THE DEVELOPMENT OF A FIBER-OPTIC PROBE FOR THE \textit{IN VIVO} RESONANCE RAMAN SPECTROSCOPY OF NEUROTRANSMITTERS

by

Georg Schulze

B. Sc. Eng. (Chemical), The University of Pretoria, 1980.


A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

INTERDISCIPLINARY STUDIES

\textit{Biotechnology Laboratory}

\textit{Chemistry}

\textit{Psychology}

We accept this thesis as conforming to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

October 1996

© H. Georg Schulze, 1996
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Psychology, Chemistry, Biotechnology Laboratory

The University of British Columbia
Vancouver, Canada

Date 15 October 1996
Abstract

The measurement of neurotransmitter secretions by living cells, both in living organisms or in preparations, constitutes an enduring and vexing problem for neuroscientists due to the large number of substances involved at very low concentrations. An ability to correlate neurotransmitter secretions with various factors including organismic behavior would greatly advance our understanding of the organization and functioning of central nervous systems. This, in turn, has many important implications for the diagnosis and treatment of disorders of central nervous systems (mainly in humans) as well as for the design and implementation of information processing and control systems.

The work presented here was undertaken in order to explore a novel approach to this demanding problem. The objective was to develop a probe capable of measuring neurotransmitter secretions in real time, at physiologically relevant concentrations, and non-invasively in situ. Data were obtained using an ultraviolet resonance Raman spectroscopic analytical technique performed via optical fibers, and were analyzed primarily with artificial neural networks. To this end, a prototype tunable ultraviolet resonance Raman system was designed, assembled, commissioned and employed.

A general introduction to the problem and a discussion of existing techniques for neurotransmitter measurement are given in Part I. In Part II, the analytical method was shown to allow discrimination between several different neurotransmitters and some of their precursors, both on the basis of their spectra and the selective resonance enhancement of their spectra. Optical fibers were characterized with regard to their suitability for use with pulsed ultraviolet radiation in Part III and on the basis thereof selected for the construction of optical fiber probes. It was found that the performance of optical fibers varied greatly when subjected to pulsed ultraviolet radiation, making the selection of fibers a crucial factor in probe construction. Various design features influencing the efficiency of optical fiber probes were investigated using both theoretical
and empirical techniques. A right-angle geometry using a small diameter excitation fiber and several larger collection fibers in close proximity produced the most efficient probe. In Part IV the use of cell secretions as samples modelling in vivo conditions were investigated. It was also shown that these probes could be inserted via surgically implanted cannulae into and operated in the crania of experimental male rats without producing discernable behavioral artifacts. In Part V some signal recovery methods were investigated and it was shown that artificial neural networks could be used to identify and quantify neurotransmitters based on their Raman spectra. Part VI contains an assessment of the neuroprobe using neurotransmitter secreting cultured cells as a model system. The thesis is concluded with a discussion of the characteristics of an ideal biosensor, reviews the work done, and highlights some future directions.

This thesis represents my contributions toward the development of a tunable ultraviolet resonance Raman neurotransmitter probe. Within the scope of this work, limitations of the available equipment and other resources precluded the complete development of a high-performance neuroprobe, however, the data presented here demonstrate proof-of-concept and feasibility. In particular, what has hitherto been considered impossible - the use of optical fibers for pulsed ultraviolet remote resonance Raman spectroscopy - has been shown to be distinctly feasible. It has further been shown that ultraviolet resonance Raman spectroscopy is well-suited to the problem of resolving a mixture of neurotransmitters in a biological matrix. With the appropriate state-of-the-art equipment, there is now a very real possibility of obtaining detection limits of $1 \times 10^{-9}$ M for the catecholamine neurotransmitters and $1 \times 10^{-6}$ M for the aliphatic neurotransmitters with 30 s exposure time, thus providing a novel and general solution to the problem of neurotransmitter measurement.
## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>List of tables</td>
<td>xii</td>
</tr>
<tr>
<td>List of figures</td>
<td>xvi</td>
</tr>
<tr>
<td>Acknowledgement</td>
<td>xxvi</td>
</tr>
<tr>
<td>Dedication</td>
<td>xxvii</td>
</tr>
</tbody>
</table>

### PART I

General Introduction

#### CHAPTER 1

1.1 INTRODUCTION

1.2. OVERVIEW OF BRAIN FUNCTION

1.3. OVERVIEW OF BIOSENSORS

1.4. OVERVIEW OF NEUROTRANSMITTER SENSORS

1.4.1. Non-selective sensors

1.4.1.1. Collection techniques

1.4.1.2. Spectrometric techniques

1.4.2. Selective sensors

1.4.2.1. Electrodes

1.4.2.2. Optrodes

1.5. THE NEED FOR A NEW NEUROSENSOR

1.5.1. The need for a fast technique

1.5.2. The need for a general technique

1.6. PROPOSAL FOR A NEUROSENSOR

1.6.1. The problem

1.6.2. The conceptual solution

1.6.3. The approach

1.6.4. General principles and axioms

1.7. THESIS OUTLINE

1.8. REFERENCES

### PART II

Spectroscopy

#### Chapter 2

2.1. INTRODUCTION

2.2. NORMAL RAMAN SPECTROSCOPY

2.3. RESONANCE AND SURFACE ENHANCED RAMAN SPECTROSCOPY

2.4. INSTRUMENTATION

2.5. REFERENCES
# Chapter 3

3.1. Introduction 30
3.2. Instrumentation 31
3.2.1. System requirements and characterization 31
3.2.2. The detector 32
3.2.3. The spectrograph 33
   3.2.3.1. The single monochromator 33
   3.2.3.2. The double monochromator 33
   3.2.3.3. The fiber-optic adaptor 34
3.2.4. The fiber-optic probe and coupling optics 34
3.2.5. The light source 35
   3.2.5.1. The second harmonic generator 35
   3.2.5.2. The dye laser 35
   3.2.5.3. The pump laser 37
   3.2.5.4. The Ar+ laser 38
3.2.6. System throughput 38
3.2.7. Synchronization 39
3.3. Methodology 40
3.4. Data Analysis 41
3.5. Discussion 42
3.6. References 43

# Chapter 4

4.1. Introduction 44
4.2. Normal Raman Spectra 45
4.2.1. Neurotransmitter spectra and band assignments 45
   4.2.1.1. Physiological saline 46
   4.2.1.2. Acetylcholine 47
   4.2.1.3. Dopamine 48
   4.2.1.4. Epinephrine 49
   4.2.1.5. Norepinephrine 50
   4.2.1.6. Serotonin 51
   4.2.1.7. Histamine 52
   4.2.1.8. Aspartate 53
   4.2.1.9. γ-Amino butyric acid 54
   4.2.1.10. Glutamate 55
   4.2.1.11. Glycine 56
4.2.2. Other spectra 57
   4.2.2.1. Tryptophan 57
   4.2.2.2. Cerebrospinal fluid 57
   4.2.2.3. Anesthetic 58
4.2.3. Difference spectra 59
   4.2.3.1. Acetylcholine in DMEM 59
   4.2.3.2. Acetylcholine in cerebrospinal fluid 60
4.3. Fiber-optic Probe Raman Spectra 61
4.3.1. Neurotransmitter spectra 62
   4.3.1.1. Acetylcholine 62
   4.3.1.2. Dopamine 63
   4.3.1.3. Serotonin 63
4.3.1.4. γ-Amino butyric acid
4.3.1.5. Mixture spectra
4.4. DISCUSSION
4.5. REFERENCES

CHAPTER 5

5.1. INTRODUCTION
5.2. RESONANCE RAMAN SPECTRA
5.2.1. Different concentrations of methyl orange
5.2.2. Methyl orange at different visible frequencies
5.2.3. Differentiation between methyl orange and chromate
5.3. UV ABSORPTION SPECTRA
5.3.1. Methyl orange
5.3.2. Some aromatic neurotransmitters and amino acids
5.3.3. Other neurotransmitters
5.3.4. Cerebrospinal fluid
5.4. UV RESONANCE RAMAN SPECTRA
5.4.1. Neurotransmitters
5.4.2. Other
5.5. DISCUSSION
5.6. REFERENCES

PART III

Fiber-optics

CHAPTER 6

6.1. INTRODUCTION
6.2. FIBER-OPTICS
6.2.1. Light propagation
6.2.2. Light coupling
6.2.3. Light delivery efficiency
6.3. REMOTE RAMAN SPECTROSCOPY
6.3.1. In vivo Raman spectroscopy
6.4. REFERENCES

CHAPTER 7

7.1. INTRODUCTION
7.2. OPTICAL FIBER CHARACTERIZATION
7.2.2. Input coupling
7.2.1. INPUT COUPLING
7.2.1.1. Continuous wave excitation
7.2.1.2. Pulsed excitation
7.2.2. Optical damage
7.2.2.1. Continuous wave light
7.2.2.2. Pulsed light
# Chapter 10

10.1. Introduction 151
10.2. The Brain-Probe Interface 151
10.2.1. Interface assembly design 151
10.2.2. Placement coordinates 153
10.2.3. Cannulations 154
10.3. In Vivo Probe Evaluation 155
10.3.1. Rayleigh scattering in the lateral ventricle 155
10.3.2. Probe fouling 157
10.4. Possible Biological Effects of In Vivo Spectroscopy 159
10.4.1. In vivo pulsed UV probe operation 159
10.4.2. Motor deficits 160
10.4.3. Brain lesions 162
10.5. Discussion 162
10.6. References 164

## Part V

Signal Processing

# Chapter 11

11.1. Introduction 165
11.2. Signal Recovery 166
11.2.1. Noise removal 166
11.2.2. Deconvolution 168
11.2.3. Baseline removal 169
11.3. Signal Identification and Quantification 170
11.3.1. Classical least squares 171
11.3.2. Artificial neural networks 171
11.3.2.1. Architecture 172
11.3.2.2. Layer computation 175
11.3.2.3. Node computation 178
11.3.2.4. Weight optimization 180
11.3.2.5. Biological neural networks 181
11.3.2.6. Modifications 183
11.3.3. Water as an internal standard 185
11.4. References 186

# Chapter 12

12.1. Introduction 191
12.2. Noise Removal 191
12.2.1. Filtering 192
12.2.2. Autoaccumulation 193
12.2.3. Moving product processing 196
12.3. Deconvolution 198
12.3.1. Maximum entropy 198
<table>
<thead>
<tr>
<th>Chapter 13</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>13.1. <strong>INTRODUCTION</strong></td>
<td>205</td>
</tr>
<tr>
<td>13.2. <strong>SIGNAL CLASSIFICATION WITH NEURAL NETWORKS</strong></td>
<td>205</td>
</tr>
<tr>
<td>13.2.1. <strong>Layer sizes</strong></td>
<td>208</td>
</tr>
<tr>
<td>13.2.2. <strong>Transfer functions</strong></td>
<td>210</td>
</tr>
<tr>
<td>13.2.3. <strong>Generalization and discrimination</strong></td>
<td>211</td>
</tr>
<tr>
<td>13.2.3. <strong>Input features</strong></td>
<td>217</td>
</tr>
<tr>
<td>13.3. <strong>DISCUSSION</strong></td>
<td>218</td>
</tr>
<tr>
<td>13.4. <strong>REFERENCES</strong></td>
<td>222</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 14</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>14.1. <strong>INTRODUCTION</strong></td>
<td>224</td>
</tr>
<tr>
<td>14.2. <strong>MIXTURE RESOLUTION</strong></td>
<td>225</td>
</tr>
<tr>
<td>14.2.1. <strong>Artificial neural networks</strong></td>
<td>225</td>
</tr>
<tr>
<td>14.2.1.1. <strong>Training</strong></td>
<td>226</td>
</tr>
<tr>
<td>14.2.1.2. <strong>Layer number and size</strong></td>
<td>228</td>
</tr>
<tr>
<td>14.2.1.3. <strong>Transfer functions</strong></td>
<td>230</td>
</tr>
<tr>
<td>14.2.1.4. <strong>Split layers</strong></td>
<td>231</td>
</tr>
<tr>
<td>14.2.1.5. <strong>Local connections</strong></td>
<td>232</td>
</tr>
<tr>
<td>14.2.2. <strong>Classical least squares</strong></td>
<td>233</td>
</tr>
<tr>
<td>14.3. <strong>DISCUSSION</strong></td>
<td>236</td>
</tr>
<tr>
<td>14.4. <strong>REFERENCES</strong></td>
<td>238</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 15</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>15.1. <strong>INTRODUCTION</strong></td>
<td>240</td>
</tr>
<tr>
<td>15.2. <strong>QUANTIFICATION WITH NEURAL NETWORKS</strong></td>
<td>241</td>
</tr>
<tr>
<td>15.2.1. <strong>Training</strong></td>
<td>241</td>
</tr>
<tr>
<td>15.2.2. <strong>Spectral removal procedure</strong></td>
<td>242</td>
</tr>
<tr>
<td>15.2.3. <strong>Binary mixtures</strong></td>
<td>244</td>
</tr>
<tr>
<td>15.2.4. <strong>Tertiary mixtures</strong></td>
<td>245</td>
</tr>
<tr>
<td>15.2.5. <strong>Measured mixture spectra</strong></td>
<td>246</td>
</tr>
<tr>
<td>15.2.6. <strong>Removal fraction coefficient</strong></td>
<td>248</td>
</tr>
<tr>
<td>15.2.7. <strong>Multi-component mixtures</strong></td>
<td>250</td>
</tr>
<tr>
<td>15.2.8. <strong>Spectral shifts</strong></td>
<td>251</td>
</tr>
<tr>
<td>15.3. <strong>DISCUSSION</strong></td>
<td>253</td>
</tr>
<tr>
<td>15.4. <strong>REFERENCES</strong></td>
<td>255</td>
</tr>
<tr>
<td>APPENDIX A</td>
<td>288</td>
</tr>
<tr>
<td>------------</td>
<td>-----</td>
</tr>
<tr>
<td>Probe simulation program in MicroSoft Quick Basic 4.0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>APPENDIX B</th>
<th>293</th>
</tr>
</thead>
<tbody>
<tr>
<td>Program to perform randomly displaced autoaccumulations of a spectrum, written in MicroSoft Quick Basic 4.0.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>APPENDIX C</th>
<th>299</th>
</tr>
</thead>
<tbody>
<tr>
<td>Program to amplify signals and reduce noise in a spectrum using a moving product method. Written in MicroSoft Quick Basic 4.0.</td>
<td></td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 3.1
The average dye laser output energy (10 pulses) (a) as a function of the oscillator dye concentration and (b) as a function of time given a dye concentration of 276 mg/L.

Table 4.1.
Table 4.1 shows some of the major peaks in the Raman spectrum of acetylcholine (0.5 M), their vibrational band assignments, and relative intensities.

