Microbial evolution, diversity, and ecology: A decade of ribosomal RNA analysis of uncultured microorganisms," Microb. Ecol. **35**(1), 1-21 (1998).

15. P. Hugenholtz, B. M. Goebel, and N. R. Pace, "Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity," J. Bacteriol. **180**(18), 4765-4774 (1998).

16. R. I. Amann, J. Stromley, R. Devereux, R. Key, and D. A. Stahl, "Molecular and microscopic identification of sulfate-reducing bacteria in multispecies biofilms," Appl. Environ. Microbiol. **58**(2), 614-623. (1992).

17. C. M. O'Hara, F. C. Tenover, and J. M. Miller, "Parallel comparison of accuracy of API 20E, Vitek GNI, MicroScan Walk/Away Rapid ID, and Becton Dickinson Cobas Micro ID-E/NF for identification of members of the family Enterobacteriaceae and common gram-negative, non-glucose-fermenting *Bacilli*," J. Clin. Microbiol **31**(12), 3165-3169 (1993).

18. E. J. Vandamme, "The search for novel microbial fine chemicals, agrochemicals and biopharmaceuticals," J. Biotechnol **37**(2), 89-108 (1994).

19. D. Birnbaum, L. Herwaldt, D. E. Low, M. Noble, M. Pfaller, R. Sherertz, and A. W. Chow, "Efficacy of microbial identification system for epidemiologic typing of coagulase-negative staphylococci," J. Clin. Microbiol **32**(9), 2113-2119 (1994).

20. P. Vandamme, B. Pot, M. Gillis, P. de Vos, K. Kersters, and J. Swings " Polyphasic taxonomy, a consensus approach to bacterial systematics," Microbiol. Rev **60**(2), 407-438 (1996).

21. D. A. Relman, "Detection and identification of previously unrecognized microbial pathogens, ." Emerg. Infect. Dis **4**(3), 382-389 (1998).

22. M. O. Schrenk, K. J. Edwards, R. M. Goodman, R. J. Hamers, and J. F. Banfield, "Distribution of thiobacillus ferrooxidans and leptospirillum ferrooxidans: implications for generation of acid mine drainage,," Science **279**(5356), 1519-1522 (1998).

23. E. Dufour and A. Riaublanc, "Potentiality of spectroscopic methods for the characterisation of dairy products. Front-face fluorescence study of raw, heated and homogenised milks," LAIT **77**(6), 657-670 (1997).

24. S. Ammor, K. Yaakoubi, and I. Chevallier, "Identification by fluorescence spectroscopy of lactic acid bacteria isolated from a small-scale facility producing traditional dry sausages," Journal of Microbiological Methods **59**(2), 271-281 (2004).

25. P. Carmona, "Vibrational spectra and structure of crtstalline dipicolinic acid and calcium dipicolinate trihydrate," Spectrochim.Acta **36A**, 705-712 (1980).

26. R. Manoharan, E. Ghiamati, R. A. Dalterio, K. A. Britton, W. H. Nelson, and J. F. Sperry "UV resonance Raman spectra of bacteria, bacterial spores, protoplasts and calcium dipicolinate," J.Microbiology Methods **11**, 1-15 (1990).

27. W. H. Nelson and J. F. Sperry in *Modern Techniques in Rapid Microorganism Analysis*, VCH, ed. (New York, 1991).

28. E. Ghiamati, R. Manoharan, W. H. Nelson, and J. F. Sperry "UV Resonance Raman-Spectra of Bacillus Spores," Appl. Spect. **46**(2), 357-364 (1992).

29. R. Manoharan, E. Ghiamati, S. Chadna, W. H. Nelson, and J. F. Sperry "Effect of cultural conditions of deep UV resonance Raman spectra of bacteria," Appl. Spectrosc **47**, 2145 - 2150 (1993).

30. T. C. Cain, D. M. Lubman, and W. J. Weber, "Differentiation of Bacteria Using Protein Profiles from Matrix assisted Laser Desorption/Ionization Time of flight Mass Spectrometry," Rapid Comm.Mass Spectrum **8**, 1026-1030 (1994).

31. T. Krishnamurthy, P. L. Ross, and U. Rajamani, "Detection of Pathogenic and Non pathogenic Bacteria by Matrix-assisted Laser Desorption/Ionization Time of flight Mass Spectrometry," Rapid Comm.Mass Spectrum **10**, 883-888 (1996).

32. R. J. Arnold and J. P. Reilly, "Fingerprint Matching of *E.coli* Strains with Matrix assisted Laser Desorption/Ionization Time of flight Mass Spectrometry of Whole Cells Using a Modified Correlation Approach," Rapid Comm.Mass Spectrum **12**(630-636)(1998).

33. B. Hanesn, T. Leser, and N. Hendriksen, "Polymerase chain reaction assay for detection of *Bacillus cereus* group cells.," FEMS Microbiology Letters **202**, 209-213 (2001).

34. P. Fach, D. Hausser, J. P. Guillou, and M. R. Popoff, "Polymerase chain reaction for the rapid identification of *Clostridium botulinum* type A strains and detection in food samples," Journal of Applied Bacteriology **75**(3), 235-239 (1993).

35. M. N. Widjojoatmodjo, A. C. Fluit, and J. Verhoef, "Molecular identification of bacteria by fluorescence based PCR-single strand conformation

polymorphism analysis of the 16S rRNA gene.," Journal of Clinical Microbiology **33**, 2601-2606 (1995).

36. M. Fauville-Dufaux, N. Maes, E. Severin, C. Farin, E. Serruys, M. Struelens, N. Younes, J. P. Vincke, M. J. DeVos, and A. Bollen, "Rapid identification of Mycobacterium xenopi from bacterial colonies or "Bactec: culture by polymerase chain reaction and a luminescent sandwich hybridization assay," Research in Microbiology **146**, 349-356 (1995).

37. D. Champiat, A. Roux, O. Lhomme, and G. Nosenso, "Biohemiluminescence and biomedical applications," Cell Biology & Toxicology **10**(345-351)(1994).

38. E. T. Arakawa, N. V. Lavrik, and P. G. Datskos, "Detection of antrax simulants with microcalorimetric spectroscopy: *Bacillus substilis* and *Bacillus cereus* spores," Applied Optics **42**(10), 1757-1762 (2003).

39. J. T. Coburn, F. E. Lytle, and D. M. Huber, "Identification of Bacterial Pathogens by Laser Excited Fluorescence," Anal. Chem **57**(8), 1669-1673 (1985).

40. T. M. Rossi and I. M. Warner, "Bacterial Identification using fluorescence spectroscopy," in *Rapid Detection and Identification of Microorganisms*, W. H. Nelson, ed. (Verlag Chemie, 1985), pp. 1-50.

41. L. Reinisch, W. P. Van de Merwe, and B. V. Bronk, "Comparative fluorescence spectra from bacteria and spores in different stages of growth," Biophys, J **59**(161a)(1991).

42. B. V. Bronk and L. Reinisch, "Variability of Steady State Bacterial Fluorescence with Respect to Growth Conditions," Appl Spect **47**, 436-440 (1993).

43. J. A. Werkhaven, L. Reinisch, M. Sorrel, J. Tribble, and R. H. Ossoff, "Non-Invasive Optical Diagnosis of Bacteria Causing Otitis Media," Laryngoscope **104**, 264 - 268 (1994).

44. M. J. Sorrel, J. Tribble, L. Reinisch, J. A. Werkhaven, and R. H. Ossoff, "Bacteria Identification of Otitis Media with Fluorescence Spectroscopy.," Lasers Surg. Med **14**, 155-163 (1994).

45. S. A. Glazier and H. H. Weetall, "Autofluorscence detection of *Escherichia coli* on silver membrane filters," J. Microbiol Methods **20**, 23-27 (1994).

46. F. S. Ligler, G. P. Anderson, P. T. Davidson, R. J. Foch, J. T. Ives, K. D. King, G. Page, D. A. Stenger, and J. P. Whelan, "Remote sensing using an airborne biosensor,." Environ. Sci. Technol **32**, 2461 (1998).

47. G. A. Luoma, P. P. Cherrier, and L. A. Retfalvi, "Real-time warning of biological-agent attacks with the Canadian Integrated Biochemical Agent Detection System II (CIBADS II)," Field Anal. Chem. Technol **3**, 260-273 (1999).

48. B. C. Spector, L. Reinisch, D. Smith, and J. A. Werkhaven, "Noninvasive fluorescent identification of bacteria causing acute otitis media in a chinchilla model.," Laryngoscope **110**, 1119-1123 (2000).

49. E. Yacoub-George, L. Meixner, W. Scheithauer, A. Koppi, S. Drost, H. Wolf, C. Danapel, and K. A. Feller, "Chemiluminescence multichannel immunosensor for biodetection," Anal. Chim. Acta **Nr.457**, S 3-12 (2002).

50. J.-K. Li, E. C. Asali, A. E. Humphery, and J. J. Horvath, "Monitoring cell concentration and activity by multiple excitation fluorometry," Biotechnol.Prog **7**, 21-27 (1991).

51. I. M. Warner, G. D. Christian, and E. R. Davidson, "Analysis of Multicomponent Fluorescence Data," Analytical Chemistry **49**(4), 564-573 (1977).

52. B. V. Bronk , W. P. Van de Merwe, and M. Stanley, "An in-vivo measure of average bacterial cell size from polarised light scattering function," Cytometry **13**, 155-162 (1992).

53. P. C. Gray, I. R. Shokair, S. E. Rosenthal, G. C. Tisone, J. S. Wagner, L. D. Rigdon, G. R. Siragusa, and R. J. Heinen, "Distinguishabilty of Biological Material by use of Ultraviolet Multispectral Fluorescence," Applied Optics **37**, 6037-6041 (1998).

54. Y. S. Cheng, E. B. Barr, B. J. Fan, P. J. Hargis, D. J. Rader, T. J. O'Hern, J. R. Torczynski, G. C. Tisone, B. L. Preeppernau, S. A. Young, and R. J. Raddoff, "Detection of Bioareosols Using Multiwavelength UV Fluorescence Spectroscopy," Aerosol Sci. Technol. **30**(3), 186-201 (1999).

55. S. C. Hill, R. G. Pinnick, S. Niles, Y. L. Pan, S. Holler, R. K. Chang, J. Bottiger, B. T. Chen, C. S. Orr, and G. Feather, "Real-time Measurement of Fluorescence Spectra from Single Airborne Particles," Field Anal. Chem. Technol. **3**(4-5), 221-239 (1999).

56. V. Sivaprakasham, A. Huston, C. Scotto, and J. D. Eversole, "Multiple UV wavelength excitation and fluorescence of bioaerosols," Optics Express **12**, 4457-4466 (2004).

57. M. O. Scully, G. W. Kattawar, and R. P. Lucht, "FAST CARS: Engineering a laser spectroscopic technique for rapid identification of bacterial spores," Proceedings of the National Academy of Sciences of the United States of America **99**(17), 10994-11001 (2002).

58. P. P. Hairston, J. Ho, and F. R. Quant, "Design of an Instrument for real time detection of areosols using simultaneous measurements of particle aerodynamics size and intrinisic fluorescence," J. Aerosol Sci **28**, 471-482 (1997).

59. J. D. Eversole, J. J. Hardgrove, W. K. Cary Jr, D. P. Choulas, and M. Seaver, "Continuous, Rapid Biological Aerosol Detection with the Use of UV Fluorescence: Outdoor Test Results," Field Anal. Chem. Tech. **3**, 249-259 (1999).

60. M. Seaver, J. D. Eversole, J. J. Hardgrove, W. K. Cary Jr, and D. C. Roselle, "Size and fluorescence measurements for field detection of bioaerosols," Aerosol Sci. Technol **30**, 174-185 (1999).

61. J. D. Eversole, W. K. Cary Jr, C. S. Scotto, R. Pierson, M. Spense, and A. J. Campillo, "Continuous bioaerosol monitoring using UV excitation fluorescence: Outdoor test results," Field Anal. Chem. Technol. **15**, 205-212 (2001).

62. J. D. Hybl, G. A. Lithgow, and S. G. Buckley, "Laser-Induced Breakdown Spectroscopy Detection and Classification of Biological Aerosols," Applied Spectroscopy **57**(10), 1207-1215 (2003).

63. A. L. Samuels, K. L. DeLucia , K. L. McNesby , and A. Miziolek "Laser-induced breakdown spectroscopy of bacterial spores, molds, pollens and protein: initial studies of discrimination potential," Applied Optics **42**, 6205-6209 (2003).

64. G. W. Faris, R. A. Copeland, K. Mortelmans, and B. V. Bronk "Spectrally resolved absolute fluorescence cross sections for *Bacillus* spores. ." Applied Optics **36**(4), 959-967 (1997).

65. J. R. Stephens, "Measurements of the ultraviolet fluorescence cross sections and spectra of *Bacillus anthracis* simulants," (Los Alamos National Laboratory, Los Alamos, NM, 1999).

66. W. F. Cheong and A. J. Welch, "A Review of the Optical Properties of Tissues," IEEE.J.Quantum Electron **26**, 2166-2185 (1990).

67. R. Bellman, *Adaptive control processes: A Guided tour* (Princeton University Press, Princeton, 1961).

68. M. A. Sharaf, D. L. llman, and B. R. Kowalski, *Chemometrics* (Wiley, New York, 1986).

69. E. R. Malinowiski and D. G. Howery, *Factor analysis in Chemistry* (John Wiley, New York, 1980).

70. N. Bonnet, "Multivariate statistical methods for analysis of microscope image series," Journal of Microscopy **190**, 2-18 (1998).

71. Y.-P. Sun, J. Sears, D.F, and J. Saltiel, "Three Component Self-Modelling Technique Applied to Luminance Spectra," Anal. Chem **59**(20), 2515 - 2519 (1987).

72. J. Saltiel, J. Sears, D. F, J.-O. Choi, Y.-P. Sun, and D. W. Eaker, "The Fluorescence, Fluorescence-Excitation and UV Absorption Spectra of trans-1-(2-Naphthyl)-2-phenylethene Conformers," J. Phys. Chem **98**, 35-46 (1994).

73. B. Vandeginste, R. Essers, T. Bosman, J. Reijnen, and G. Kateman, " Three-component curve resolution in liquid chromatography with multiwavelength diode array detection," Anal. Chem **57**, 971-985 (1985).

74. B. Hessling, G. Souvignier, and K. Gerwert, "A model-independent approach to assigning bacteriorhodopsin's intramolecular reactions to photocycle intermediates," Biophys. J **65**, 1929 -1941 (1993).

75. L. Leblanc and E. Dufour, "Monitoring the identity of bacteria using their intrinsic fluorescence," FEMS Microbiology Letters **211**(2), 147-153 (2002).

76. B. S. Everitt and G. Dunn, *Applied Multivariate Data Analysis* (Oxford University Press, New York, 1992).

77. A. Basilevsky, *Statistical Factor Analysis and related Methods, Theory and applications* (John Wiley & Sons, Inc, New York, 1994).

78. D. L. Massart, B. G. Vandeginste, S. N. Deming, Y. Michotte, and L. Kaufman, *Chemometrics: A Textbook* (Elsevier, Amasterdam, 1988).

79. I. Morell, E. Gimenez, and V. Esteller, "Application of principal components to the study of salinisation on the Castellon Plain(spain)," The Science of the Total Environment **177**, 161-171 (1996).

80. R. P. Ashley and J. W. Lyoyd, "An example of the use of factor analysis and cluster analysis in groundwater chemistry interpretation," Journal of Hydrology **39**, 355-364 (1978).

81. N. Bonnet, "Artificial intelligence and pattern recognition techiniques in microscope image processing and analysis.," Advances in imaging and Electron physiscs **114**, 1-77 (2000).

82. M. B. Eisen, P. T. Spellman, P. O. Brown, and D. Botstein, "Cluster analysis and display of genome-wide expression patterns," Proc. Natl. Acad. Sci. **95**, 14863-14868 (1998).

83. P. Hansen and B. Jaumard, " Cluster analysis and mathematical programming," Mathematical programming **79**, 191-215 (1997).