Table 4.2.
Some of the major peaks in the Raman spectrum of dopamine (0.5 M), their vibrational band assignments, and relative intensities.

Table 4.3.
Table 4.3 shows some of the major peaks in the Raman spectrum of epinephrine (0.5 M), their vibrational band assignments, and relative intensities.

Table 4.4.
Table 4.4 shows some of the major peaks in the Raman spectrum of norepinephrine (0.5 M), their vibrational band assignments, and relative intensities.

Table 4.5.
Table 4.5 shows some of the major peaks in the Raman spectrum of serotonin (0.1 M), their vibrational band assignments, and relative intensities.

Table 4.6.
Table 4.6 shows some of the major peaks in the Raman spectrum of histamine (0.5 M), their vibrational band assignments, and relative intensities.

Table 4.7.
Table 4.7 shows some of the major peaks in the Raman spectrum of aspartate (0.5 M), their vibrational band assignments, and relative intensities.

Table 4.8.
Table 4.8 shows some of the major peaks in the Raman spectrum of γ-aminobutyric acid (0.5 M), their vibrational band assignments, and relative intensities.

Table 4.9.
Table 4.9 shows some of the major peaks in the Raman spectrum of glutamate (0.1 M), their vibrational band assignments, and relative intensities.

Table 4.10.
Table 4.10 shows some of the major peaks in the Raman spectrum of glycine (0.5 M), their vibrational band assignments, and relative intensities.

Table 7.1.
The effects of placing the beam waist outside (A) or inside (B) the fiber. Excitation at 225 nm and 10 Hz.

Table 7.2.
The effects of increasing pulse energies on the transmission through a 4x300 μm input bundle of 6 cm length. Excitation at 266 nm and 10 Hz.
Table 7.3.  
The effects of increasing pulse energies on the transmission efficiency (average of 10 pulses) through a 15 cm section of 300 μm diameter Fiberguide Industries fiber (Superguide G). Excitation at 266 nm and 10 Hz.

Table 8.1.  
The effects of varying excitation and collection fiber diameters (μm) on the relative collection efficiencies (%) of dual fiber probes for maximum probe depths of 500 μm. The separation distances between fibers were a function of their diameters.

Table 8.2.  
The effects of probe tip geometry on the relative collection efficiencies of dual fiber probes for maximum probe depths of 500 μm and in solutions with varying absorbances. The separation distances between fibers were 20 μm.

Table 8.3.  
The effects of probe geometry on the relative collection efficiencies of dual fiber probes for maximum probe depths of 500 μm. The separation distances between fibers were 20 μm and that between collection fiber endface and excitation fiber apex varied as indicated.

Table 9.1.  
The number of viable and non-viable PC12 cells per hemocytometer counting area before depolarization.

Table 9.2.  
The concentration of neurotransmitters secreted by approximately 8.30x10^7 cells into 1 ml depolarizing solution.

Table 10.1.  
The means, standard errors of the means, and probabilities (α = 0.05) of 3 different behavioral measures obtained on 4 consecutive days before and after in vivo operation of the UV resonance Raman neuroprobe.

Table 12.1.  
The characteristics of signal, noise, and approximate SNR from an acetylcholine spectrum autoaccumulated with varying bandwidths and using uniform or normal distributions. Abbreviations: N = normal distribution; U = uniform distribution; n = number of samples; sd = standard deviation.

Table 12.2.  
The characteristics of signal, noise, and approximate SNR from an acetylcholine spectrum processed with autoaccumulation and varying bandwidths of a zero-order Savitsky-Golay filter. Abbreviations: SG = Savitsky-Golay filter; others as in Table 12.1.

Table 14.1.  
This table shows the desired output node values for the neurotransmitter Raman spectra used in the training data set. Abbreviations: ACH, A = acetylcholine; DOP = dopamine; EPI, EP = epinephrine; NOR = norepinephrine; SER = serotonin; HIS = histamine; ASP = aspartate; GAB = y-amino butyric acid (GABA); GLU = glutamate; GLY = glycine; and combinations of these.
Table 14.2. 228
This table shows the desired output node values for the neurotransmitter Raman spectra in the BASIC, NOISE, SH5, difference spectra, and INDEP testing data sets. See text for further explanation. Abbreviations as in the text and in Table 1; MIX = equimolar mixture of ACH, DOP, and EPI (the spectrum was not normalized and the table values reflect the relative Raman signal intensities for these three neurotransmitters).

Table 14.3. 229
The test results of three artificial neural networks when given the BASIC testing set as input. The networks had different configurations but used only hyperbolic tangent transfer functions. Consult the text for further details. Tfunction = transfer function; RMS = root mean square error value.

Table 14.4. 231
The test results of three artificial neural networks when given the BASIC testing set as input. The networks used different hidden layer transfer functions: sigmoidal (SGSG); hyperbolic tangent (TNSG); and sine (SNSG) transfer functions. Consult the text for further details. Tfunction = transfer function; RMS = root mean square error value.

Table 14.5. 232
The test results of artificial neural networks when given the BASIC testing set as input. (a) The networks had configurations and used transfer functions as indicated in the table. (b) The table shows a detailed breakdown of the network performance (RMS error values) on the binary and ternary mixtures in the BASIC set plus two additional binary mixtures. Also shown are RMS error values based only on the average values of the training set (AVG). Consult the text for further details. Tfunction = transfer function (or network name as it appears in the text); RMS = root mean square error value.

Table 14.6. 234
The mixture composition estimates of data in the BASIC testing set obtained with the classical least squares method (CLS) and a neural network (LC1) using local processing.

Table 14.7. 235
The RMS values and composition estimates of the acetylcholine difference spectra and data in the INDEP testing set obtained with the classical least squares method (CLS) and a neural network (LC2) using local processing. Also shown are the RMS values for the NOISE (5 runs) and SH5 testing sets.

Table 15.1. 245
Consecutive network estimates of the components present in the mixture and the highest estimate for components not present in the mixture (Other) for the AEPiGLY spectrum and residual spectra after component removal. Also shown are spectral average values. Value of α used: 80%. Abbreviations: ach = acetylcholine; epi = epinephrine; gly = glycine; nor = norepinephrine; ser = serotonin; Resid. = Residual.

Table 15.2. 246
Consecutive network estimates of the components present in the mixture and the highest estimate for components not present in the mixture (Other) for the AEPiGLY spectrum and residual spectra after component removal. Also shown are spectral average values. α used: 100%. Abbreviations: ach = acetylcholine; epi = epinephrine; gly = glycine; nor = norepinephrine; ser = serotonin.
Table 15.3.
Consecutive network estimates of largest component in mixture (Estimate), the fraction of that component removed from the mixture spectrum RF (RF = \( \alpha \) x Estimate), the scaling factor SF used to renormalize the remaining mixture spectrum, and CFs obtained for every component recovered from the MIX2 spectrum. Value of \( \alpha \) used: 50%.
Abbreviations: ach = acetylcholine; dop = dopamine; epi = epinephrine.

Table 15.4.
The table shows the network estimates for the three components present in the mixture of the MIX2 spectrum when the spectrum was shifted towards the excitation frequency by 0, 1, 2, 3, and 4 cm\(^{-1}\).

Table 16.1.
The concentration of neurotransmitters secreted by approximately 4.27x10\(^{-6}\) cells into 0.5 ml depolarizing solution as determined by HPLC.

Table 17.1.
The estimated improvement in detection limit that could be obtained with the use of different equipment or modifications to the existing system.

Table 17.2.
The results (after 2000 trials) of two networks with sine transfer functions trained by initial random assignment of either the weights (W) or the hidden layer values (H) followed by optimization with the simplex method.
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1.</td>
<td>3</td>
</tr>
<tr>
<td>A schematic representation of a neuron.</td>
<td></td>
</tr>
<tr>
<td>Figure 1.2.</td>
<td>14</td>
</tr>
<tr>
<td>A schematic representation of the <em>in vivo</em> Raman sensing process. It is for conceptual purposes only and not drawn to scale. In particular, the probe diameter is about 500 μm and the synaptic cleft about 50 nm.</td>
<td></td>
</tr>
<tr>
<td>Figure 1.3.</td>
<td>18</td>
</tr>
<tr>
<td>Schematic diagram showing the organization of the thesis and the major research areas involved.</td>
<td></td>
</tr>
<tr>
<td>Figure 2.1.</td>
<td>25</td>
</tr>
<tr>
<td>Energy level diagram showing the relationships between (a) IR absorption; (b) Rayleigh (R), Raman, and resonance Raman scattering; (c) fluorescence; and (d) resonance fluorescence.</td>
<td></td>
</tr>
<tr>
<td>Figure 2.2.</td>
<td>26</td>
</tr>
<tr>
<td>Absorption and Raman spectra of a compound with two chromophores (A and B) showing how selective resonance enhancement of the chromophoric Raman bands can be obtained by the appropriate choice of the excitation frequency.</td>
<td></td>
</tr>
<tr>
<td>Figure 3.1.</td>
<td>31</td>
</tr>
<tr>
<td>A schematic representation of the Raman spectroscopic system.</td>
<td></td>
</tr>
<tr>
<td>Figure 3.2.</td>
<td>36</td>
</tr>
<tr>
<td>The output of the dye laser as a function of the oscillator voltage of the pump laser after optimization of the dye laser using coumarin 460.</td>
<td></td>
</tr>
<tr>
<td>Figure 3.3.</td>
<td>37</td>
</tr>
<tr>
<td>The output energy from the pump laser at 532 nm as a function of the oscillator voltage setting: 500-700 V (dotted line) and 500-590 V Q-switched (solid line).</td>
<td></td>
</tr>
<tr>
<td>Figure 3.4.</td>
<td>39</td>
</tr>
<tr>
<td>A schematic representation showing the throughput of the Raman spectroscopic system at various points.</td>
<td></td>
</tr>
<tr>
<td>Figure 3.5.</td>
<td>40</td>
</tr>
<tr>
<td>A schematic representation of the optical path and electronic circuit time lags used to synchronize the optical and gating pulses.</td>
<td></td>
</tr>
<tr>
<td>Figure 4.1.</td>
<td>46</td>
</tr>
<tr>
<td>The Raman spectrum of physiological saline solution. The prominent peak is due to water deformation.</td>
<td></td>
</tr>
<tr>
<td>Figure 4.2.</td>
<td>47</td>
</tr>
<tr>
<td>The Raman spectrum of acetylcholine (0.5 M) in physiological saline. Inset shows structural formula.</td>
<td></td>
</tr>
<tr>
<td>Figure 4.3.</td>
<td>The Raman spectrum of dopamine (0.5 M) in physiological saline. Inset shows structural formula.</td>
</tr>
<tr>
<td>Figure 4.4.</td>
<td>The Raman spectrum of epinephrine (0.5 M) in physiological saline. Inset shows structural formula.</td>
</tr>
<tr>
<td>Figure 4.5.</td>
<td>The Raman spectrum of norepinephrine (0.5 M) in physiological saline. Inset shows structural formula.</td>
</tr>
<tr>
<td>Figure 4.6.</td>
<td>The Raman spectrum of serotonin (0.1 M) in physiological saline. Inset shows structural formula.</td>
</tr>
<tr>
<td>Figure 4.7.</td>
<td>The Raman spectrum of histamine (0.5 M) in physiological saline. Inset shows structural formula.</td>
</tr>
<tr>
<td>Figure 4.8.</td>
<td>The Raman spectrum of aspartate (0.5 M) in physiological saline. Inset shows structural formula.</td>
</tr>
<tr>
<td>Figure 4.9.</td>
<td>The Raman spectrum of γ-amino butyric acid (0.5 M) in physiological saline. Inset shows structural formula.</td>
</tr>
<tr>
<td>Figure 4.10.</td>
<td>The Raman spectrum of glutamate (0.1 M) in physiological saline. Inset shows structural formula.</td>
</tr>
<tr>
<td>Figure 4.11.</td>
<td>The Raman spectrum of glycine (0.5 M) in physiological saline. Inset shows structural formula.</td>
</tr>
<tr>
<td>Figure 4.12.</td>
<td>The Raman spectrum of tryptophan (0.1 M) in physiological saline.</td>
</tr>
<tr>
<td>Figure 4.13.</td>
<td>The Raman spectrum of cerebrospinal fluid taken during surgery (solid line), 24 h later (dashed line), and 72 h later (dotted line) Traces offset for clarity.</td>
</tr>
<tr>
<td>Figure 4.14.</td>
<td>The Raman spectrum of the anesthetic used during surgery.</td>
</tr>
<tr>
<td>Figure 4.15.</td>
<td>The Raman spectrum of DMEM (solid top line), DMEM spiked with acetylcholine to 0.5 M (dotted line), and the difference spectrum (bottom line).</td>
</tr>
<tr>
<td>Figure 4.16.</td>
<td>The Raman spectrum of cerebrospinal fluid taken during surgery (see Figure 4.13) spiked with acetylcholine to 0.2 M (top trace), and the difference spectrum (bottom trace).</td>
</tr>
</tbody>
</table>
Figure 4.17. The Raman spectrum of acetylcholine dissolved in physiological saline (0.5 M) obtained with a fiber-optic probe (baseline adjusted).

Figure 4.18. The Raman spectrum of dopamine dissolved in physiological saline (0.5 M) obtained with a fiber-optic probe (baseline adjusted).

Figure 4.19. The Raman spectrum of serotonin dissolved in physiological saline (0.1 M) obtained with a fiber-optic probe (baseline adjusted).

Figure 4.20. The Raman spectrum of γ-amino butyric acid dissolved in physiological saline (0.5 M) obtained with a fiber-optic probe (baseline adjusted).

Figure 4.21. The Raman spectrum of an equimolar mixture of serotonin and acetylcholine dissolved in physiological saline (0.5 M) obtained with a fiber-optic probe (baseline adjusted).

Figure 5.1. Spectra of different methyl orange concentrations obtained at 514.7 nm: (a) $1\times10^{-3}$; (b) $1\times10^{-4}$; (c) $1\times10^{-5}$, and (d) $1\times10^{-6}$ M.

Figure 5.2. Spectra of methyl orange ($1\times10^{-5}$ M) excited at different visible frequencies: (a) 488 nm; (b) 472.7 nm; and (c) 457.9 nm. Resonance enhancement occurred with excitation at 472.7 nm. The traces are offset for clarity.

Figure 5.3. Spectra of $20\times10^{-6}$ M methyl orange (a) and $200\times10^{-6}$ M chromate (b) excited at 364 nm and 472.7 nm showing differential resonance enhancement.

Figure 5.4. The UV absorption spectra of $3.3\times10^{-5}$ M dopamine (solid line), $3.8\times10^{-5}$ M epinephrine (dotted line), $108\times10^{-5}$ M tryptophan (dashed line), and $110\times10^{-5}$ M tyrosine (dash/dot). All analytes were dissolved in water.