84. A. Guenoche, "Partitions optimises selon differents criteres: evaluation et Comparison," Mathematiques, Informatique et Sciences humaines **161**(2003).

85. P. Hansen, B. Jaumard , and M. Mlandenovic, "How to choose K entities among N," DIMAC Series in Discrete Mathematicas and Theortical Computer Science. **19**, 109-115 (1995).

86. R. Weichert, W. Klemm, K. Legenhausen, and C. Pawellek, "Determination of Fluorescence Cross-Sections of Biological Aerosols," Particle & Particle Syststems Characterization **19**(3), 216-222 (2002).

87. W. Saliga, R. Burnham, T. Deely, W. Gavert, M. Pronkao, G. Verdun, and H. Verdun, "Short Range Biological Standoff Detector System (SR-BSDS)," Proc.Soc.Photo-Opt. Instrum.Eng **3855**, 72-81 (1998).

88. K. L. Schroder, P. J. Hargis, R. L. Schmitt, D. J. Rader, and I. R. Shokair, "Development of an Unattended Ground Sensor for Ultraviolet Laser Induced

Fluorescence Detection of Biological Agent Aerosols," Proc.Soc. Photo-Opt. Instrum.Eng. **3855**, 82-91 (1998).

89. D. C. Shelly, J. M. Quarles, and I. M. Warner, "Identification of fluorescent Pseudomonas species," Clin. Chem **26**, 1127-1132 (1960).

90. J. K. Li, E. C. Asali, A. E. Humphery, and J. J. Horvath, "Monitoring Cell Contentration and Activity by Multiple Excitation Fluorometry," Biotech.Prog **7**, 21-27 (1991).

91. J. R. Lakowicz, *Principles of Fluorescence Spectroscopy* (Plenum Press, New York., 1999).

92. R. A. Dalterio, W. H. Nelson, D. Britt, J. F. Tanguay, and S. L. Suib, "Steady state and decay characteristic of protein tryptophan fluorescence from live bacteria," Appl. Spec **40**, 86-90 (1986).

93. R. A. Dalterio, W. H. Nelson, D. Britt, J. F. Sperry, J. F. Tanguay, and S. L. Suib, "The steady state and decay characteristics of primary from live bacteria.." Appl. Spec **41**, 234-241 (1987).

94. I. Munro, I. Pecht, and L. Stryer, "Subnanosecond motions of tryptophan residues in proteins," Proc. Natl. Acad. Sci.USA **76**, 56-60 (1979).

95. I. D. Campbell and R. A. Dwek, *Biological Spectroscopy* (Benjamin/Cummings Publishing Co, Menlo park, CA, 1984).

96. J. B. Birks, *Photophysics of Aromatic Molecules* (John Wiley & Sons, New York, 1970).

97. M. Eftink, "Fluorescence Techniques for Studying Protein Structure," in *Protein Structure Determination: Methods of Biochemical Analysis*, H. Clarence and Suelter, eds. (John Wiley and Sons, Inc., New York, 1990).

98. J. M. Beechem and L. Brand, "Time resolved fluorescence of proteins," Annu. Rev. Biochem **54**, 43-71 (1985).

99. D. L. Massart and L. Kaufman, *The Interpretation of Analytical Chemical Data by the use of Cluster Analysis* (New York, 1983).

100. J. Murrell, *Theory of Electronic Spectra of Organic Molecules* (Methuen, London, 1963).

101. N. J. Turro, *Molecular Photochemistry* (W.A. Benjamin, New York, 1965).

102. C. A. Parker, *Photoluminescence of Solutions with Appplications to Photochemistry and Analytical Chemistry* (Elsevier, Amsterdam, 1968).

103. J. G. Calvert and J. N. Calvert, Jr, *Photochemistry* (John Wiley & Sons, New York, 1966).

104. R. Richards-Kortum and E. Sevick-Muraca, "Quantitative Optical Spectroscopy for Tissue Diagnosis," Annu. Rev. Phys.Chem. **47**, 555-606 (1996).

105. M. B. Silberberg, H. E. Savage, G. C. Tang, P. G. Sacks, R. R. Alfano, and S. P. Schantz, "Detection of Retinoic Acid Induced Biochemical Alterations in Squamous Cell Carcinoma Using Intrinsic Fluorescnce Specroscopy," Laryngoscopy **104**(3), 278-282 (1994).

106. L. Stryer, *Biochemistry* (W.H Freeman and Company, New York, 1988).

107. T. C. Creighton, *Proteins, Structures and Molecular Properties* (W.H Freeman and Company, New York, 1993).

108. F. W. J. Teale and G. Weber, "Ultraviolet fluorescence of the aromatic amino acids," Biochem. J **67**, 476-482 (1957).

109. Chou-pong, P. G. Patonay, C. W. Moss, D. Hollis, G. M. Carlone, B. D. Plikaytis , and I. M. Warner, "Comparison of Flavobacterium and Shpingobacterium species by enzyme profiles, with use of pattern recognition of two dimensional fluorescence data," Clin. Chem **33**, 377-380 (1987).

110. M. Brunori, "The structural dynamics of myoglobin," Biophys. Chem. **86**, 221-230 (2000).

111. S. Dhami, A. J. de Mello, G. Rumples, S. M. Bisshop, D. Philips, and A. Beeby, "Phthalocymine fluorescence at high concentration: dimmers or reabsorption effect?," Photochem. photobiol. **61**(4), 341-346 (1995).

112. A. T. R. Williams, S. A. Winfield, and J. N. Miller, "Relative fluorescence quantum yields using a computer controlled fluorescence spectrometer," Analyst **108**, 1067 (1983).

113. P. C. Jurs and T. C. Isenhour, *Chemical Application of Pattern Recognition* (John Wiley, New York, 1975).

114. C. Ho, N. G. D. Christian, and E. R. Davidson, " Application of the method of rank annihilation to quantitative analyses of multicomponet fluorescence data from video fluorometer," Anal.Chem. **50**, 1108-1113 (1978).

115. M. Mccue and E. R. Malinowiski, "Rank Annihilation Factor Analysis of Unresolved LC Peaks," J.Chromatogr.Sci. **21**(5), 229-234 (1983).

116. A. Lorber, "Quantifying chemical composition from two-dimensional data arrays," Anal. Chim. Acta **164**, 293-297 (1984).

117. R. D. Jiji, G. A. Cooper, and K. S. Booksh, " Excitation-Emission Matrix Fluorescence Based Determination of Carbamate Pesticides and Polycyclic Aromatic Hydrocarbons," Anal. Chim. Acta **397**, 61-72 (1999).

118. W. Chen, P. Westerhoff, J. A. Leenheer, and K. Booksh, "Fluorescence excitation-emission matrix regional integration to quantify spectra for dissolved organic matter," Environ. Sci. Technol. **37**(24), 5701-5710 (2003).

119. I. M. Warner, J. B. Callis, E. R. Davidson, and G. D. Christian, "Multicomponent analysis in clinical chemistry by use of rapid scanning fluorescence spectroscopy," Clin. Chem **22**, 1483-1492 (1976).

120. D. W. Johnson, J. B. Callis, and G. D. Christian, "Rapid Scanning fluorescence Spectroscopy," Anal. Chem. **49**(8), A747 (1977).

121. J. W. Longworth, *Luminescence of polypeptide and proteins*, Excited states of proteins and nucleic acids (Plenum, New York, 1971), pp. 319-487.

122. L. P. Giering, "Multicomponent Analysis of Mixtures," Industrial Research & Development **20**(9), 134-140 (1978).

123. G. R. Haugen, B. A. Raby, and L. P. Rigdon, "Online Computer Control of Luminescent Spectrometer with Off- Line Computer Data Processing and Display," Chem.Inst. **6**(3), 205-225 (1975).

124. A. P. Bentz, "Oil-Spil identification," Anal. Chem **48**(6), A455 (1976).

125. F. F. Hartline, "Three Dimensional Fuorescence Spectroscopy," Science **203**(4387), 1330-1331 (1979).

126. D. W. Johnson, J. A. Gladder, J. B. Callis, and G. D. Christian, "Video Fluorometer," Rev.Sci.Instrum. **50**, 118-126 (1979).