Figure 5.5. A compilation of the UV absorption spectra of some of the 10 small-molecule neurotransmitters: (a) dopamine, epinephrine, norepinephrine, serotonin; (b) histamine; (c) glutamate; (d) aspartate; (e) acetylcholine; (f) glycine.

Figure 5.6. The UV absorption spectrum of cerebrospinal fluid diluted approximately 20 times with deionized water.

Figure 5.7. The resonance Raman spectrum of $1\times10^{-3}$ M dopamine (dotted line) dissolved in distilled water obtained at 227 nm, 20 Hz, 90 s exposure time, a monochromator setting of 235.3 nm and with = 17 mJ/pulse. The spectrum of the solvent only (solid line) and obtained under the same conditions.
Figure 5.8. The UV resonance Raman spectrum of 1x10^{-3} M epinephrine (dotted line) dissolved in distilled water measured at 227 nm, 20 Hz, 40 s exposure time, a monochromator setting of 234.8 nm and with ≈ 17 mJ/pulse. The spectrum of the solvent only taken under the same conditions (solid line).

Figure 5.9. The UV resonance Raman spectrum of 1x10^{-3} M serotonin dissolved in distilled water obtained at 227 nm, 20 Hz, 90 s exposure time, a monochromator setting of 235.3 nm and with ≈ 17 mJ/pulse.

Figure 5.10. The UV resonance Raman spectrum of 1x10^{-3} M histamine dissolved in distilled water obtained at 227 nm, 20 Hz, 90 s exposure time, a monochromator setting of 234.3 nm and with ≈ 20 mJ/pulse.

Figure 5.11. The UV resonance Raman spectrum of 50x10^{-6} M melatonin (top), 820x10^{-6} M tryptophan (center), 100x10^{-6} M tyrosine dissolved in distilled water (solid line) obtained at 227 nm, 20 Hz, 40 s exposure time, a monochromator setting of 234.8 nm and with ≈ 18 mJ/pulse. The spectrum of water obtained under the same conditions (dotted line).

Figure 5.13. Lorentzian curves fitted to the measured spectra of dopamine, epinephrine and norepinephrine indicating peak positions.

Figure 6.1. A diagram showing the coupling critical half-angle α and the transmission critical angle β of an optical fiber.

Figure 6.2. Transmission curves of fused silica (a) and glass (b), redrawn from the Newport Catalog (1993).

Figure 7.1. The pulse energy (mean ± S.E.M.; n = 10) as a function of oscillator setting when measured at 266 nm and 10 Hz before (solid line) and after (dotted line) a 150 mm focal length lens.

Figure 7.2. The attenuation for a 15 cm section of 300 μm diameter fused silica fiber excited at 266 nm, 10 Hz, and averaging 100 pulses per point, as a function of time. The recovery exhibited by the fiber after a 5 minute rest period is also shown (last measurement).

Figure 7.3. The attenuation for a 1 m section of 600 μm diameter fused silica fiber excited at 266 nm, 10 Hz, and averaging 10 pulses per point as a function of the distance between the input face and the beam waist.
Figure 7.4.  
The attenuation for a 15 cm section of 300 µm diameter fused silica fiber excited at 266 nm, 10 Hz, and averaging 100 pulses per point as a function of time. The fiber tip extended freely either 2 mm (top) or 6 mm (bottom) from its support. In neither case was the dynamic attenuation affected.

Figure 7.5.  
The attenuation (mean ± S.E.M.; n = 10) for 15 cm sections of 300 µm fused silica fibers from Polymicro Technologies (dotted line), 3M (dashed line), Fiberguide Industries (dash/dot), and the new experimental fiber from Polymicro Technologies (200 µm, solid line) at 230 nm and 10 Hz as a function of time. The last point in each of the two top traces were obtained after a 5-minute recovery period.

Figure 7.6.  
The transmission efficiency (mean ± S.E.M.; n = 10) of an experimental fiber (20 cm section, 200 µm) as a function of recovery time at 230 nm and 10 Hz.

Figure 7.7.  
The transmission efficiency (dotted line) and output pulse energies (solid line) of an experimental fiber (20 cm section, 200 µm) as a function of input energy at 230 nm and 10 Hz (mean ± S.E.M.; n = 10).

Figure 7.8.  
The output energy as a function of time for fibers tested under 3 conditions: increasing the input energy with small increments (e.g. oscillator voltage from 640-646; solid line); with a moderate increment (590-615 V; dashed line), and with a large increment (590-639 V; dotted line) to the final operating level. Data were scaled to the same initial values.

Figure 7.9.  
The central section emission beam profile of a 300 µm fused silica fiber with a Gaussian fit superimposed for comparison.

Figure 8.1.  
The simulated intensity distribution of light emerging from the excitation fiber of a front-casting probe.

Figure 8.2.  
The simulated intensity distribution of light emerging from the excitation fiber of a side-casting probe.

Figure 8.3.  
Two 300 µm fibers with polished enfaces before being assembled into a probe.

Figure 8.4.  
A silastic collar being stretched with fine wire to accommodate two optical fibers.

Figure 8.5.  
The aligned endfaces of two optical fibers of 300 µm diameter being held in place by the silastic collar.

Figure 8.6.  
End view of an optical fiber probe consisting of a central collection fiber (300 µm) surrounded by 8 angled excitation fibers (200 µm each).
Figure 8.7. 126
Side view of the optical fiber probe consisting of a central collection fiber surrounded by 8 angled excitation fibers shown in Figure 8.6.

Figure 8.8. 127
End view of the input bundle of the optical fiber probe consisting of a central collection fiber surrounded by 8 angled excitation fibers (200 μm each) shown in Figure 8.6. The single lighter fiber functions as a spacer to improve alignment of the fibers inside the glass collar.

Figure 8.9. 128
A dual front-casting probe glued beyond the silastic collar with a suspension of silicone rubber, cyclohexane, and tetradeacane, to produce a thin film of glue.

Figure 8.10. 130
The working curves of 3 probes with 300 μm excitation fiber diameters and 300 (dotted line), 600 (dashed line), and 1000 μm (solid line), collection fiber diameters, respectively. The working curves were normalized with respect to output power.

Figure 8.11. 131
The working curves of two probes with the same collection fiber diameter (1000 μm) and varying excitation fiber diameters, 300 μm (dotted line) and 1000 μm (solid line), respectively, normalized with respect to output power.

Figure 8.12. 132
The working curves of a 200x200 front-casting probe (dotted line) and a 200x200 side-casting probe (solid line) with aluminum pigmented acrylic paint applied to the angled facet of the excitation fiber to provide a reflective surface. The working curves were normalized with respect to output power.

Figure 8.13. 133
A diagrammatic cross-section of a probe consisting of an input bundle of 10 excitation fibers of 100 μm diameter each and a 400 μm collection fiber. Some excitation fibers were misaligned.

Figure 8.14. 133
A probe consisting of a 300 μm to 100 μm tapered excitation fiber surrounded by 4 collection fibers of 200 μm diameter each.

Figure 8.15. 134
The working curves of the 2 probes shown in Figures 8.13 (input bundle, dotted line) and 8.14 (tapered input, solid line), respectively. The working curves were normalized with respect to output power.

Figure 8.16. 136
The working curve (closed symbols) and detailed simulated working curve (open symbols) of a probe with 300 μm excitation fiber diameter and 600 μm collection fiber diameter. The working curves were normalized with respect to output power.

Figure 8.17. 137
The working curve for a side-collection probe using pulsed ultraviolet resonance Raman spectra of tryptophan dissolved in water excited at 227 nm, 20 HZ, and 300 μW average
power. The tryptophan signal (solid line) shows a maximum near 80 μM while the water signal (dotted line) declines monotonically.

Figure 9.1. 143
Secretions of neurotransmitters (dopamine, top; serotonin, center solid line; epinephrine, dotted line; norpinephrine, dashed line) from rat PC12 cells, measured with an HPLC, as a function of time.

Figure 9.2. 147
HPLC trace of neurotransmitter secretions from rat PC12 cells taken at 30 minutes after depolarization. Sample A (1.5 μl, solid line) was filtered with a 0.2 μm filter and sample B (3 μl, dotted line) with an ultrafilter.

Figure 9.3. 147
HPLC trace of neurotransmitter secretions from rat PC12 cells taken at 30 minutes after depolarization, as in Figure 9.2, but enlarged 4 times.

Figure 9.4. 148
Resonance Raman spectra of neurotransmitter secretions from rat PC12 cells taken after 30 minutes of depolarization (dotted line) and depolarizing agent (solid line, scaled down by a factor of 2 for comparison). Spectra were obtained with 227 nm excitation.

Figure 9.5. 149
Resonance Raman spectra of neurotransmitter secretions from rat PC12 cells taken after 30 minutes of depolarization: 8 spectra obtained with 227 nm excitation added together (top line); 3 spectra obtained with 218 nm excitation added together (bottom line). Backgrounds were removed.

Figure 10.1. 152
A schematic diagram of the protective sleeve designed for inserting the fiber-optic probe into the centrally implanted guide-cannula.

Figure 10.2. 154
A digitized histological slide from a trial implantation into the lateral ventricle of a male rat showing the cannula tract entering the left lateral ventricle. The figure was digitally enhanced to demonstrate the relative scales of the cannula tract and the ventricles.

Figure 10.3. 156
An anesthetized rat with fiber-optic probe inserted into the guide-cannula.

Figure 10.4. 157
Rayleigh scattering obtained in vivo from an anesthetized rat with 473 nm excitation and different monochromator settings: 896 (top), 1080 (middle), and 1187 cm⁻¹ (bottom).

Figure 10.5. 158
Two spectra obtained in vivo with 473 nm excitation and the same monochromator setting: (top) initial spectrum and (bottom) 25 minutes later.

Figure 10.6. 160
Spectra obtained in vivo from an anesthetized rat with pulsed 230 nm excitation and 3 different monochromator settings after processing: 1000 cm⁻¹ (top), 2000 cm⁻¹ (middle), and 3000 cm⁻¹ (bottom). Spectra are vertically displaced for ease of viewing.
Figure 11.1. 173
This figure shows an artificial neural network with 250 input nodes, 30 hidden nodes, and 10 output nodes. Each output node corresponds to one of 10 neurotransmitters. Also shown are data from a dopamine spectrum being given as input to the network and a network output classifying the input as dopamine.

Figure 11.2. 179
This figure shows (from left to right) the computations performed at hidden layer node 1. The data transferred from each of the 4 input nodes, multiplied by the respective connection weights (in small rectangles), are combined into a single value at this hidden node. This value is passed through the transfer function $F(x)$ to determine the hidden node's output. As before, this output is multiplied with the connection weights (in small rectangles) and transferred to the 3 output nodes.

Figure 11.3. 182
A simplified rendition of the primate retina showing a 3-layered structure similar to the backpropagation artificial neural network discussed in the text. The input layer consists of the rod receptor cells, the hidden layer consists of the bipolar cells, and the output layer of the ganglion cells. The axons of the ganglion cells form the optic nerve.

Figure 12.1. 192
The effects of filtering with a zero-order Savitsky-Golay filter with 10 (dotted line) and 50 (dashed line) point spread on a section of dopamine Raman spectrum.

Figure 12.2. 195
The effects of 50-point normally (dotted line) and uniformly (dashed line) displaced autoaccumulations on an acetylcholine spectrum.

Figure 12.3. 197
The effects of using a moving product to smooth a spectrum with SNR of 1. The Gaussian signal after Gaussian noise addition (top trace). The effects of smoothing with a 6 (dotted line), 20 (dashed line) and 25-point (solid line) moving product on the noisy spectrum. Original signal (before noise addition) shown at the bottom. Spectra were normalized (using factors of 4, 300, and 10,000 respectively) and offset for comparison.

Figure 12.4. 200
First stage in the baseline correction of a dopamine spectrum showing a polynomial fit (dashed line) and the removed outliers (Raman signals, dotted lines). The baseline-corrected spectrum is shown at the bottom.

Figure 12.5. 201
The effects of polynomial order on baseline fitting, displaced vertically for ease of viewing: 3rd order (top), 5th order (middle), and 6th order (bottom).

Figure 12.6. 202
The recovered signal (solid line) from a 0.01 M acetylcholine in physiological saline Raman spectrum (dotted line) after a 3rd order polynomial background removal and 1000 autoaccumulations.

Figure 13.1. 207
Typical learning curves for neural networks using sine transfer functions (solid line) and sigmoid transfer functions (dotted line).
Figure 13.2. Performance of neural networks with the same ratio of hidden to input layer nodes but different numbers of input layer nodes (IN136: 136 input nodes, IN150: 150 input nodes, IN250: 250 input nodes). Testing sets consisted of the baseline-corrected Raman spectra of neurotransmitters manipulated as follows: smoothed with a 10-point moving average (SM10); smoothed with a 5-point moving average and also used for training (SM5); not processed (RAW); shifted towards the exciting frequency by 5 cm\(^{-1}\) (SH5); and shifted towards the exciting frequency by 10 cm\(^{-1}\) (SH10).

Figure 13.3. Performance of neural networks with the same number of input nodes but different ratios of hidden to input nodes: ratio = 0.10 (RATIO10); ratio = 0.16 (RATIO16); and ratio = 0.25 (RATIO25). Testing sets as for Figure 13.2.

Figure 13.4. This figure shows the consequences of a relaxation of the training criterion on SNSG networks from a low root mean square (RMS) error value (SNSGL) to a higher RMS error value (SNSGH). The performance of a SNSG network with the more relaxed criterion (SNSGH) is shown for comparison. The discrimination index is defined in the text and the testing sets in the caption of Figure 13.2. SNSG networks use sine and sigmoid transfer functions for the hidden and output layers, respectively, and SGSG networks use only sigmoid transfer functions.

Figure 13.5. Performance of neural networks using different transfer functions: a sigmoid function (SGSG) and a sine function (SNSG). Testing sets are as given in Figure 13.2.

Figure 13.6. Abilities of two neural networks (SNSG, SGSG) with different transfer functions (sine, sigmoid) to discriminate between various spectra: 5-hydroxytryptophan (TRP, not in the training set) and serotonin (in the training set) as well as between epinephrine (ENP) and norepinephrine (NNP) and vice versa (both in the training set).

Figure 13.7. Discrimination index (see text) of two neural networks (SNSG, SGSG) with different transfer functions (sine, sigmoid) when presented with acetylcholine difference spectra (not included in the training set) obtained in cerebrospinal fluid (CSF) and Dulbecco's modified eagle medium (DMEM).

Figure 13.8. Performance of neural networks using different transfer functions: sine (SNSG) or sigmoid (SGSG); and different input features: interval means (IM) or uniform sampling (US). Testing sets as for Figure 13.2.