127. I. M. Warner, M. P. Forgarty, and D. C. Shelly, "Design considerations for a 2-Dimensional Rapid Scanning fluorometer," Anal.Chim.Acta **109**(2), 361-372 (1979).

128. I. M. Warner, *Contemporary Totpics in Analytical and Clinical Chemistry* (Plenum Publishing Corp., New York, 1982), Vol. 4.

129. J. C. Wright and M. J. Wirth, "Lasers and Spectroscopy," Anal. Chem **52**, 988A (1980).

130. K. Pearson, "On lines and planes of closet fit to system of points in space," Phili. Mag **6**(2), 559-572 (1901).

131. H. Hotelling, "Analysis of a complex of Statistical variable into principal components," J.Educ.Psysch **24**, 417-441 (1933).

132. S. K. Jenson and F. A. Walty, "Principal components analysis and canonical analysis in remote sensing," presented at the Proc. An. Soc. of Photogrammetry, 1979.

133. W. W. Cooley and P. R. Lohnes, *Multivariate data Analysis* (John Wiley & Sons, Inc, New York, 1971).

134. J. C. Davis, *Statistics and data analysis in geology,* 2nd ed. (Wiley, New York, 1986).

135. B. F. J. Manly, *Multivariate statistical methods:a primer,* 2nd ed. (Chapman and Hall, London, 1994).

136. I. H. Berstein, C. P. Garbin, and G. K. Teng, *Applied multivariate analysis* (Springer-Verlag, New York, 1988).

137. L. Kaufman and P. J. Rousseeuw, *Finding Groups in Data. An Introduction to Cluster Analysis* (Willey-Interscience, New York, 1990).

138. I. B. Berlman, *Handbook of Fluorescence Spectra of Aromatic molecules,* 2nd ed ed. (Accademic Press, New York and London, 1971).

139. W. R. Dawson and M. W. Windsor, "Fluorescence yields of aromatic compounds," J. Phys. Chem. **72**, 3251-3260 (1968).

140. D. F. Eaton, "Reference materials for fluorescence measurement," Pure Appl. Chem. **60**, 1107-1114 (1988).

141. J. Kunnil, B. Swartz, and L. Reinisch, "Changes in the Luminescence Between Dried and Wet *Bacillus* Spores," Applied Optics **43**, 5404-5409 (2004).

142. S. Sarasanandarajah, J. Kunnil, E. Chacko, B. V. Bronk, and L. Reinisch, "Reversible changes in fluorescence of bacterial endospores found in aerosols due to hydration/drying," J.Aerosol Sci. (2005).

143. D. B. Wetlaufer, "Ultraviolet spectra of proteins and amino acids," Adv.Protein Chem **17**, 303-390 (1962).

144. R. J. Robbins, G. R. Fleming, G. S. Beddard, G. W. Robinson, P. J. Thistlethwaite, and G. J. Woolfe, "Photophysics of aqueous tryptophan: pH and temperature effects," J. Am. Chem. Soc. **102**, 6271–6279 (1980).

145. G. J. Smith and W. H. Melhuish, "Effect of the temperature and viscosity on the excited singlet state of indoles in polar solvents," J. Phys. Chem. **95**, 4288–4291 (1991).

146. C. Saavedra, C. Vasquez, and M. V. Encinas, "Structural studies of the *Bst* V1 restriction-modification proteins by fluorescence spectrscopy," Eur.J.Biochem. **263**, 65-70 (1999).

147. P. Marco, J. L. Terrance, E. W. Katherine, and J. M. Alexander, "The high-resolution architecture and structural dynamics of *Bacillus*," Biophys J BioFAST (doi:10.1529/biophysj.104.049312)(2004).

148. A. J. Westphal, P. B. Price, T. J. Leighton, and K. E. Wheeler, " Kinetics of size changes of individual Bacillus thuringiensis spores in response to changes in relative humidity," PNAS **100**, 3461-3466 (2003).

149. Y. Cao, D. W. Williams, and N. E. Williams, "Data transformation and standardizationin multivariaten analysis of river water quality.," Ecological Applications **9**(2), 669-677 (1999).

150. H. C. Romesburg, *Cluster analysis for researchers* (Lifetime Learning Publications, Belmont, CA, 1984).

151. B. R. Kowalski and C. F. Bender, "Pattern Recognition. A Powerful Approach to Interpreting Chemical Data," J.Am.Chem.Soc. **94**, 5632-5639 (1972).

152. J. Kunnil, S. Sarasanandarajah, E. Chacko, B. Swartz, and L. Reinisch, " Identification of *Bacillus* Spores Using Clustering of Principal Components of Fluorescence Data," (2005).

153. H. F. Kaiser, "The application of electronic computers to factor analysis," Educational and Psychological Measurement **20**(141-151)(1960).

154. K. Y. Yeung and W. L. Ruzzo, "Principal component analysis for clustering gene expression data," Bioinformatics **17**, 763-774 (2001).

155. C. Estes, A. Duncan, B. Wade, C. Lloyd, W. Ellis Jr, and L. Powers, "Reagentless detection of microorganisms by intrinsic fluorescence," Biosens.Bioelectron. **18**, 511-519 (2003).

156. R. Neihof, J. K. Thomson, and V. R. Deitz, "Sorption of water vapour and nitrogen gas by bacterial spores," Nature **216**, 1304-1306 (1967).

157. B. J. Marshall and W. G. Murrell, "Symposium on bacterial spores: IX. Biophysical Analysis of the endospore.," Appl. Bacteriol **33**, 103-129 (1970).

158. Z. Wasylewski and M. R. Eftink, "Frequency-domain fluorescence studies of an extracellular metalloproteinase of *Staphylococcus aureus*," Biochim Biophys Acta **13**, 331-341 (1987).

159. V. N. Petushkov, B. G. Gibson, and J. Lee, "Direct measurement of excitation transfer in the protein complex of bacterial luciferace hydroxyflavin and the associated yellow fluorescence proteins from *Vibrio fischeri* Y1," Biochemistry **35**, 8413-8418 (1996).

160. B. D. Flury, "Developments in principal component analysis," in *Descriptive multivariate analysis.*, J. W. Krzanowski, ed. (Clarendon Press, Oxford, United Kingdom, 1994), pp. 14-33.

161. E. Holmes, A. W. Nicholls, J. C. Lindon, S. Ramos, M. Spraul, P. Neidig, S. C. Connor, J. Connelly, S. J. P. Damment, J. Haselden, and J. K. Nicholson, "Development of a model for classification of toxin-induced lesions using 1H

NMR spectroscopy of urine combined with pattern recognition," NMR Biomed. **11**(235-244)(1998).

162. K. Maquelin, L. P. Choo-Smith, T. van Vreeswijk, H. P. Endtz, B. Smith, R. Bennett, H. A. Bruining, and G. J. Puppels, "Raman spectroscopic method for identification of clinically relevant microorganisms growing on solid culture medium," Anal. Chem. **72**, 12-19 (2000).

163. K. Maquelin , L. P. Choo-Smith, H. P. Endtz, H. A. Bruining, and G. J. Puppels, "Rapid identification of Candida species by confocal Raman microspectroscopy," J Clin. Microbiol. **40**(2), 594-600 (2002).

164. J. O. Klein, "Epidemiology of Otitis Media," Pediatr Infec Dis J. **8:S9**(1989).

165. C. o. t. I. o. M. a. t. N. R. Council, in *Chemical and Biological Terrorism: Research and Development to Improve Civilian Medical Response* (National Academy Press, Washington, DC, 1999).