Figure 13.9. A schematic illustration of the discriminating abilities of ideal networks on hypothetical data sets. The solid line represents the performance of a discriminating network on noiseless data while the dotted line represents the performance of the same network on data with an arbitrary noise content. The dashed line represents the performance of a network with a better ability to generalize between data sets while the dotted line represents a network with a better ability to discriminate between data sets. The negative numbers on the abscissa represent an arbitrary increase in the loss of spectral detail (e.g. increasing the degree of smoothing) and the positive numbers an arbitrary increase in the loss of spectral peak positions (e.g. increasing the extent of spectral shifts). The '0' point
on the abscissa represents the spectral characteristics of the training data set. The discrimination index is given in arbitrary units and defined in the text. The figure is intended to show general trends only.

Figure 15.1. The Raman spectra of acetylcholine (a) and two experimentally measured mixtures, MIX1 (b) and MIX2 (c), both containing acetylcholine. In the mixture spectra, considerable overlap of the major component peaks occur.

Figure 15.2. The sum of all the recovered component fractions are plotted against the coefficient of the removal fraction used during spectral removal iterations. This graph illustrates that component fraction sums tend to converge to a minimum which is indicative of correct spectral recovery. The results shown are for the AEPIGLY (broken line, left scale) and MIX2 (solid line, right scale) spectra.

Figure 15.3. The number of iterations required for mixture decomposition as a function of the $\alpha$ value used. The results shown are for the AEPIGLY (broken line, right scale) and MIX2 (solid line, left scale) spectra. This figure indicates that the number of iterations required to resolve a mixture into its individual components is inversely proportional to the value of $\alpha$ used.

Figure 16.1. The UV resonance Raman spectra of cell secretion samples F (5 minute depolarization, top trace) and G (45 minute depolarization, bottom trace) obtained at 227 nm, 20 Hz, and approximately 400 $\mu$W employing a fiber-optic probe.

Figure 16.2. The UV resonance Raman difference spectrum of cell secretion samples F (5 minute depolarization) and G (45 minute depolarization) obtained at 227 nm, 20 Hz, and approximately 400 $\mu$W employing a fiber-optic probe.

Figure 16.3. The full-scale chromatograms of cell secretion samples F (5 minute depolarization, dotted line) and G (45 minute depolarization, solid line).

Figure 16.4. The scale-expanded chromatograms of cell secretion samples F (5 minute depolarization, dotted line) and G (45 minute depolarization, solid line).

Figure 16.5. The average output values produced by the neural network for the difference spectrum of cell secretion samples F (5 minute depolarization) and G (45 minute depolarization). Negative values (epinephrine and serotonin) were set to zero.

Figure 17.1. The working curves for front-(dotted lines) and side-casting probes (solid lines) plotted on a linear scale to show the linear ranges of the two probes for resonant (positive slope) and non-resonant (negative slope) analytes in an absorbing medium.
ACKNOWLEDGEMENTS

I wish to thank my committee members - Drs. Michael Blades, Alan Bree, Boris Gorzalka, and Robin Turner - for having given me this opportunity to pursue an interdisciplinary project and for their encouragement and support. Special thanks to my colleague, Shane Greek, for his collaboration, support, and patience, in spite of the many hours spent with severely attenuated visibility below 550 nm. I have also received courteous and expert help from technical professionals in the mechanical, electrical and glassblower's workshops of the Department of Chemistry, from technician Chris Sherwood in the Biotechnology Laboratory, and Dr. Alina Kulpa, a scientific engineer in the Department of Electrical Engineering. I wish to thank Dr. Ariane Coury in the Department of Psychology for analyzing some of my samples with her HPLC and Dr. Karl Klein from Friedberg, Germany, for his shared insights into optical fiber transmission processes. Finally I acknowledge with appreciation the many graduate, and sometimes undergraduate students, in the various departmental laboratories where I spent research time, for their patience with a novice and their willingness to show me the ropes.
Dedicated

with love

to the memory of

Neil Lewis Andrews:

PART I

General Introduction

CHAPTER 1

1.1 INTRODUCTION
1.2. OVERVIEW OF BRAIN FUNCTION
1.3. OVERVIEW OF BIOSENSORS
1.4. NEUROTRANSMITTER SENSORS
1.4.1. Non-selective sensors
    1.4.1.1. Collection techniques
    1.4.1.2. Spectrometric techniques
1.4.2. Selective sensors
    1.4.2.1. Electrodes
    1.4.2.2. Optrodes
1.5. THE NEED FOR A NEW NEUROSENSOR
1.5.1. The need for a fast technique
1.5.2. The need for a general technique
1.6. PROPOSAL FOR A NEUROSENSOR
1.6.1. The problem
1.6.2. The conceptual solution
1.6.3. The approach
1.6.4. General principles and axioms
1.7. THESIS OUTLINE
1.8. REFERENCES

1.1 INTRODUCTION

The fourfold requirement of performing rapid simultaneous measurements of many low concentration analytes in situ makes the development of an ideal sensor for investigating brain function an extremely difficult task. When to this requirement are added those of identification, quantification, biocompatibility, dynamic range, stability, ease of use, etc., the design of an ideal brain sensor becomes truly daunting. In spite of these considerable demands, the study of brain function is of such enormous importance to many disciplines that several techniques have been developed to investigate brain activity. Although these techniques have often lead to great advances in our understanding of the brain and its operation, their many limitations have denied researchers the opportunity to investigate numerous important aspects of brain activity effectively. The functioning of the normal and diseased brain therefore remains largely uncharted territory.
The aim of the interdisciplinary work undertaken and reported in this thesis was to conceptualize, design, and lay the foundations for the development of a brain sensor capable of meeting the general requirements of sensitivity, speed, and versatility while operating non-invasively in an aqueous in situ environment.

In this Chapter, brief overviews of general brain function, biosensors, and brain sensors are given. In addition, the need for and advantages of a new brain sensor are described, the central problem formulated, and the approach taken to solve this problem sketched. The chapter concludes with an outline of the thesis organization indicating the important areas of research involved in this work.

1.2. Overview of Brain Function

Brain function generally depends on two factors: the organization of brain cells and the operation of these cells. One of the most important classes of brain cells are nerve cells or neurons which operate by the propagation of intercellular electrical and chemical signals. Due to the electrochemical character of nerve signals, investigative techniques aimed at the elucidation of brain function typically detect electromagnetic (e.g. electroencephalograph) or chemical (e.g. in vivo dialysis) events associated with neuronal operation. In order to limit the scope of inquiry, only chemical sensing will be considered here.

A neuron, shown in Figure 1.1, consists of three parts: the cell body or soma, an input region of dendritic protrusions from the soma, and a single output conduit, the axon. The axon terminals abut on the somata and dendrites of other cells but for a minute gap (about 50 nm), the synaptic cleft.
Neurons are bathed in extracellular fluid rich in ions, primarily $\text{Na}^+$ and $\text{Cl}^-$. The presence of a selectively permeable cell membrane, intracellular protein anions, and an ion pump, result in the generation of a resting membrane potential of about $-70$ mV due to osmotic and electrostatic gradients and the activity of a transmembrane ion pump. Ion channels situated in the membrane can be gated electrically or molecularly to allow the free movement of $\text{Na}^+$ into the cell followed by the movement of $\text{K}^+$ out of the cell across the membrane. These ion fluxes locally change the membrane potential and signals are transmitted within a cell by a sequence of such changes. When potential changes are transmitted down the axon they are called action potentials and cause the release of gating molecules (neurotransmitters) from the axon terminals which then diffuse across the synaptic cleft and act on the ion channels of neighboring cells.

Such neurotransmission causes fluxes of cations and anions in the neighboring cell which produce excitatory and inhibitory postsynaptic potentials. If a summation of
these potentials exceeds a cellular threshold, an action potential is triggered at the axon near the soma and propagated toward the axon terminals.

Neurons are often highly organized into structures in many brain areas. The cortex, for instance, clearly consists of six layers (Kelly, 1985). The cerebellum, which plays an important role in motor control, shows exquisite organization in the arrangement of Purkinje cells and lateral fibers (e.g. Pellionisz and Llinas, 1979). In the tectum, input neurons terminate in a highly organized manner to form topographical maps of various sensory modalities (see Camhi, 1984).

1.3. OVERVIEW OF BIOSENSORS

Hall (1992) defines a biosensor as consisting of an analyte selective interface that relays the interaction between the surface and the analyte via some transducer to a detection system. This definition makes no reference to biology. In contrast, Buerk (1993) defines a biosensor as any measuring device that contains a biological (sensing) element. A third option is to define a biosensor as any measuring device that has been specifically designed or adapted to perform measurements in biological systems and this definition will be employed here. Sensors are often categorized on the basis of their transducers into electrochemical (amperometric, conductimetric, and potentiometric) and electromagnetic (optical and mass) sensors, but can also be categorized on the basis of their application (glucose sensors, pH sensors, etc.). Brain sensors, hence termed neurotransmitter sensors or neurosensors, can be seen as an application-specific subcategory of the set of biosensors.

1.4. OVERVIEW OF NEUROTRANSMITTER SENSORS

Chemical neurosensors can be divided into two broad groups: non-selective and selective sensors.
1.4.1. Non-selective sensors

1.4.1.1. Collection techniques

Ventricular perfusion uses artificial cerebrospinal fluid to perfuse the ventricular system of an experimental animal. This can be done while the animal is unanesthetized and freely moving (Pappenheimer et al., 1962). Samples can be taken at specific intervals and analyzed. It is impossible to be sure of the precise site of release of the neurotransmitters identified. Using recording electrodes in conjunction with this technique may allow for the identification of some of the sites where neurotransmitters were released (Obata and Takeda, 1969).

*In vivo* dialysis can overcome this problem to some extent, but suffers from another. This method involves the intracranial implantation of a probe into a specific site. A selectively permeable membrane in this probe allows certain neurotransmitters to diffuse upon release into a minute carrier stream which is collected at regular intervals (e.g. 12 minutes) before analysis with a high performance liquid chromatograph (Ungerstedt, 1986; Ungerstedt, 1984). This long time interval severely complicates the correlation of behavior with neurotransmitter release. The time spent in transport may also allow the breakdown of neurotransmitters and render the temporal resolution of transmitters released in close temporal contiguity to one another difficult. Another problem is that the selectivity of the membrane limits the number of neurotransmitters that can be detected. This technique is also rather invasive.

Suck-blow micropipettes are closed systems where material is collected during the period when a cell actively secretes a substance. This is then later released at the same position to determine the effects on the cell. Electrodes incorporated into the pipette are used for this purpose. The minute amounts of neurotransmitter involved can be collected and pooled for analysis. Push-pull cannulae are implanted cannulae allowing the withdrawal or application of substances centrally. A disadvantage is the fact that tissue damage is bound to occur and that collected material could be drawn from large
numbers of perhaps heterogeneous types of neurons (see Mitchell, 1975 for a discussion of these methods).

A definite disadvantage to all the above techniques is the time required for analysis. During the period between collection and analysis, the compounds in the sample may naturally degrade or may be degraded by other substances in the same sample.

1.4.1.2. Spectrometric techniques.

Nuclear magnetic resonance (NMR) spectroscopy is based on the absorption of radio frequency radiation by analytes subjected to an intense magnetic field (e.g. Skoog, 1985). NMR, employing surface coils, can be used to detect the formation of metabolites from radiolabeled precursor molecules introduced into living organisms (den Hollander et al., 1984; Tungsal et al., 1990). Although there is a fairly wide range of compounds that can be investigated with this method, it has several disadvantages. These include the lack of fine spatial localization and potential heating effects of the sample being investigated (Hammer et al., 1990), the difficulty in differentiating among compounds with a common precursor (Tungsal et al., 1990), the low signal-to-noise ratio (Sekihara and Ohyama, 1990), and the need to inject radio-active substances.

Surface-enhanced Raman spectroscopy (see Chapter 2) has been applied to the analysis of neurotransmitters (McGlashen et al., 1990; Morris et al., 1990) and peptides (Garrell et al., 1990). It is capable of detecting the biological amines at stimulated release concentrations (Morris et al., 1990). At present, some problems exist with the electrode coating to prevent protein fouling: thick coatings are more robust but increase response times, while thinner coatings have faster response times but are more fragile (Morris et al., 1990). Detection is dependent on the adsorption of neurotransmitters on metal surfaces and the adsorption kinetics (whether fast or slow adsorption occurs) as well as the degree of adsorption and the intensity of spectra generated by adsorbates would have
a bearing on the usefulness of this approach. This method, however, may have the potential for the development of a fast, general in vivo neurotransmitter probe.

Infrared absorption (IR) spectroscopy is based on the absorption of infrared radiation. Organic molecules produce complex spectra and IR spectroscopy is primarily used for their identification (e.g. Skoog, 1985). In vivo near-IR spectroscopy may also have the potential for a general method but appears to be currently used mainly to determine tissue and blood oxygenation through the detection of hemoglobin (Dickensheets and Cheung, 1989; Thorniley et al., 1988) and cytochrome c oxidase (Thorniley et al., 1988). In this method, near-infrared radiation is relayed to the sample (e.g. the patient's head) by optic fiber, passed through the sample (e.g. dermis, cranium, meninges etc.) and, upon emergence from the sample, collected and relayed to the sensors by optic fiber (Rea et al., 1985). Since tissue is radiated, the effective optical pathlength has to be estimated (Delpy et al., 1988). This technique, as currently used, has poor spatial resolution and does not appear to have been extended to investigate a wider range of substances. In addition, IR methods have poor detection limits.

1.4.2. Selective sensors.

Selective techniques are highly sensitive to a particular molecular species. This quality is generally gained by designing the measuring device so as to interact selectively with a particular type of molecule. Such an interaction then provides, either directly or indirectly, a signal indicating the presence of the analyte. Selectivity can be accomplished through a variety of ways: specific molecules can be 'selected' on the basis of their physical properties (e.g. size - permeable membranes and molecular sieves, natural fluorescence, light scattering, etc.) or on the basis of their chemical properties (e.g. interactions with antibodies, specific redox reactions, specific enzymatic reactions, etc.) or a combination of these methods. Many of the selective techniques combine fiber optics and spectrometry.
1.4.2.1. Electrodes.

*In vivo* voltammetry is a technique selectively sensitive to several of the monoamines (Blaha and Lane, 1983). This technique uses a chemically modified graphite paste electrode. Currents recorded are those arising as a result of monoamine oxidation subsequent to their release. The use of electrode arrays can produce very low detection limits (Monita and Niwa, 1996). However, should a cell release other substances, those remain unidentified. Furthermore, it is difficult to differentiate between the monoamines and/or their metabolites, e.g. dopamine can only be discriminated from norepinephrine after pharmacological pretreatment to suppress interfering metabolites (see Gonzalez-Mora et al., 1991).

There are two types of ion-selective microelectrodes: ion-selective glass microelectrodes and ion-specific liquid ion-exchanger electrodes. They measure ionic activity and not concentration. The ionic activity coefficient is also not known in and around living cells and has to be estimated. Other ions, intracellular diffusion barriers and compartmentalization could further complicate such measurements. Although these electrodes allow for more specific measurements, potentials recorded are still a summation of all potentials occurring between the microelectrode and the reference electrode and transmitter identification is problematic (for a discussion of these recording methods see Kelly et al., 1975).

Iontophoresis is related to the above techniques and is essentially used to record potentials resulting from the intracranial administration of neuroactive substances (Kelly et al., 1975). Such administration can be controlled with great accuracy using small electrical currents, but a major problem remains the identification of the actual neurotransmitters released and the temporal aspects of their release.
1.4.2.2. Optrodes.

Biological molecules with natural fluorescence can be detected with fluorescence spectroscopy which is much more sensitive and more selective than absorption spectroscopy (Bright, 1988). Optical fiber is used to relay exciting incident radiation and to collect emitted radiation (Kulp et al., 1988). Non-fluorescent species can frequently be fluorescently labeled (Bright, 1989) or transformed into a fluorescent molecule by chemical reaction which would render these substances detectable. Such modification of non-fluorescent molecules has to occur \textit{in vivo} and is accomplished through the use of an immobilized indicator phase on the sensing surface of the optic fiber. The analyte is made fluorescent when it combines with a reagent adsorbed on the indicator phase. The same general principle is used to construct optrodes that employ immune (Tromberg et al., 1988) and enzymatic (Kulp et al., 1988; Pantano and Kuhr, 1995) reactions to indicate the presence of analyte. In the latter cases, chemical reaction is not aimed at generating a fluorescent molecule, but to modify incident radiation to indicate the presence of the analyte (e.g. evanescent wave attenuation). The reader is referred to several reviews of this technology (Schultz, 1985; Seitz, 1989; Seitz, 1984; Sepaniak et al., 1988). The very selectivity of these methods also constitutes a limitation. Although some of these fiber optic probes can be very small, it is perhaps not feasible to use more than a few of them, each sensitive to a different analyte, in close proximity \textit{in vivo}. Furthermore, some of these methods cannot yield real time results due to the duration of diffusion and reaction processes (Sepaniak et al., 1988).

1.5. THE NEED FOR A NEW NEUROSENSOR.

Ideally, an investigative tool would be able to indicate which events are occurring and when they are occurring. That is, real time identification of neuronal events is desirable. At present, no technique can accomplish this. Existing techniques can generally be divided into two groups. In the first group are those techniques capable of
detecting and identifying a wide variety of neurotransmitters (e.g. *in vivo* dialysis). These techniques are generally slow, which is one of their major drawbacks. In the second group are techniques that are generally faster, but limited to the detection of single compounds or families of compounds (*in vivo* voltammetry). Their inherent specificity thus also constitutes an important deficit. A brief discussion of these two groups of techniques has been given above.

There does not seem to exist an *in vivo* technique that can be used with unanesthetized, freely moving animals to allow the rapid identification and quantification of all organic substances released in the process of neuronal signaling occurring at a specific central site. This view is in agreement with that of Morris et al. (1990). These authors concluded that there was a need for a general (amine) neurotransmitter probe providing resting-level sensitivity and real-time response; usable *in vivo*, with brain slices, extracts, and dialysates; and that was easy to fabricate and operate.

1.5.1. The need for a fast technique.

Many behaviors occur at timescales in the millisecond range. For instance, the lordosis reflex of the female rat (an index of sexual receptivity) occurs approximately 160 milliseconds (ms) after mounting by the male (Pfaff and Lewis, 1974) and lasts from about 200 ms to several seconds. Our understanding of this behavior would be greatly improved if we could determine the chemical signaling events associated with its onset, maintenance and termination (e.g. Schulze and Gorzalka, 1995). Another example is provided by Parkinson's disease, which is neurologically primarily characterized by a deterioration of the dopaminergic nigrostriatal pathway and behaviorally characterized by, among other symptoms, a rhythmic tremor of 3 - 6 Hz (e.g. Cote and Crutcher, 1985). Detection of the chemical signaling events responsible for these tremors will enhance our understanding of this disease and so facilitate treatment efforts. The vast majority of interneural signaling events are very rapid and only two examples from a
possible multitude have been presented. Clearly, to gain insight into the temporal characteristics of these processes, these events have to be detected as they occur. This will necessitate a technique with an appropriately fine temporal resolution.

1.5.2. The need for a general technique.

To date dozens of neurotransmitters have been identified. These fall into two broad classes: small-molecule neurotransmitters (about 10 different compounds) and neuropeptides (more than 30 different compounds). Several of these compounds coexist and are released concomitantly (e.g. Schwartz, 1985). A method that can detect a wide variety of unrelated neurotransmitters (e.g. peptides and small-molecule transmitters) is necessary to investigate this type of concomitant signaling. A consequence of concomitant signaling is that at least some postsynaptic neurons must be sensitive to two or more different types of neurotransmitters. Furthermore, many neurons receive inputs from a variety of sources perhaps using a variety of transmitters. In addition, many functional central neural circuits involve neurons that use different transmitters. For instance, an imbalance in the dopaminergic-cholinergic-GABA (γ-aminobutyric acid)-ergic loop, which is biochemically distinct but functionally integrated, can cause movement disorders. This happens in Huntington's disease where acetylcholine synthesizing and gamma-aminobutyric acid synthesizing neurons of the striatum are lost (e.g. Cote and Crutcher, 1985). Understanding the interactions between different transmitters will have undeniable advantages and demands a method capable of detecting them. Here too, a clear need exists.

1.6. PROPOSAL FOR A NEUROSENSOR

1.6.1. The problem

The eventual aim of this work is to furnish researchers with a sophisticated technique for the effective correlation of behavior and cognition with neurotransmitter
release in health and disease. In order to formulate the requirements for such a technique, some important issues will be considered in this section.

Chemical neurotransmission at the synapse takes about 10 ms (see Kandel and Siegelbaum, 1985) while behaviors become manifest within about 200 ms (e.g. Hyman, 1953; Pfaff and Lewis, 1974). For the adequate real time correlation of brain function with behavior and cognition, a temporal resolution of at least 10 ms is required.

Baseline levels of dopamine in the rat nucleus accumbens are about 1 nM as determined with in vivo microdialysis (Fiorino et al., 1993) while release levels around 1 µM can be obtained in the striatum under some conditions (Wood et al., 1992). Other neurotransmitters can be expected to have similar resting and release concentrations, thus establishing the required limit of detection at approximately 1 nM.

Up to now, approximately 40 chemicals have been identified as neurotransmitters (e.g. Schwartz, 1985). If one supposes at least one precursor and one metabolite for each of these, in excess of a 120 substances are directly or indirectly involved in neural signaling. As mentioned above, these different compounds are often jointly involved in some brain function. Furthermore, many neurotransmitters and their metabolites are often simultaneously present in some brain areas (e.g. Fiorino et al., 1993; Wood et al., 1992). Therefore, an adequate understanding of brain function cannot be obtained without the ability to measure several neurotransmitters, their precursors, and their metabolites concurrently.

In addition, measurements should take place in the extracellular fluid of the living brain under natural conditions (Crespi, 1990; Ungerstedt, 1986). This requirement implies remote measurement in an aqueous environment with minimal invasiveness.

Taken together, the following 4 stringent performance requirements apply to the ideal neurosensor:

(i) real time measurements (10 ms resolution);

(ii) physiological detection limits (1 nM);
(iii) concurrent multi-component detection (120 plus); and

(iv) *in situ* operation (remote aqueous samples).

To these are added the need to identify and quantify the various concurrently measured substances and the requirements of the ideal biosensor (e.g. linearity, dynamic range, biocompatibility, etc. - see Buerk 1993).

1.6.2. The conceptual solution

The best candidates for a fast, general technique appear to be the various analytical spectrometric methods. These methods involve the detection of radiation, can be extremely fast, and can be used with optical fibers, thus rendering remote data collection possible. Moreover, if the data thus collected can be appropriately processed and interpreted, the need to limit or select the data can be eliminated. Theoretically at least, spectrometry offers the possibility of fast and complete data collection *in vivo*.

There are several potential spectrometric techniques, e.g. fluorescence, absorption, reflectance, infrared absorption, Raman, nuclear magnetic resonance, to name but a few. Of these methods, Raman and especially resonance Raman spectroscopy appear to be the most feasible for a fast, general *in vivo* neurotransmitter analytical method for the following reasons: (i) they permit fast data collection due to the inherent rapidity of the scattering process; (ii) organic molecules have unique and relatively narrow Raman band profiles in the fingerprint 1500 cm$^{-1}$ to 500 cm$^{-1}$ region (unlike the broad bands produced by fluorescence); (iii) a low detection limit can be realized with resonance Raman spectroscopy (unlike infrared absorption spectroscopy); (iv) interfering fluorescence is absent below 260 nm; (v) significant interference by water, glass, and silica Raman scattering does not occur; (vi) selective resonance enhancement can be obtained with wavelength tuning and (vii) moderately uncomplicated instrumentation is required (unlike nuclear magnetic resonance spectroscopy).
In vivo Raman spectroscopy could be effected by the use of optical fibers to connect a living animal to a Raman spectrometer. An optical fiber probe could be constructed and inserted through an implanted intracranial cannula. When the probe is implanted and all the equipment operational, the release of a neurotransmitter or neurotransmitters from their nerve terminals would lead to a local increase in the concentration of such neurotransmitter(s). Diffusion would then occur away from the site of release. This process would bring the neurotransmitter molecules into the sensing area of the probe as shown in Figure 1.2.

![Figure 1.2](image)

Figure 1.2. A schematic representation of the in vivo Raman sensing process. It is for conceptual purposes only and not drawn to scale. In particular, the probe diameter is about 500 μm and the synaptic cleft about 50 nm.

The sensing area of the probe could either continuously or periodically be subjected to electromagnetic radiation of the required frequencies delivered by the excitation fiber. Interaction between the molecules and the radiation would change the frequency characteristics of some of the scattered radiation which can be collected and
relayed to a spectrometer by the collection fiber. The collected light can be decoded into its constituent frequencies by the spectrograph before delivery to a detector. The radiation so measured can then be mathematically processed to enable the identification of the different neurotransmitter species and their concentrations.

The dynamic nature of in vivo neurotransmitter levels makes it possible in principle to employ difference spectroscopy and matrix algebra to identify and quantify individual neurotransmitters. If neurotransmitters fluctuate independently in vivo (given a certain time scale), then a number of spectra taken at suitable intervals can be used to identify those neurotransmitters fluctuating most rapidly with the standard methods of matrix algebra (e.g. Gaussian elimination). For instance, if correlated with a certain behavior, 5 neurotransmitters show independent fluctuations in their levels (i.e. released at different times and/or rates), then 5 spectra taken at different intervals can be used to identify and quantify these 5 components. The fact that the tissue/extracellular fluid ratio for neurotransmitters can increase by up to 3 orders of magnitude upon pharmacological stimulation while those of precursors and metabolites show little change (see Westerink et al., 1987), provides added support for using difference spectroscopy. Methods such as the blind identification of independent source signals in sensor arrays (Cichocki and Unbehauen, 1993) and principal component analysis (e.g. Erickson et al., 1992) could also be applied to this problem. In addition, methods exist (e.g. classical least squares regression and artificial neural network processing - Schulze et al., 1995) for the resolution of the individual components of static mixtures which could be applied to individual spectra (a dynamic mixture could be considered static for a very short period of time during which a spectrum is obtained). Finally, for quantification purposes, water could serve as a convenient and practical internal standard.

Using fiber optic based resonance Raman spectroscopy (Chapter 2 contains a brief discussion of Raman spectroscopy and its instrumentation) has been considered difficult due to the problems detailed in Part III. These mostly concern the ability to
transmit pulsed ultraviolet radiation with optical fibers (see Chapter 7) and designing an appropriate probe for such work (see Chapter 8).

1.6.3. The approach

In order to establish practical goals, it was proposed to develop a Raman neuroprobe capable of identifying and quantifying the 10 small-molecule neurotransmitters (acetylcholine, dopamine, epinephrine, norepinephrine, serotonin, histamine, aspartate, γ-amino butyric acid, glutamate, and glycine) and mixtures of these in vivo. This consisted of 11 basic goals:

(i) the characterization of the individual neurotransmitters with normal and resonance Raman spectroscopy;

(ii) the determination of the viability of resonance Raman spectroscopy for obtaining selectivity and sensitivity;

(iii) the spectroscopic characterization of the intended environment;

(iv) the design and development of a suitable fiber-optic probe for in vivo operation;

(v) the determination of the conditions for optimum probe use;

(vi) the testing of the probe in vitro with neurotransmitters: individually, in mixtures, and using cell culture secretions to simulate in vivo conditions;

(vii) the testing and development of the brain/probe interface;

(viii) the testing for possible behavioral effects due to probe operation in vivo;

(ix) the selection/development of post collection processing methods to improve signal-to-noise ratios in order to help achieve stringent detection limits;

(x) neurotransmitter identification and quantification from complex spectra obtained under difficult conditions; and, if possible,
(xi) the fiber-optic based detection of these neurotransmitters *in vivo* including a comparison with *in vivo* voltammetry and *in vivo* dialysis to determine the viability of *in vivo* spectroscopy.

These goals were advanced in parallel, as much as circumstances permitted, rather than serially. This was done to avoid solving earlier problems in ways inconsistent with later requirements. Of these 11 basic goals, all but the last one have been completed. Contrary to the situation when this work was commenced, it is now for the first time possible, and as a direct consequence of this work, to perform *in vivo* resonance Raman spectroscopy with state-of-the-art equipment.

1.6.4. General principles and axioms

The development of this neuroprobe is based on the following general principles: That all data available *in situ* be collected and mathematically post processed for species identification and quantification before chemical and/or mechanical selection is attempted. In other words, it is preferable that selectivity be obtained mathematically (which retains all collected information) and not physically (which rejects much useful information before collection). The following axioms were postulated:

(i) neurotransmitters will not all be released at the same time;

(ii) neurotransmitters will not all be released at the same rate;

(iii) neurotransmitters will not all diffuse away from the site of release at the same rate;

(iv) neurotransmitters can all be differentiated by spectroscopic means, and

(v) failing (iv), neurotransmitters can be mechanically separated (e.g. with permeable membranes) until spectroscopic differentiation is possible.

Assumptions (i)-(iii) collectively assert the independence of neurotransmitter release (which may only be true by first approximation), while assumption (iv) is
justified by the uniqueness of Raman spectra (Chapter 2) and (v) by existing techniques (Chapter 1).