Chapter 7

APPENDIX

7.1 List of Publications

- J. Kunnil, B. Swartz, L. Reinisch. *Changes in the Luminescence between Dried and Wet Bacillus Spores*. Applied Optics, Volume 43, Issue 28, p.5404- 5409 (2004).
- J. Kunnil, S. Sarasanandarajah, E. Chacko, B. Swartz, L. Reinisch. *Identification of Bacillus Spores Using Clustering of Principal Components of Fluorescence Data.* Aerosol Science and Technology, Volume 39, Issue 9, p842-848 (2005).
- J. Kunnil, S. Sarasanandarajah, E. Chacko, L. Reinisch. *Fluorescence Quantum Efficiency of Bacillus Spores*. Optics Express, Volume 13, No 22 (2005).
- J. Kunnil, S. Sarasanandarajah, E. Chacko, B. Swartz, L. Reinisch. *Effect of Washing on the Identification of Bacillus spores by principal Components Analysis of Fluorescence Data* (Submitted to Applied Optics, 2005).
- L. Reinisch, J. Kunnil and S. Sarasanandarajah. *Anthrax Detector Development.* New Zealand Science Review, Volume 62, Issue 1-2, p23-25(2005).
- S. Sarasanandarajah, J. Kunnil, B. V. Bronk, Lou Reinisch. *Two Dimensional Multi Wavelength Fluorescence Spectra of Dipicolinic Acid and Calcium Dipicolinate*. Applied Optics, Volume 44, Issue 7, 1182-1187 (2005)
- S. Sarasanandarajah, J. Kunnil, E. Chacko, B. V. Bronk, L. Reinisch. *Reversible changes in fluorescence of bacterial endospores found in aerosols due to hydration/drying*. J.Aerosol Sci. Volume 36, Issue 5-6, p689-699(2005).

- J. Kunnil, S. Sarasanandarajah, E. Chacko, L. Reinisch. Detection of Bacillus Spores based on Stability and Influence in Clustering of Fluorescence data Using Principal Components Analysis (In Preparation).
- J. Kunnil, S. Sarasanandarajah, E. Chacko, L. Reinisch. Solvent Effect on Autofluorescence of Bacillus Spores: A Comparative Study of Dry Spores and Spores in Aqueous, Glycerol and Ethanol Suspensions (In Preparation).
- J. Kunnil, S. Sarasanandarajah, E. Chacko, B. V. Bronk, L. Reinisch. *Auto fluorescence of Bacillus globigii: A Comparative study of dry spores and spores in aqueous and ethanol suspensions.* The Pittsburgh Conference (pittcon) on Analytical Chemistry and Applied Spectroscopy, Orlando, FL,USA, Feb 27 –March 4, 2005.
- S. Sarasanandarajah, J. Kunnil, E. Chacko, B. V. Bronk, L. Reinisch. *Reversible effect of Hydration on Bacillus globigii probed with Fluorescence Spectroscopy*. The Pittsburgh Conference (pittcon) on Analytical Chemistry and Applied Spectroscopy, Orlando, FL, USA, Feb 27 –March 4, 2005.
- L. Reinisch, E.Chacko, J. Kunnil, S. Sarasanandarajah. *The Robust Nature of Fluorescence Identification of Bacterial Spores.* EPSM 29th Annual Conference, Australia, 23rd -27th October, 2005.

7.2 *Mathematica* Program for Smoothing the Spectrum

```
ResetDirectory[];
SetDirectory["C:\Joseph U\NEW BG FL FOR QE\1BG NEW FL/"];
f = OpenRead["BG18.PRN"];
dd = ReadList[f, {Real}];
Close[f];
ld = Length[dd];
d = Table[{i, dd[[i, 1]]}, {i, 20, ld}];
{f1a, f1b} = Transpose[d];
lf1 = Length[f1a];
fmin = Min[f1b];
f1b = f1b - fmin;
f1a = (((280 + 10) + f1a) / 280);
f1 = Transpose[{f1a, f1b}];
tl = Table[Exp[-((x * (lf1 / 4)) - i) ^2], {i, First[f1a] * (lf1 / 4), Last[f1a] * (lf1 / 4)}];
ff1 = Fit[f1, tl, x];
ListPlot[f1]
Plot[ff1, {x, 1, 2}]
Show[%, %%]
```

7.3 *Mathematica* Program for Plotting 2D Fluorescence Fingerprints

```
ResetDirectory[];
SetDirectory["C:\Joseph U\NEW BG FL FOR QE\1BG NEW FL/"];
readit[file_String] := Module[{f, d},
f = OpenRead[file];
dd = ReadList[f, {Real}];
Close[f];
ld = Length[dd];
d = Table[{i, dd[[i, 1]]}, {i, 20, ld}]
];
doit[file_String, lambda_] := Module[{f1, lf1, f1a, f1b, fmin, tl, i, ff1, x},
{f1a, f1b} = Transpose[readit[file]];
lf1 = Length[f1a];
fmin = Min[f1b];
f1b = f1b - fmin;
f1a = (((lambda + 10) + f1a) / lambda) - 1;
f1 = Transpose[{f1a, f1b}];
tl = Table[Exp[-((x * (lf1 / 4)) - i) ^ 2], {i, First[f1a] * (lf1 / 4), Last[f1a] * (lf1 / 4)}];
ff1 = Fit[f1, tl, x];
N[Table[ff1 /. x -> i / 42, {i, 1, 41}]]
];

files = {
"BG10.PRN",
"BG11.PRN",
"BG12.PRN",
"BG13.PRN",
"BG14.PRN",
"BG15.PRN",
"BG16.PRN",
"BG17.PRN",
"BG18.PRN",
"BG19.PRN",
"BG20.PRN",
"BG21.PRN",
"BG22.PRN",
"BG23.PRN",
"BG24.PRN",
"BG25.PRN",
"BG26.PRN",
"BG27.PRN",
"BG28.PRN",
"BG29.PRN",
"BG30.PRN",
"BG31.PRN",
"BG32.PRN",
"BG33.PRN",
"BG34.PRN",
"BG35.PRN",
```

```
"BG36.PRN",
"BG37.PRN",
"BG38.PRN",
"BG39.PRN",
"BG40.PRN",
"BG41.PRN",
"BG42.PRN",
"BG43.PRN",
"BG44.PRN",
"BG45.PRN",
"BG46.PRN",
"BG47.PRN",
"BG48.PRN",
"BG49.PRN",
"BG50.PRN"};
waves = Table[200 + 10 * i, {i, 0, 40}];

q0 = Thread[doit[files, waves]];
Min[q0]
Max[q0]

ListContourPlot[q0, Contours -> 15, PlotRange -> All, ColorFunction -> (Hue[2 #/3] &)]
```

7.4 Principal Components Model

The major step in PCA is the extraction of the eigenvectors from variance-covariance matrix to get uncorrelated new variables called principal components (PCs). Many research projects gather a large number of variables; we would like some objective means of reducing the variables to a manageable number. One of the primary purposes of PCA is to reduce the number of variables. The logic of PCA is as follows. We would like to reduce the number of variables but preserve as much of the variability in the data as possible. To do this, let us create a new variable from the data that is a linear combination of the original variables. Let A denote this new variable and let x_1, x_2, ... x_n denote the original variables. By making A a linear combination of the original variables, we imply the mathematical formula

$$A = a_1x_1 + a_2x_2 + ... a_nx_n \tag{7.0}$$

where a_1, a_2, ... a_n are weights assigned to the original n variables. To preserve as much variability as possible, we want to make the variance of A as large as possible. Thus, the goal of PCA is to select values of the a's so that the variance of A is as large as possible. The variance of A will begin approaching infinity as the individual a's go to either plus or minus infinity. To avoid this, PCA selects the a's subject to the constraint that the sum of the squared a's equals unity. Technically, this is called "normalizing" a vector, the vector in this case being the vector of a's. Once the weights are calculated, and then the variable A is called the first principal component or first PC.