1.7. THESIS OUTLINE

The present thesis describes the progress made in the development of a resonance Raman neuroprobe for in vivo use. A tree diagram outlining the structure of the thesis and the important areas of research involved is given in Figure 1.3.

![Thesis structure diagram]

The thesis is divided into 6 parts. Part I, consisting of Chapter 1, serves as a general introduction to neurotransmission and the problem of neurotransmitter measurement. Each of the next 4 parts has an introductory chapter discussing the basic theory applicable to that part, hence serving as a reference chapter. Part II describes
Raman and resonance Raman spectroscopy and instrumentation; Part III describes the characterization of optical fibers and the development of optical fiber probes; Part IV describes the measurement of neurotransmitter secretions from cell cultures, the development of the brain/probe interface, and the behavioral effects of in vivo probe operation; and Part V deals with signal analysis. Part VI contains the penultimate chapter which attempts to draw all 4 primary areas of investigation together in a single application and the last chapter, Chapter 17, contains a summary and discussion of the progress made and addresses issues for future research.

1.8. REFERENCES.


PART II
Spectroscopy

Chapter 2

2.1. INTRODUCTION

In Part II the focus is on the spectroscopic work done for this multidisciplinary thesis. This work served to establish the uniqueness of Raman and resonance Raman spectra of neurotransmitter for identification purposes, the feasibility of using resonance Raman spectroscopy to obtain signal enhancement and selectivity, and to briefly characterize the intended measuring environment. In addition spectra were obtained for use in the signal processing part of the thesis (e.g. difference spectra). Part II also contains a description of the Raman system designed for this work. The present chapter establishes the theoretical basis for Part II. It provides brief introductions to Raman, surface enhanced Raman, and resonance Raman spectroscopy and the instrumentation used for Raman measurements.

2.2. NORMAL RAMAN SPECTROSCOPY

Rayleigh scattering of light by matter consists of the elastic scattering of incident radiation. In contrast, the Raman effect discovered in 1928 by the Indian physicist C.V. Raman, consists of the species-characteristic inelastic fractional (about $1/10^9$) scattering of incident radiation. In the particle theory view of Raman scattering, molecules are seen as being excited to non-quantized virtual energy states from where they immediately ($10^{-15}$ s) relax to vibrational ground electronic states. The scattered radiation is Stokes shifted if the molecule was originally in a ground vibrational state and relaxation
occurred to an excited vibrational state and is anti-Stokes shifted \textit{vice versa}, hence the Boltzmann distributions of thermally occupied states account for the relative intensities of these differently shifted frequencies. In the wave theory view, a proportional electric dipole moment, \( m \), with proportionality constant, \( \alpha \), is induced in a molecule subjected to an electric field \( E \). The size of the proportionality constant depends on the polarizability of the molecular bond involved in the vibration, which is in general dependent on the bond length. The intensity of scattering due to the Raman-active vibrational mode \( (I_S, \text{erg}) \) is proportional to \( \text{(proportionality constant K)} \) the squared derivative of the polarizability tensor elements \( (\alpha'^2) \) induced by the molecular vibration \( (\text{of frequency } \nu_1, \text{Hz}) \) caused by the excitation light, the concentration of the analyte \( (\text{proportional to } n_i = e^{-E/\text{kT}}) \) where \( E \) is energy (erg), \( k \) is Boltzmann's constant \( (1.3805 \times 10^{-6} \text{ erg/degree}) \), \( T \) is absolute temperature (degree), \( n_i \) are the number of molecules in state \( i \), the intensity of the incident radiation \( (I_0, \text{erg}) \), and the fourth power of the frequency of the incident radiation \( (\nu_0, \text{Hz}) \) as shown in Equation 2.1 (e.g. Asher, 1988):

\[
I_S = KI_0(\nu_0 \pm \nu_1)^4 \alpha'^2 n_i e^{-E/\text{kT}}
\]  

(2.1)

Energy exchanges of vibrational quanta between radiation and matter involve infrared frequencies, thus accounting for the close relationship between infrared and Raman spectroscopy. However, Raman spectra (dependent on the change in molecular polarizability) often provide information not available from IR spectra (dependent on the induced molecular dipole moment), rendering the two methods mostly complementary. In particular, water and glass exhibit weak Raman scattering while nonpolar bonds (e.g. organics) show strong Raman scattering (Gerrard, 1991), thus making Raman spectroscopy well suited for the fiber-optic based analysis of biological samples. The relationships between IR absorption, Raman and resonance Raman scattering, and fluorescence and resonance fluorescence, are shown in Figure 2.1.
2.3. Resonance and Surface Enhanced Raman Spectroscopy

The normally weak Raman signals can be enhanced in a number of ways, the two most important ones being resonant Raman scattering (RRS) and surface-enhanced Raman scattering (SERS). When the energy of the incident radiation approaches that of an electronic transition of the molecule, increases in the intensity of Raman scattering, up to 6 orders of magnitude, are observed (e.g. Vickers et al., 1991). This occurs because extra intensity is "mixed" into vibrational modes that are otherwise weak (e.g. Stencel, 1990). Consider again the change in polarizability term ($\alpha'$) from Equation 2.1 which can be rewritten as:

$$\alpha' = \frac{1}{h} \Sigma \left( \frac{(M_{me} M_{en})}{(v_{em} - v_0 + i\Gamma_e)} + \frac{(M_{me} M_{en})}{v_{em} + v_0 + i\Gamma_e} \right)$$

(2.2)

where $h$ is Planck's constant ($6.62 \times 10^{-27}$ erg s), the M's are electric transition moments (erg), the $v_{em}$ is the frequency (Hz) corresponding to the energy difference between
states $e$ and $m, m$ and $n$ index ground electronic vibrational states, $e$ indexes an excited electronic vibrational state, $i\Gamma_e$ is a damping constant (Hz), and $v_0$ has been defined in Equation 2.1 (Ferraro and Nakamoto, 1994). In normal Raman scattering, the excitation frequency is much lower than that of a molecular electronic transition and Raman bands are proportional to $(v_0 - v_1)^4$. However, when the excitation frequency is close to the absorption frequency of a particular chromophore in a molecule, the Raman band associated with that chromophore would get selectively enhanced (Asher et al., 1986; Ferraro and Nakamoto, 1994; Johnson et al., 1984). This occurs because $v_0$ approaches $v_{em}$ and the denominator of the first or "resonance" term of Equation 2.2 becomes very small, leading to a large increase in the value of the resonance term. This is illustrated in Figure 2.2.

![Figure 2.2](image_url)

**Figure 2.2.** Absorption (dotted line) and Raman spectra (solid lines) of a compound with two chromophores (A and B) showing how selective relative resonance enhancement of the chromophoric Raman bands can be obtained by the appropriate choice of the excitation frequency.
Resonance Raman scattering ($10^{-12}$ s) is slower than Raman scattering because relaxation now occurs, not from a virtual state, but from an excited electronic state. It differs from fluorescence ($10^{-9}$ s) in the respect that no prior relaxation to the lowest vibrational level of the excited electronic state occurs and is therefore a faster process (Skoog, 1985). Fluorescence is also absent in Raman scattering excited by wavelengths below about 260 nm (Johnson et al., 1984; Vickers et al. 1991).

Extraordinary enhancements of Raman signals, up to 6 orders of magnitude, can be obtained from analytes adsorbed on some rough metal surfaces such as Cu, Ag, Au, Ni, Pt, Pd, Ti, Hg, and others (Chang and Furtak, 1982; Kerker, 1990; Stencel, 1990). Initially, two general models have been advanced to account for this phenomenon: the electromagnetic model and the chemical model (Kerker, 1987; Pettinger, 1986). Enhancement according to the electromagnetic model occurs as a result of intensified local electromagnetic fields at or near metal rough spots. Chemical models attribute the enhancement to a modification of the molecular polarizability due to an interaction with the metal surface thus producing molecular resonances. Currently, it is widely accepted that there are at least two contributions to SERS: electromagnetic enhancement and an enhancement due to changes in the electro-optical parameters of the scattering process.

2.4. INSTRUMENTATION

The typical Raman spectroscopic system consists of a light source, excitation optics, the sample and sample holder, collection optics, a frequency decoder, detector, and data processing equipment.

The advent of the laser, providing intense monochromatic light, has facilitated spectroscopy based on the Raman effect. Continuous wave (cw) visible lasers are frequently used and the excitation wavelength should be chosen to avoid analyte absorption (except for resonance Raman spectroscopy), solvent absorption, sample
fluorescence and photodecomposition. Although less common due to the difficulties encountered with high peak energies and low duty cycles, pulsed lasers can also be used for Raman spectroscopy. These are presently the only means of providing continuously tunable UV radiation for UV resonance Raman work.

The excitation optics often consist of a single lens to focus the light on the sample but could also include an interference filter to remove unwanted laser radiation and/or optical fiber to deliver the light to the sample. In general, small samples with little preparation can be used for Raman spectroscopy. Glass and silica can be used as sample holders and capillary tubes are often employed for this purpose.

The collection optics generally consist of a high quality lens to collect and focus the scattered light onto the entrance slit of a monochromator and often include a notch filter to eliminate Rayleigh scattered light. Optical fibers can also be used for collection. The frequencies of the scattered light are spatially separated with a spectrograph. The use of a triple spectrograph significantly reduces stray light and is common.

Photomultiplier tubes, intensified photodiode arrays, and intensified charge coupled devices are detectors in frequent use for Raman spectroscopy. Photomultiplier tubes are highly sensitive but bulky devices and cannot easily be used for multichannel spectroscopy. Both photodiode and charge coupled device arrays permit multiplex spectroscopy with the general difference that photodiodes are more sensitive in the ultraviolet and charge coupled devices in the visible range of the electromagnetic spectrum. A controller is associated with the detector and allows the experimenter to vary a number of detector parameters such as temperature, exposure time, synchrony, etc.

Microcomputers with appropriate software are widely used for experimental control, as well as data collection and manipulation.

General descriptions of Raman and resonance Raman spectroscopy and its instrumentation can be found in Skoog (1985), Asher (1988), and Ingle and Crouch (1988), amongst others.
2.5. REFERENCES


CHAPTER 3

3.1. INTRODUCTION

3.2. INSTRUMENTATION

3.2.1. System requirements and characterization

3.2.2. The detector

3.2.3. The spectrograph

3.2.3.1. The single monochromator

3.2.3.2. The double monochromator

3.2.3.3. The fiber-optic adapter

3.2.4. The fiber-optic probe and coupling optics

3.2.5. The light source

3.2.5.1. The second harmonic generator

3.2.5.2. The dye laser

3.2.5.3. The pump laser

3.2.5.4. The Ar+ laser

3.2.6. System throughput

3.2.7. Synchronization

3.3. METHODOLOGY

3.5. DISCUSSION

3.6. REFERENCES

3.1. INTRODUCTION

The aim of the interdisciplinary work undertaken and reported in this thesis was to conceptualize, design, and lay the foundations for the development of a brain sensor capable of meeting the general requirements of sensitivity, speed, and versatility while operating non-invasively in an aqueous in situ environment as stated in Chapter 1. As part of this effort, it was necessary to design, procure, and commission equipment, individually and in concert, to create an analytical system suitable for in vivo continuously tunable UV resonance Raman investigations.

The following chapter contains the specifications established for this system including the appropriate justification. They also contain a description of the selected equipment, their characterization, and integrated operation. This chapter does not include the design and testing of fiber-optic probes, which are presented in Part III. For an introduction to Raman spectroscopy and instrumentation the reader may refer to Chapter 2.
3.1. INSTRUMENTATION.

Technically speaking, the design of a spectrometric system should proceed from the detector backwards to the light source to ensure that the entire system matches the performance requirements (e.g. Heiman et al., 1989). Given the fact that funds were not available to purchase all the constituent items simultaneously and that a number of important items were shared with or borrowed from other users, this was not feasible. However, the system, shown in Figure 3.1, will be discussed this way for illustrative purposes.

![Diagram of Raman spectroscopic system](image)

Figure 3.1. A schematic representation of the Raman spectroscopic system.

3.2.1. System requirements.

In order to identify the various neurotransmitters from their Raman spectra, it was necessary for these spectra to cover the infrared "fingerprint" region (500 cm\(^{-1}\) to 2000 cm\(^{-1}\)) at a sufficiently high resolution (4 cm\(^{-1}\)) (e.g. Tanabe et al. 1992). For quantification purposes, water, which has a peak near 1630 cm\(^{-1}\), could be used as an
internal standard. An ability to collect a full spectrum every 10 ms would provide the required temporal resolution (see Chapter 1.6) and make it equal to or better than that of in vivo voltammetry (Stamford, 1989). Taken together, a bandwidth of 1000 cm\(^{-1}\) to 2000 cm\(^{-1}\), a spectral resolution of 2 cm\(^{-1}\) to 5 cm\(^{-1}\), and a temporal resolution of 10 ms were desired. In addition, a tunable light source to generate light in the 200 - 300 nm region was required for UV resonance Raman spectroscopy (Asher, 1988). Finally, multi-mode, small-diameter optical fiber with high UV transmission and a high damage threshold to withstand the high pulse energies from the light source was needed.

3.2.2. The detector.

An intensified diode array detector (EG&G model 1455B-700-G, Princeton, NJ) was selected and purchased for the Raman neuroprobe system. This model consisted of a linear array of 700 intensified diodes, was blue enhanced to give a quantum efficiency of about 12% between 200 and 300 nm, gateable to 5 ns, had a 14 bit dynamic range, was thermoelectrically cooled to below -25 °C to reduce dark current, had a spectral range of 180 - 910 nm, and a 500:1 variable gain. The detector was controlled with a ST120 controller from the same supplier, enabling up to 1000 s integration time, 33 ms array readout time, and external triggering and synchronization. The detector was gated with an FG100 pulse generator, also from EG&G, with variable time delay and variable pulse width for use with pulsed light sources. Gating allows the experimenter to collect a spectrum for only that period of time during which the sample is being excited by a pulse of light. This reduces the number of dark counts and hence gives better signal-to-noise ratios.

The detector was thermoelectrically cooled and with the circulation of water coolant temperatures of -25 °C or less could be reached. In order to obtain noise levels, the baseline standard deviations of spectra of 1 s integration measured at 5 minute intervals were calculated. Spectra were collected starting immediately after commencing
detector cooling. The noise values decayed exponentially and stabilized to ~ 5 counts/s about 30 minutes after the start of detector cooling with 1 L/minute water flow.

3.2.3. The spectrograph.

The spectrograph (capable of multichannel spectroscopy as opposed to the spectrometer which is a scanning instrument) could be operated as a single monochromator. With the addition of a double monochromator in subtractive mode to function as an optical filter for Rayleigh rejection, a triple monochromator was formed.