To see how the weights are identified, let C denote the covariance or correlation matrix for the xs. The variance of A then equals $\mathbf{a}^T\mathbf{C}\mathbf{a}$ where $\mathbf{a}$ is the vector of a weights. We want to maximize the variance of A subject to the constraint that $\mathbf{a}^T\mathbf{a} = 1$. Thus, we can construct the augmented Lagrangian function

$$F(\mathbf{a}) = \mathbf{a}^T\mathbf{C}\mathbf{a} + \lambda(\mathbf{a}^T\mathbf{a} - 1) \tag{7.1}$$

where λ is the Langrangian multiplier. We want to maximize $F(\mathbf{a})$ with respect to vector $\mathbf{a}$. We may do this by taking the first derivative of $F(\mathbf{a})$ with respect to $\mathbf{a}$ and setting it to a vector of 0s.Doing this gives

$$\partial F(a)/\partial a = 2Ca - \lambda a = 0 \quad (7.2)$$

or

$$(C - \lambda I)\,\mathbf{a} = 0. \quad (7.3)$$

One solution to (7.3) is to set all the *a*s equal to zero. This, however, is called a *trivial* solution. We require a nontrivial solution or, in other words, substantive values for the elements of **a**. To solve for this, suppose for the moment that we have some real value for **a** such that the matrix (C - λ I) is known and has an inverse. Premultiplying both sides of (7.3) by the inverse of (C - λ I) gives

$$\mathbf{a} = (C - \lambda I)^{-1} x\ 0 \quad (7.4)$$

or a must equal 0. Hence, if matrix (C - λ I) has an inverse, the only solution to (7.3) is the trivial solution. We must conclude then that in order to have a nontrivial solution, the matrix
(C - λ I) must NOT have an inverse. That is, (C - λ I) must be singular and its determinant must equal 0,
or

$$|C - \lambda I| = 0 \quad (7.5)$$

Equation (7.5) is called the *characteristic equation* for matrix C and that, because C is symmetric, there will be *n* real values of λ. This *n* λ's are the *eigenvalues* of C and for any given eigenvalue, the solution of (7.3) is that *a* is an *eigenvector* of C and equation (7.3) rewrite I as

$$C\mathbf{a} = \lambda\,\mathbf{a} \quad (7.6)$$

Premultiply both sides by $\mathbf{a}^T$ giving

$$\mathbf{a}^T C\mathbf{a} = \mathbf{a}^T \lambda\,\mathbf{a} = \lambda\,\mathbf{a}^T\mathbf{a} \quad (7.7)$$

But $\mathbf{a}^T C\mathbf{a}$ is the variance of variable *A* and we have constrained $\mathbf{a}^T\mathbf{a}$ to equal unity. Consequently, (7.7) implies that

$$\lambda = \text{Var}(A) \tag{7.8}$$

In other words, *the variance of a principal component is actually an eigenvalue of the matrix C.* Hence, to maximize Var (*A*), all we have to do is select the largest eigenvalue of C and take its associated eigenvector as the weights in vector a.

The second PC may be defined as that linear combination of the *x*s that has the second largest variance, or

$$A_2 = a_{12}x_1 + a_{22}x_2 + \ldots + a_{n2}x_n \tag{7.9}$$

Here, we subscript the *A* and double subscript the *a*s so that a_{ij} is the weight assigned to the i^{th} variable for the j^{th} PC. Since Var (A_2) is an eigenvalue of C, we simply select the second largest eigenvalue and its associated eigenvector as the solution to $\mathbf{a}_2$. We may continue with this logic, each time selecting the next highest eigenvalue and its associated eigenvector as the next PC. A consequence of using the eigenvalues as the variance for the PCs and the eigenvectors as the weights is that the PCs will be uncorrelated. Recall that the matrix of eigenvectors is an *orthogonal* or *orthonormal* matrix. That is, if **M** is the matrix of eigenvectors, then $\mathbf{M}^T\mathbf{M} = \mathrm{I}$. We may write the equation for the PCs in matrix form as

$$\mathbf{z} = \mathrm{M}\mathbf{x} \tag{7.10}$$

where **z** is a *n* by 1 vector of PCs and **x** is a *n* by 1 vector of variables. Premultipling both sides by $\mathbf{M}^T$, giving

$$\mathbf{M}^T\mathbf{z} = \mathbf{M}^T\mathbf{M}\mathbf{x} = \mathbf{I}\mathbf{x} = \mathbf{x}, \tag{7.11}$$

reveals that the variables may be written as a linear combination of the PCs. Now take the covariance matrix of **x** in (7.11),

$$\mathbf{M}^T C_{zz}\,\mathbf{M} = \mathrm{C} \tag{7.12}$$

where C_{zz} is the covariance matrix among the PCs. But $\mathrm{C} = \mathbf{M}^T \lambda\, \mathbf{M}$ where λ is the diagonal matrix of eigenvalues. Hence, the covariance matrix among the PCs is a diagonal matrix, making all the components uncorrelated with one another.

7.5 The R Program for PCA and Cluster analysis

The R program read the fluorescence data, formed a data matrix, smoothed, calculated and formed the correlation matrix, calculated the eigenvectors and eigen values and plot the PCA plot.

For example, PCA of three *Bacillus* samples: *B. globigii* (i), *B. cereus* and *B. popilliae* measures three times each from three different dust sources. Also show the cluster analysis program for hierarchical and partition around medoids clustering

```
readat <-
 function (filedir,filepat="BG",startno=0,ext="asp",nrows=2000,nfiles=41,step=5){
## first form the names of the files
 fnames<-array(,nfiles) ## empty array to take the file names
 for(i in 1:nfiles) { ## loop to form the file names
 extn <- paste(".",ext,sep="")
 fnam <- paste(paste(filepat,startno+1+(i-1)*step,sep=""),extn,sep="")
 fnames[i]<- paste(filedir,fnam, sep="")
 }
 d <- matrix (,nrows,nfiles) ## a matrix to read the values into
 for(i in 1:nfiles) {
 dt<-read.table(fnames[i],nrow=nrows) ## read a file
 d[1:(dt$V1[1]+6),i] <-dt$V1 ## store in the matrix
 }
 list(d,fnames) ## output the matrix and file names as
}
seldat <- function(datin,skipstart=200,datlen=500){
lengthm <-max(datin[1,])
nrows=6+datlen
ncols=ncol(datin)
datout <- matrix(,nrows,ncols)
datout[2:6,] <- datin[2:6,]
##datout[1,] <- (datin[1,]-(skipstart+datlen))
```

```
datout[1,] <- datlen
datout[7:nrows,] <- datin[(7+skipstart):(6+skipstart+datlen),]
datout
}
plot1axis<-function(dat in){
 lengthm <-max(datin[1,])
 minfrq <- min(datin[2,])
 maxfrq <- max(datin[3,])
 delta <- min(datin[4,])
 xlabl <- "Wavelength" ## labels for the x and y axis
 ylabl <- "intensity"
 xax <- seq(minfrq-delta,maxfrq+delta,by=delta)
 miny <- min(datin[-(1:6),],na.rm=T)
 maxy <- max(datin[-(1:6),],na.rm=T)
 yax <- seq(miny,maxy,length=length(xax))
 plot(xax,yax,type="n",xlab=xlabl,ylab=ylabl)
 for (i in 1:nf) {
 lngth <- datin[1,i]
 dat <- datin[7:(lngth+6),i]
 xax <- seq(datin[2,i],datin[3,i],length=length(dat))
 lines(xax,dat,col=i)
 }}
readata<-
function(filedir,filepat="BG",startno=0,ext="asp",nrows=2000,nfiles=41,step=5,
omitstart=200,lendat=500){
tmpl <- readat(filedir,filepat,startno,ext,nrows,nfiles)
tmpm <- tmpl[[1]]
tmpf <- tmpl[[2]] ## separate out the matrix of data and the names
seldat(tmpm,omitstart,lendat) ## omit first 20 and have 900
}
```

```
smoothitsp<- function(datin,order){ ## spline smoothing
al1<-datin[1:500,]
ax1 <- 1:500
 alls1<-fitted(lm(al1~ns(ax1,df=order)))
al1<-datin[501:1000,]
ax1 <- 501:1000
 alls1<-rbind(alls1,fitted(lm(al1~ns(ax1,df=order))))
al1<-datin[1001:1500,]
ax 1 <- 1001:1500
 alls1<-rbind(alls1,fitted(lm(al1~ns(ax1,df=order))))
al1<-datin[1501:2000,]
ax 1 <- 1501:2000
 alls1<-rbind(alls1,fitted(lm(al1~ns(ax1,df=order))))
al1<-datin[2001:2500,]
ax 1 <- 2001:2500
 alls1<-rbind(alls1,fitted(lm(al1~ns(ax1,df=order))))
alls1
}
smoothitls<- function(datin){ ## loess smoothing; fixed smoothing
alls1 <- datin
al1 <- datin[1:500,]
ax1 <- 1:500
nrow <- dim(datin)[2]
for (i in 1:nrow) alls1[ax1,i]<-predict(loess(al1[,i]~ax1))
al1<-datin[501:1000,]
ax1 <- 501:1000
 for (i in 1:nrow) alls1[ax1,i]<-predict(loess(al1[,i]~ax1))
al1<-datin[1001:1500,]
ax1 <- 1001:1500
 for (i in 1:nrow) alls1[ax1,i]<-predict(loess(al1[,i]~ax1))
al1<-datin[1501:2000,]
ax1 <- 1501:2000
```

```
 for (i in 1:nrow) alls1[ax1,i]<-predict(loess(al1[,i]~ax1))
al1<-datin[2001:2500,]
ax1 <- 2001:2500
 for (i in 1:nrow) alls1[ax1,i]<-predict(loess(al1[,i]~ax1))
alls1
}
showall <- function(data,prts,colst) { ## plot the data in prts parts by colst
 len <- length(data)
 sp = len/prts
 plot(1:len,data,type="n")
 for (i in 1:prts) lines(((i-1)*sp+1):(i*sp),data[((i-1)*sp+1):(i*sp)],
 col = colst+i)
}
### end of functions Only change below
nf <- 5
tmps1 <- readata("C:/ Joseph/alldata/DUST X B.GLOBIGII FL
1/","B1.", 8, nfiles=nf,step=3)
plot1axis (tmps1)
tmps2<- readata("G:/Easaw/postgrads/Joseph/alldata/DUST Y B.GLOBIGII FL
1/","D1.",8,nfiles=nf,step=3)
plot1axis(tmps2)
tmps3<- readata("G:/Easaw/postgrads/Joseph/alldata/DUST Z B.GLOBIGII FL
1/","F1.",8,nfiles=nf,step=3)
plot1axis(tmps3)

tmps4 <- readata("G:/Easaw/postgrads/Joseph/alldata/DUST X BC FL
1/","T",8,nfiles=nf,step=3)
plot1axis(tmps4)
tmps5<- readata("G:/Easaw/postgrads/Joseph/alldata/DUST Y BC FL
1/","R1.",8,nfiles=nf,step=3)
plot1axis(tmps5)
tmps6<- readata("G:/Easaw/postgrads/Joseph/alldata/DUST Z BC FL
```

```
1/","S",8,nfiles=nf,step=3)
plot1axis(tmps6)
tmps7 <- readata("G:/Easaw/postgrads/Joseph/alldata/DUST X BP FL
1/","H",8,nfiles=nf,step=3)
plot1axis(tmps7)
tmps8<- readata("G:/Easaw/postgrads/Joseph/alldata/DUST Y BP FL
1/","J1.",8,nfiles=nf,step=3)
plot1axis(tmps8)
tmps9<- readata("G:/Easaw/postgrads/Joseph/alldata/DUST Z BP FL
1/","L",8,nfiles=nf,step=3)
plot1axis(tmps9)
## will plot the spectrum and will co add the data from five different excitation
wavelength, 280, 310,340,370 and 400 and will delect 200 data points from  the star
and retain 500 data point from fluorescence emission of each sample
library(mva)
all<-cbind(as.vector(tmps1[-c(1:6),]),as.vector(tmps2[-c(1:6),]))
all<-cbind(all,as.vector(tmps3[-c(1:6),]),as.vector(tmps4[-c(1:6),]))
all<-cbind(all,as.vector(tmps5[-c(1:6),]),as.vector(tmps6[-c(1:6),]))
all<-cbind(all,as.vector(tmps7[-c(1:6),]),as.vector(tmps8[-c(1:6),]))
all<-cbind(all,as.vector(tmps9[-c(1:6),]))
## will co add the five selected fluorescence ranges for each sample, form nine
columns corresponding to 9 variables(samples)
colst<- c(rgb(0,0,255,max=255),rgb(0,0,255,max=255),rgb(0,0,255,max=255),
rgb(255,0,0,max=255),rgb(255,0,0,max=255),rgb(255,0,0,max=255),
rgb(0,255,0,max=255),rgb(0,255,0,max=255),rgb(0,255,0,max=255))
## use different colour code for different samples
  cora<-cor(all,use="complete")
  eva<-eigen(cora)
  evav<-eva$vectors
  cora ## this plots out the full correlation matrix.
evav[,1:9] ## to see just the first 9 eigenvectors write evav[,1:9]
eva$values ## to see eigenvalues
```

```
X11()
plot(evav[,1],evav[,2],col=colst,pch=as.character(1:9),cex=1.2,
xlab="Eigen Vecor 1",ylab="Eigen Vector 2",main="PCA 2D Plot")
## will plot 2D PCA plot

X11()

cloud(evav[,1]~evav[,2]*evav[,3]) ## 3d plot
par(mfrow=(c(1,1))) ## splits screen to 4
cloud(evav[,1]~evav[,2]*evav[,3],col=colst,pch=as.character(LETTERS),cex=1.2,xl
ab="Eigen Vector 1",ylab="Eigen Vector 3",zlab="Eigen Vector 2")

## will plot 3D PCA plot

X11()
 hc <- hclust(dist(cbind(evav[,1],evav[,2])), "ave")
 plot(hc)
 plot(hc, hang = -1)

## will plot the hierarchical clustering

clc <- cl$cluster
par(mfrow=c(2,2))
tmp<-cbind(evav[,1],evav[,2])
plot(pam(tmp,3),cor=T,stand=F,shade=F,labels=2)

## will plot partition around medoid  with three clusters

X11()
clc <- cl$clustering
cl<-kmeans(cbind(evav[,1],evav[,2],evav[,3]),4,20)
plot(cbind(evav[,1],evav[,2],evav[,3]), col=cl$cluster)
points(cl$centers,col=1:27,pch=8)
lines(predict(ellipsoidhull(unname(cbind(evav[clc==1,1],evav[clc==1,2])))))
lines(predict(ellipsoidhull(unname(cbind(evav[clc==2,1],evav[clc==2,2])))))
lines(predict(ellipsoidhull(unname(cbind(evav[clc==3,1],evav[clc==3,2])))))
## will plot the k-means cluster
```

7.6 PCA Output

Example 1: PCA output from R program for PCA of three samples(with three measurements each) of *B. globigii* (i), *B. cereus*, *B. popilliae* from three different dust sources, A, B and C and is represented by numbers 1,2, ….9 respectively.