3.2.3.1. The single monochromator.

A Model 207 spectrograph from McPherson (Acton, MA), was acquired for the Raman system. This instrument was of the Czerny-Turner arrangement, had a 0.67 m focal length, accommodated a 120 mm by 140 mm 3600 G/mm holographic grating to give a f/4.7 throughput, had 2 entrance slits with folding mirror, an exit port that could accommodate the detector, a linear dispersion of 0.415 nm/mm in first order to give a bandwidth of 1203 cm⁻¹ and resolution of 1.7 cm⁻¹ (with 3600 G/mm grating operating at 250 nm).

The incoming light was dispersed in the single monochromator by a grating and then focused onto the detector. The calibration of the instrument was checked with visible Hg lines and a 1200 G/mm grating. For a given setting, the center pixel was found to be 530 and the dispersion 0.031 nm/pixel (equivalent to 0.402 nm/mm with a 3600 G/mm grating).

3.2.3.2. The double monochromator.

The Model 275D double monochromator with 2 holographic gratings (1200 G/mm) from the same supplier was purchased to function as a stray light rejecter (specified better than 10⁻⁸ at 1.5 bandpasses from a given line). This unit was attached to
the Model 207 such that the exit port of the Model 275D mated with the entrance port of the Model 207, had an aperture ratio of f/4.2, a linear dispersion of 4 nm/mm to give a maximum bandwidth of 2576 cm\(^{-1}\) at 250 nm, and covered the spectral range 185 - 1000 nm.

The double monochromator functioned by dispersing the incoming light onto a central slit that could be adjusted to give a certain bandpass. The light passing the central slit was then recombined and imaged onto the entrance slit of the single monochromator. When the instrument was in operation, it was found that the exit slit affected the bandpass of the unit. This was anomalous. Upon questioning, the manufacturer modified the instrument by inserting inversion optics at the position of the central slit which corrected the unit's operation.

3.2.3.3. The fiber-optic adapter.

A fiber-optic adapter, Model 132, with a SMA connector, x-y micropositioners for aligning the fiber with the entrance slit of the spectrograph, UV grade optics to aperture-match the fiber (f/4.5) and spectrometer (f/4.7 or f/4.2), adjustable focus, and filter receptacle was also obtained from McPherson. This unit made possible the use of a single optical fiber input to either the single or double monochromator entrance slit assemblies.

3.2.4. The fiber-optic probe and coupling optics.

Fiber-optic probes were mostly constructed from glass (F-MLD-10) from Newport (Fountain Valley, CA), UV grade fused silica (Superguide G) from Fiberguide Industries (Stirling, NJ), and high UV transmission fused silica (Polymicro Technologies, Phoenix, AZ), but occasionally from other suppliers (see Chapter 8). These were step-index fibers with numerical apertures of 0.22 (f/4.5), core diameters from 100 - 1000 \(\mu\)m, cladding thicknesses typically 10% of the core diameter, and mostly acrylate
jacketing material. The input ends of the excitation fibers were held in a chuck and positioned with a fiber-optic positioner (Newport Model F915). Light was coupled into the fiber with a microscope objective (Newport M-10X, visible; U-27X, UV) or a 150 mm focal length plano-convex silica lens.

3.2.5. The light source.

The tunable ultraviolet light source consisted of the frequency doubled output from a dye laser pumped by a Nd:YAG laser. For visible light excitation, the lines from an Ar⁺ laser was used.

3.2.5.1. The second harmonic generator.

Second harmonics were generated with a CSK Optronics Super Doubler (Culver City, CA) using β-barium borate crystals (205-220 nm/ 220-250 nm), cut at phase-matching angles, and separated from the fundamental with a fused silica prism. The efficiency of the frequency doubler was specified as being larger than 5% in the 205 - 250 nm range. When measured at 225 nm, the efficiency was found to be 6.4%.

3.2.5.2. The dye laser.

Tunable pulsed visible light was obtained with a Quanta Ray PDL-1 narrow-band (less than 0.25 cm⁻¹ at 560 nm) dye laser (Mountain View, CA), pumping coumarin 460, Exciton 417, or Exciton 418 dye (Exciton, Dayton, OH), and cited efficiencies from 5 to 20%.

It was difficult to establish the efficiency of the dye laser due to the lack of a suitable energy meter head to measure the Q-switched output of the pump laser at 355 nm and high oscillator energy settings. However, the output of the dye laser was measured as a function of the oscillator energy setting of the pump laser after optimization of the dye laser using coumarin 460 and the results are shown in Figure 3.2.
By placing the energy meter head a suitable distance beyond the focal point of a 50 mm lens, the active area of the energy meter head could be filled with radiation, thus avoiding possible damage to the head. However, the formation of an air plasma at high pulse energies (plateau in Figure 3.2) clearly reduced the throughput thus limiting this approach. The dye laser was further optimized with regard to the oscillator dye concentration (initially 300 mg/L oscillator and 60 mg/L amplifier) and the initial concentration found to be near optimal. Subsequently, the output (shown in Table 3.1) was measured over a 30 minute period at high pump energies (pump laser oscillator setting at 710, 5 Hz) and found to be stable.

**Table 3.1**

*The average dye laser output energy (10 pulses) (a) as a function of the oscillator dye concentration and (b) as a function of time given a dye concentration of 276 mg/L.*

<table>
<thead>
<tr>
<th>Concentration (mg/L)</th>
<th>Output (mJ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>324</td>
<td>4.40</td>
</tr>
<tr>
<td>314</td>
<td>4.50</td>
</tr>
<tr>
<td>303</td>
<td>4.60</td>
</tr>
<tr>
<td>300</td>
<td>4.60</td>
</tr>
<tr>
<td>293</td>
<td>4.65</td>
</tr>
<tr>
<td>276</td>
<td>4.60</td>
</tr>
<tr>
<td>273</td>
<td>3.90</td>
</tr>
</tbody>
</table>
3.2.5.3. The pump laser.

The 355 nm output (shown in Figure 3.3) from a Lumonics HY-400 Nd:YAG laser (Rugby, England) operating at 10 Hz was used to pump the dye laser. A half wave plate in a rotary mount (CVI, Albuquerque, NM) was interposed between the pump and dye lasers to adjust the vertically polarized pump light to the horizontal orientation required by the dye laser.

<table>
<thead>
<tr>
<th>Time (minutes)</th>
<th>Output ± S.E.M. (mJ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>4.65 ± 0.15</td>
</tr>
<tr>
<td>5</td>
<td>4.60 ± 0.11</td>
</tr>
<tr>
<td>10</td>
<td>4.72 ± 0.10</td>
</tr>
<tr>
<td>15</td>
<td>4.76 ± 0.08</td>
</tr>
<tr>
<td>20</td>
<td>4.67 ± 0.10</td>
</tr>
<tr>
<td>25</td>
<td>4.66 ± 0.07</td>
</tr>
<tr>
<td>30</td>
<td>4.65 ± 0.10</td>
</tr>
</tbody>
</table>

![Figure 3.3. The output energy from the pump laser at 532 nm as a function of the oscillator voltage setting: 500-700 V (dotted line) and 500-590 V Q-switched (solid line).]
The doubled (532 nm) and quadrupled (266 nm) output from the pump laser was sometimes used directly for Raman excitation. The output of the pump laser was re-optimized whenever a different wavelength was used.

3.2.5.4. The Ar⁺ light source.

A Spectra Physics Stabilite 2017 Ar⁺ laser (Mountain View, CA) was used to provide continuous wave visible and UV excitation. The 488 nm, 473 nm, and 364 nm, lines were mostly used. The output from this laser was measured with a power meter and the readings found to correspond with those of the instrument's built-in power meter. Furthermore, plasma lines from ion lasers can interfere with Raman measurements (see Figure 10.4, top trace, about pixel 730) and interference filters are often used to avoid this problem. Since an interference filter was only available for the 488 nm line, and since several of the other lines were periodically used for excitation, the laser was occasionally detuned to verify the existence of plasma lines. When the laser was detuned, lasing ceased, but the plasma lines persisted to the extent that they could be easily and safely detected with the spectroscopic equipment.

3.3.1.8. System throughput.

Using 200 mW at 514 nm, the throughput of the system as shown in Figure 3.4 was measured. With the single monochromator the throughput to the detector was about 12%, but with a triple monochromator arrangement the throughput to the detector was too low to measure and estimated at 1%, an additional 93% (of 12%) loss. Values reported in the literature range from 1% - 5% (e.g. Chase, 1994). The most severe losses occurred at the gratings. These results suggested that a prism monochromator system may be more advantageous for Raman spectroscopy, notably in the ultraviolet region where the dispersion of quartz would be more pronounced.
3.3.1.9. Synchronization.

In order to test the synchronization of the gating of the detector with the arrival of the optical pulse from the laser, the system was set up as shown in Figure 3.5. The Pockels cell synchronization output started 100 ns before the optical pulse from the laser. From the moment of triggering by the Pockels cell to the arrival of the light at the detector, about 120 ns elapsed, while the electrical signal took about 64 ns to reach the detector. The electrical path consisted of two sections of coaxial cable (30 foot and 4 foot, respectively) and the fixed internal delay of the pulse generator (30 ns). The variable delay feature of the pulse generator (part of the electrical pathway) was used to compensate for the difference. Furthermore, the delay required proved to vary with the oscillator energy setting. In general, the gate width was set at a maximum to bracket the signal whereafter it was reduced to optimize the signal-to-noise ratio. The gating could perhaps be improved by splitting a small fraction of the beam to allow triggering from a
separate detector. Due to small pulse-to-pulse variations in the Q-switch operation, using the Q-switch for triggering demands a wider gate width allowing for the build-up of unnecessary dark charges. However, the gate pulse generator has an internal delay of 30 ns while the optical path is traversed in 19 ns thus requiring a delay of the main optical beam.

3.3. METHODOLOGY.

The general manner in which the equipment was operated and data collected is described in this section. The construction of the optical fiber probe is described in III. More detailed information will be provided in subsequent chapters describing spectroscopic measurements.

Prior to data collection, the input fiber endface was rinsed with deionized water, and subsequently visually aligned with the laser beam and optimized with the aid of a

Figure 3.5. A schematic representation of the optical path and electronic circuit time lags used to synchronize the optical and gating pulses.
piece of white paper (which reflected visible radiation and produced violet fluorescence from UV radiation). Optimizing the coupling visually was generally satisfactory as verified with a power/energy meter.

Optimizing the coupling of light from the collection fiber(s) into the spectrometer was done with the 1050 cm\(^{-1}\) peak of KNO\(_3\) (visible excitation) and the Rayleigh line (UV excitation). The slit width on the spectrometer was generally on the order of the collection fiber/bundle diameter and the intermediate slit on the double monochromator fully open for maximum bandpass. A 1200 G/mm holographic grating was used for frequency decoding visible wavelengths and a 3600 G/mm holographic grating for UV wavelengths.

A variety of sample holder shapes and sizes were employed for Raman measurements. With some probe designs (ends aligned flush) and visible excitation, a measuring cylinder was used. It was modified by covering the bottom with a piece of black plastic to attenuate the amount of stray light collected by the probe (and hence the generation of large spectral backgrounds).

When using UV excitation, a background was collected after the detector had been cooled sufficiently to minimize dark current noise and this background was subtracted from the measured spectra to remove the deterministic readout noise of the detector. This was rarely done with visible excitation because signals were mostly clearly detectable.

The components comprising the light source were optimized only initially or after some change in set-up or configuration had been implemented (e.g. changing to UV optics on the Ar\(^+\) laser).

3.4. DATA ANALYSIS.

For the identification of neurotransmitters from their Raman spectra, artificial neural networks were used. Difference spectroscopy was investigated with the intention
to remove static and slowly changing spectral components and thus simplify complex spectra before identification. For quantification, the water peak at 1630 cm\(^{-1}\) served as an internal standard and working curves for some individual neurotransmitters were established. The data analysis is discussed in detail in Part V.

3. DISCUSSION.

This discussion serves to briefly outline some of the operational implications inherent in a system as described above. In general, one wishes to collect as much scattered light as possible, this would give spectra in the shortest possible time. Collecting spectra for a longer time generally reduces the noise level and increases the signal strength (see Chapter 11) so that even very weak signals could be detected if collection continues for an adequate length of time. Due to the reduction in noise level, the resolution of the collected spectra are better. If a higher resolution of the spectra is required, the entrance slit to the monochromator could be narrowed (e.g. Skoog, 1985), thus requiring more collection time. The resolution could also be increased by using a dispersive element (grating or prism) with greater dispersion. This would also lead to a reduction of throughput to the detector and require more collection time. Hence improved resolution and better signal-to-noise ratios are generally gained at the expense of collection speed.

Higher intensity signals could be obtained by increasing the laser power thus producing more intense scattering from the sample. Under some conditions, samples and/or fiber optic probes (see Part III, Chapter 7) may be sensitive to high powers, thus establishing effective upper limits to the laser power that could be used. Higher intensity signals could also be obtained by using more sensitive detectors, or a better cooling of a detector. Furthermore, a better throughput of the system (i.e. using more efficient dye lasers, doubling optics, beam steering optics and optical fiber probe, as well as spectrometer - see Chapters 8 and 17) would also shorten the collection time required to
obtain high resolution spectra. Finally, the application should be kept in mind in order to
determine the appropriate balance to maintain between these variables. In the present
case, a resolution sufficient for spectral identification is required which is perhaps not as
severe a requirement as that necessary to detect spectral shifts caused by changes in
physical environment (e.g. change of solvent, protein denaturation, isotope substitution,
temperature broadening, etc.).

3.6. REFERENCES.


14A-19A.


Neuroscience 12: 407-412.
CHAPTER 4

4.1. INTRODUCTION

The spectroscopic characterization of the individual small-molecule neurotransmitters, the intended environment, and the testing of the probe in vitro with individual neurotransmitters and mixtures with normal Raman spectroscopy, some of the 11 basic goals established in Chapter 1, is the purpose of the research described in Chapter 4.

Measuring the normal Raman spectra of the 10 small-molecule neurotransmitters would determine their degree of uniqueness and hence their identifiability. Considerable difficulty in differentiating the individual neurotransmitter spectra would render the
method unfeasible. The 10 small-molecule neurotransmitters can structurally be subdivided into 3 groups: (i) acetylcholine; (ii) the biogenic amines consisting of dopamine, epinephrine, norepinephrine, serotonin, and histamine; and (iii) the amino acid neurotransmitters aspartate, glutamate, glycine, and γ-amino butyric acid. It was therefore expected that the neurotransmitters belonging to the same group would exhibit similar spectra. These spectra were also required as benchmarks for later measurements with optical fiber probes in the event that fiber-based measurements produced distortions of the spectra. They could therefore aid in probe development. Both regular collection and fiber-based collection normal Raman spectra are reported in this chapter. Band assignments were made for the reference spectra based on tables of characteristic Raman frequencies by Dollish et al. (1974) and Lin-Vien et al. (1991).