cora ## this plots out the full correlation matrix.

```
[,1]                 [,2]             [,3]          [,4]             [,5]
[1,]  1.000000000  0.984722353  0.99312237  0.067336167  0.07645990
[2,]  0.984722353  1.000000000  0.99326645  0.005293833  0.02128887
[3,]  0.993122370  0.993266449  1.00000000 -0.023893428 -0.01111374
[4,]  0.067336167  0.005293833 -0.02389343  1.000000000  0.98643952
[5,]  0.076459902  0.021288869 -0.01111374  0.986439525  1.00000000
[6,] -0.008975902 -0.035544159 -0.07483373  0.680306904  0.72981802
[7,] -0.490583415 -0.558908482 -0.55657989  0.624545825  0.63100596
[8,] -0.341941654 -0.381390331 -0.41567028  0.881443899  0.88385201
[9,] -0.348389495 -0.369769049 -0.41502365  0.871585813  0.88236798

[,6]                 [,7]       [,8]             [,9]
[1,] -0.008975902 -0.4905834 -0.3419417 -0.3483895
[2,] -0.035544159 -0.5589085 -0.3813903 -0.3697690
[3,] -0.074833735 -0.5565799 -0.4156703 -0.4150237
[4,]  0.680306904  0.6245458  0.8814439  0.8715858
[5,]  0.729818017  0.6310060  0.8838520  0.8823680
[6,]  1.000000000  0.6037112  0.6855760  0.6889685
[7,]  0.603711157  1.0000000  0.8619996  0.8010106
[8,]  0.685575968  0.8619996  1.0000000  0.9881296
[9,]  0.688968489  0.8010106  0.9881296  1.0000000
                                        >
> evav[,1:9] ## to see just the first 9 eigenvectors write evav[,1:9]

          [,1]       [,2]        [,3]       [,4]       [,5]       [,6]
 [1,]  0.2046043 -0.50953212  0.006141274  0.21368362 -0.14615009  0.01463434
 [2,]  0.2237498 -0.49512901 -0.011171213  0.04589586  0.37969557  0.09531347
 [3,]  0.2365369 -0.48480466 -0.001905256  0.18837562  0.12440565 -0.05443206
 [4,] -0.3529403 -0.30266092  0.317104423 -0.12748637 -0.50004359  0.57237411
 [5,] -0.3557292 -0.31314792  0.212934821 -0.15290948 -0.28567935 -0.75686355
 [6,] -0.3087474 -0.22895130 -0.873246757 -0.28231464 -0.05714724  0.07043739
 [7,] -0.3888886  0.07050062 -0.172957654  0.86942021 -0.04487000 -0.06101891
 [8,] -0.4236120 -0.08028257  0.170346337  0.06543897  0.34219126  0.26885976
 [9,] -0.4183381 -0.08178727  0.180125742 -0.19297176  0.60412283 -0.07945734
```

```
          [,7]        [,8]        [,9]
 [1,]  0.51343320 -0.01977834  0.60600068
 [2,] -0.71071863 -0.10859815  0.17995597
 [3,]  0.27115895  0.15166146 -0.74736720
 [4,] -0.16130203  0.23740481 -0.08358483
 [5,] -0.14120894 -0.17743153 -0.02820529
 [6,]  0.03266387 -0.01232809 -0.02342495
 [7,] -0.16383415  0.15810034  0.02038286
 [8,]  0.23076157 -0.72774868 -0.09250927
 [9,]  0.17498696  0.56763125  0.15657961

> eva$values ## to see eigenvalues

[1] 5.3421644118 2.9428487809 0.4181640315 0.2178284717 0.0627672456
[6] 0.0106667836 0.0027636426 0.0019909369 0.0008056955
```

Example 2: PCA output from R program for PCA of four samples(with three measurements each) of *B. cereus*, *B. globigii* (iii), pig weed, mixture of *B. globigii* (iii) and pig weed and is represented by numbers 1,2, ….12 respectively.

cora ## this plots out the full correlation matrix.

```
          [,1]       [,2]       [,3]       [,4]       [,5]       [,6]
 [1,]  1.00000000  0.98098978  0.87958541 -0.6071597 -0.61133952 -0.59626526
 [2,]  0.98098978  1.00000000  0.93796824 -0.6837688 -0.69000910 -0.67285869
 [3,]  0.87958541  0.93796824  1.00000000 -0.7603125 -0.77455552 -0.74736664
 [4,] -0.60715972 -0.68376880 -0.76031249  1.0000000  0.99782396  0.99837330
 [5,] -0.61133952 -0.69000910 -0.77455552  0.9978240  1.00000000  0.99553921
 [6,] -0.59626526 -0.67285869 -0.74736664  0.9983733  0.99553921  1.00000000
 [7,]  0.30419322  0.31204603  0.37527408  0.2147188  0.17400900  0.23907296
 [8,]  0.55736076  0.57251158  0.61437737 -0.0479923 -0.08565913 -0.02360082
 [9,]  0.36563726  0.37106738  0.41820730  0.1755194  0.13546921  0.20010142
[10,]  0.01599049  0.03417260  0.01181680  0.4871506  0.45993007  0.50094256
[11,] -0.01688948 -0.04043139 -0.14933748  0.6220353  0.60453077  0.63152123
[12,] -0.16530706 -0.17702540 -0.23235555  0.7019973  0.68170442  0.71144757
          [,7]       [,8]      [,9]      [,10]       [,11]      [,12]
 [1,] 0.3041932  0.55736076 0.3656373 0.01599049 -0.01688948 -0.1653071
 [2,] 0.3120460  0.57251158 0.3710674 0.03417260 -0.04043139 -0.1770254
 [3,] 0.3752741  0.61437737 0.4182073 0.01181680 -0.14933748 -0.2323556
 [4,] 0.2147188 -0.04799230 0.1755194 0.48715055  0.62203530  0.7019973
 [5,] 0.1740090 -0.08565913 0.1354692 0.45993007  0.60453077  0.6817044
 [6,] 0.2390730 -0.02360082 0.2001014 0.50094256  0.63152123  0.7114476
 [7,] 1.0000000  0.95424155 0.9961059 0.76537595  0.64105000  0.6604875
 [8,] 0.9542416  1.00000000 0.9696850 0.65932669  0.52356271  0.4983176
 [9,] 0.9961059  0.96968495 1.0000000 0.76620644  0.64690671  0.6518847
```

```
[10,] 0.7653760  0.65932669 0.7662064 1.00000000  0.95085853  0.9581048
[11,] 0.6410500  0.52356271 0.6469067 0.95085853  1.00000000  0.9747448
[12,] 0.6604875  0.49831757 0.6518847 0.95810484  0.97474476  1.0000000

> evav[,1:12] ## to see just the first 9 eigenvectors write evav[,1:9]

           [,1]       [,2]        [,3]        [,4]        [,5]        [,6]
 [1,] -0.1891262 -0.3478375  0.390854272 -0.485237125  0.27661347  0.27888695
 [2,] -0.2062054 -0.3607723  0.344311783 -0.261101634 -0.08797483  0.05621291
 [3,] -0.2242728 -0.3655773 -0.012099081 -0.005633526 -0.78280096 -0.32582507
 [4,]  0.3752495  0.1432591  0.008862791 -0.358291164 -0.16765731  0.01154736
 [5,]  0.3688464  0.1565269  0.048449397 -0.380135924 -0.21213132 -0.07646134
 [6,]  0.3778534  0.1338121 -0.008891583 -0.362053930 -0.14254459  0.10203249
 [7,]  0.2075086 -0.3571611 -0.435361102 -0.043373002  0.05443579 -0.03606031
 [8,]  0.1086430 -0.4182749 -0.300750774 -0.072030998  0.18648911 -0.05449029
 [9,]  0.1947806 -0.3734935 -0.363996514 -0.054908151  0.15656972  0.02129828
[10,]  0.3186462 -0.2391804  0.217993422  0.435454864 -0.18011807  0.58970990
[11,]  0.3444330 -0.1786390  0.450237314  0.178450102  0.30405256 -0.66264503
[12,]  0.3725683 -0.1454017  0.251619638  0.246637874 -0.14964773  0.07021777

           [,7]       [,8]       [,9]       [,10]       [,11]        [,12]
 [1,]  0.51942905  0.09510080 -0.13028059  0.01700819  0.067618321 -
0.006326736
 [2,] -0.69353289 -0.21661778  0.29160380  0.12898276 -0.007599656 -
0.007770381
 [3,]  0.23193486  0.17283016 -0.05558470 -0.09000079 -0.039245198
0.023312096
 [4,] -0.02657084  0.03302065 -0.05241370  0.20388752 -0.118456515
0.788559169
 [5,] -0.08114553 -0.13771074 -0.44242386  0.30781280 -0.152245526 -
0.550036183
 [6,] -0.04217994  0.27657992  0.41727280 -0.57706622  0.225937106 -
0.202445626
 [7,]  0.05170797  0.03007385  0.16153369  0.45723599  0.626505859 -0.049461327
 [8,] -0.25477893 -0.13718776 -0.55980513 -0.51201298  0.048529683
0.135014525
 [9,]  0.12218549  0.02527377  0.35703317  0.06859980 -0.708670632 -
0.110119591
[10,] -0.13852772  0.40285350 -0.17845238  0.09148638 -0.041143702 -
0.022941192
[11,] -0.02233848  0.28587317  0.01767196  0.02716231  0.008616399
0.001641074
[12,]  0.29006519 -0.74379645  0.15230469 -0.14357659  0.082010778
0.025997726
> eva$values ## to see eigenvalues
 [1] 5.9998099998 5.0048611268 0.5357899793 0.3931218844 0.0459665237
 [6] 0.0064397885 0.0049879535 0.0035611860 0.0021700832 0.0013451733
[11] 0.0010575234 0.0008887781
```