As part of the characterization of the intended measuring environment, the normal Raman spectra of cerebrospinal fluid samples were obtained as well. This was necessary in order to ensure that a probe capable of functioning under these conditions was developed. Furthermore, one of these spectra, along with that of a cell culture medium, were used to obtain neurotransmitter difference spectra in biological matrices for use in the signal processing investigations. The matrix, as well as difference spectra, are also reported in this chapter.

4.2. NORMAL RAMAN SPECTRA.
4.2.1. Neurotransmitter spectra and band assignments.

At the outset, while a Raman/resonance Raman system was not available in-house, access to another Raman instrument was obtained to measure the normal Raman spectra of the small-molecule neurotransmitters. They were measured between 2000 cm\(^{-1}\) and 500 cm\(^{-1}\) from the excitation frequency. This range is considered to be effective for the identification of infrared spectra, and, due to the similar information probed by Raman spectroscopy, thus also for Raman spectra. The neurotransmitters were dissolved
in physiological saline (0.9% NaCl) to concentrations between 0.1 M and 0.5 M. Raman scattering was excited with the 488 nm line from an Ar+ laser operating at 200 mW and measured with a 1 m focal length spectrometer (JASCO Model NR-1100, Tokyo, Japan) scanning at 120 cm\(^{-1}\)/min and with slits set at 500 µm. The spectra were digitized at 1 cm\(^{-1}\) intervals resulting in 1501 points per spectrum with approximate signal-to-noise ratios (maximum peak height/noise standard deviation) between 20 and 50. Between 3 and 8 scans of each spectrum were accumulated. The backgrounds of the spectra were removed with a polynomial fitting procedure. More detail about data analysis is given in Part V.

4.2.1.1. Physiological saline.

The spectrum of physiological saline (0.9 % NaCl in water) is shown in Figure 4.1.

![Raman spectrum of physiological saline solution. The prominent peak is due to water deformation.](image)
The spectrum shows only a single prominent feature near 1630 cm$^{-1}$ assigned to H$_2$O deformation. Physiological saline is routinely used as a vehicle for the administration of drugs to living animals and was often used as solvent for neurotransmitters.

4.2.1.2. Acetylcholine.

The Raman spectrum of 0.5 M acetylcholine dissolved in physiological saline is shown in Figure 4.2 and the band assignments of some of the major features of this spectrum are made in Table 4.1. Note the relative intensity of the H$_2$O deformation peak near 1630 cm$^{-1}$.

![Raman Spectrum of Acetylcholine](image)

**Figure 4.2.** The Raman spectrum of acetylcholine (0.5 M) in physiological saline. Inset shows structural formula.
Table 4.1.  
Table 4.1 shows some of the major peaks in the Raman spectrum of acetylcholine (0.5 M), their tentative vibrational band assignments, and relative intensities.

<table>
<thead>
<tr>
<th>Raman shift</th>
<th>Band assignment</th>
<th>Intensity(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>646</td>
<td>O=C=O in plane def.</td>
<td>38</td>
</tr>
<tr>
<td>722</td>
<td>C-N symmetric stretch</td>
<td>100</td>
</tr>
<tr>
<td>840</td>
<td>C-C stretch</td>
<td>16</td>
</tr>
<tr>
<td>877</td>
<td>C-O-C stretch</td>
<td>22</td>
</tr>
<tr>
<td>950</td>
<td>--</td>
<td>23</td>
</tr>
<tr>
<td>1451</td>
<td>CH₃ deformation</td>
<td>25</td>
</tr>
<tr>
<td>1742</td>
<td>C=O stretch</td>
<td>13</td>
</tr>
</tbody>
</table>

4.2.1.3. Dopamine.

The Raman spectrum of 0.5 M dopamine dissolved in physiological saline is shown in Figure 4.3 and the band assignments of some of the major features of this spectrum are made in Table 4.2. A ring stretch vibrational mode at 1617 cm⁻¹ is superimposed on the H₂O deformation peak near 1630 cm⁻¹.

![Figure 4.3. The Raman spectrum of dopamine (0.5 M) in physiological saline. Inset shows structural formula.](image)
Table 4.2.
Some of the major peaks in the Raman spectrum of dopamine (0.5 M), their tentative vibrational band assignments, and relative intensities.

<table>
<thead>
<tr>
<th>Raman shift</th>
<th>Band assignment</th>
<th>Intensity(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>599</td>
<td>ring vibration</td>
<td>22</td>
</tr>
<tr>
<td>717</td>
<td>C-N stretch</td>
<td>33</td>
</tr>
<tr>
<td>756</td>
<td>ring vibration</td>
<td>36</td>
</tr>
<tr>
<td>788</td>
<td>ring vibration</td>
<td>100</td>
</tr>
<tr>
<td>953</td>
<td>C-C stretch</td>
<td>24</td>
</tr>
<tr>
<td>1293</td>
<td>ring vibration</td>
<td>42</td>
</tr>
<tr>
<td>1617</td>
<td>ring stretch</td>
<td>25</td>
</tr>
</tbody>
</table>

4.2.1.4. Epinephrine.

The Raman spectrum of 0.5 M epinephrine dissolved in physiological saline is shown in Figure 4.4. The ring stretch vibrational mode at 1617 cm\(^{-1}\) is again superimposed on the H\(_2\)O deformation peak near 1630 cm\(^{-1}\). The band assignments of some of the major features of this spectrum are made in Table 4.3.

Figure 4.4. The Raman spectrum of epinephrine (0.5 M) in physiological saline. Inset shows structural formula.
Table 4.3.

<table>
<thead>
<tr>
<th>Raman shift</th>
<th>Band assignment</th>
<th>Intensity(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>762</td>
<td>ring vibration</td>
<td>78</td>
</tr>
<tr>
<td>788</td>
<td>ring vibration</td>
<td>100</td>
</tr>
<tr>
<td>944</td>
<td>C-C stretch</td>
<td>28</td>
</tr>
<tr>
<td>1294</td>
<td>ring vibration</td>
<td>74</td>
</tr>
<tr>
<td>1613</td>
<td>ring stretch</td>
<td>46</td>
</tr>
</tbody>
</table>

4.2.1.5. Norepinephrine.

The Raman spectrum of 0.5 M norepinephrine dissolved in physiological saline is shown in Figure 4.5. The ring stretch vibrational mode at 1613 cm\(^{-1}\) is superimposed on the H\(_2\)O deformation peak near 1630 cm\(^{-1}\). Note the similarity to the spectrum of epinephrine in Figure 4.4. The band assignments of some of the major features of this spectrum are made in Table 4.4.

![Structural formula of norepinephrine](image)

Figure 4.5. *The Raman spectrum of norepinephrine (0.5 M) in physiological saline. Inset shows structural formula.*
Table 4.4 shows some of the major peaks in the Raman spectrum of norepinephrine (0.5 M), their tentative vibrational band assignments, and relative intensities.

<table>
<thead>
<tr>
<th>Raman shift</th>
<th>Band assignment</th>
<th>Intensity(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>763</td>
<td>ring vibration</td>
<td>60</td>
</tr>
<tr>
<td>781</td>
<td>ring vibration</td>
<td>100</td>
</tr>
<tr>
<td>1294</td>
<td>ring vibration</td>
<td>58</td>
</tr>
<tr>
<td>1613</td>
<td>ring stretch</td>
<td>46</td>
</tr>
</tbody>
</table>

4.2.1.6. Serotonin.

The Raman spectrum of 0.1 M serotonin dissolved in physiological saline is shown in Figure 4.6 and the band assignments of some of the major features of this spectrum are made in Table 4.5. The peak near 1630 cm\(^{-1}\) is due to H\(_2\)O deformation.

Figure 4.6. The Raman spectrum of serotonin (0.1 M) in physiological saline. Inset shows structural formula.
Table 4.5. 
Table 4.5 shows some of the major peaks in the Raman spectrum of serotonin (0.1 M), their tentative vibrational band assignments, and relative intensities.

<table>
<thead>
<tr>
<th>Raman shift</th>
<th>Band assignment</th>
<th>Intensity(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>762</td>
<td>phenyl ring vibration</td>
<td>79</td>
</tr>
<tr>
<td>829</td>
<td>C-N symm. stretch</td>
<td>46</td>
</tr>
<tr>
<td>937</td>
<td>C-C stretch</td>
<td>86</td>
</tr>
<tr>
<td>1240</td>
<td></td>
<td>47</td>
</tr>
<tr>
<td>1348</td>
<td>indole ring vibration</td>
<td>100</td>
</tr>
<tr>
<td>1435</td>
<td>NH bending of ring</td>
<td>51</td>
</tr>
<tr>
<td>1550</td>
<td>indole ring vibration</td>
<td>72</td>
</tr>
</tbody>
</table>

4.2.1.7. Histamine.

The Raman spectrum of 0.5 M histamine dissolved in physiological saline is shown in Figure 4.7 and the band assignments of some of the major features of this spectrum are made in Table 4.6. The shoulder near 1630 cm⁻¹ is due to H₂O deformation.

![Figure 4.7. The Raman spectrum of histamine (0.5 M) in physiological saline. Inset shows structural formula.](image)
Table 4.6 shows some of the major peaks in the Raman spectrum of histamine (0.5 M), their tentative vibrational band assignments, and relative intensities.

<table>
<thead>
<tr>
<th>Raman shift</th>
<th>Band assignment</th>
<th>Intensity(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>650</td>
<td>--</td>
<td>42</td>
</tr>
<tr>
<td>860</td>
<td>C-N stretch symm.</td>
<td>26</td>
</tr>
<tr>
<td>990</td>
<td>C-C stretch</td>
<td>27</td>
</tr>
<tr>
<td>1010</td>
<td>C-N stretch antisym.</td>
<td>29</td>
</tr>
<tr>
<td>1108</td>
<td>--</td>
<td>26</td>
</tr>
<tr>
<td>1160</td>
<td>ring breathing</td>
<td>46</td>
</tr>
<tr>
<td>1236</td>
<td>--</td>
<td>41</td>
</tr>
<tr>
<td>1272</td>
<td>ring vibration</td>
<td>100</td>
</tr>
<tr>
<td>1310</td>
<td>--</td>
<td>65</td>
</tr>
<tr>
<td>1452</td>
<td>C=N stretch</td>
<td>53</td>
</tr>
<tr>
<td>1495</td>
<td>--</td>
<td>42</td>
</tr>
<tr>
<td>1575</td>
<td>C=C stretch</td>
<td>79</td>
</tr>
</tbody>
</table>

4.2.1.8. Aspartate.

The Raman spectrum of 0.5 M aspartate dissolved in physiological saline is shown in Figure 4.8 and the band assignments of some of the major features of this spectrum are made in Table 4.7. The peak near 1630 cm\(^{-1}\) is due to H\(_2\)O deformation.

Figure 4.8. The Raman spectrum of aspartate (0.5 M) in physiological saline. Inset shows structural formula.
Table 4.7 shows some of the major peaks in the Raman spectrum of aspartate (0.5 M), their tentative vibrational band assignments, and relative intensities.

<table>
<thead>
<tr>
<th>Raman shift</th>
<th>Band assignment</th>
<th>Intensity(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>825</td>
<td>C-C stretch</td>
<td>100</td>
</tr>
<tr>
<td>1738</td>
<td>C=O stretch</td>
<td>85</td>
</tr>
</tbody>
</table>

4.2.1.9. γ-Amino butyric acid.

The Raman spectrum of 0.5 M γ-amino butyric acid dissolved in physiological saline is shown in Figure 4.9 and the band assignments of some of the major features of this spectrum are made in Table 4.8. The peak near 1630 cm\(^{-1}\) is due to H\(_2\)O deformation.

Figure 4.9. The Raman spectrum of γ-amino butyric acid (0.5 M) in physiological saline. Inset shows structural formula.
Table 4.8.
Table 4.8 shows some of the major peaks in the Raman spectrum of γ-amino butyric acid (0.5 M), their tentative vibrational band assignments, and relative intensities.

<table>
<thead>
<tr>
<th>Raman shift</th>
<th>Band assignment</th>
<th>Intensity(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>867</td>
<td>C-C stretch</td>
<td>58</td>
</tr>
<tr>
<td>897</td>
<td>C-C stretch</td>
<td>49</td>
</tr>
<tr>
<td>944</td>
<td>--</td>
<td>32</td>
</tr>
<tr>
<td>976</td>
<td>--</td>
<td>42</td>
</tr>
<tr>
<td>1052</td>
<td>C-N stretch</td>
<td>26</td>
</tr>
<tr>
<td>1314</td>
<td>-(CH₂) twist</td>
<td>58</td>
</tr>
<tr>
<td>1408</td>
<td>CO₂⁻ symm. stretch</td>
<td>100</td>
</tr>
</tbody>
</table>

4.2.1.10. Glutamate.

The Raman spectrum of 0.1 M glutamate dissolved in physiological saline is shown in Figure 4.10 and the band assignments of some of the major features of this spectrum are made in Table 4.9. The peak near 1630 cm⁻¹ is due to H₂O deformation; note the relative intensity.

Figure 4.10. The Raman spectrum of glutamate (0.1 M) in physiological saline. Inset shows structural formula.
Table 4.9.
Table 4.9 shows some of the major peaks in the Raman spectrum of glutamate (0.1 M), their tentative vibrational band assignments, and relative intensities.

<table>
<thead>
<tr>
<th>Raman shift</th>
<th>Band assignment</th>
<th>Intensity(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>938</td>
<td>C-C stretch</td>
<td>100</td>
</tr>
<tr>
<td>1351</td>
<td>CH$_2$ twist</td>
<td>64</td>
</tr>
<tr>
<td>1412</td>
<td>CO$_2^-$ symm. stretch</td>
<td>100</td>
</tr>
</tbody>
</table>

4.2.1.11. Glycine.

The Raman spectrum of 0.5 M glycine dissolved in physiological saline is shown in Figure 4.11 and the band assignments of some of the major features of this spectrum are made in Table 4.10. The peak near 1630 cm$^{-1}$ is due to H$_2$O deformation.

Figure 4.11. The Raman spectrum of glycine (0.5 M) in physiological saline. Inset shows structural formula.

Table 4.10.
Table 4.10 shows some of the major peaks in the Raman spectrum of glycine (0.5 M), their tentative vibrational band assignments, and relative intensities.

<table>
<thead>
<tr>
<th>Raman shift</th>
<th>Band assignment</th>
<th>Intensity(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>873</td>
<td>C-C stretch</td>
<td>100</td>
</tr>
<tr>
<td>1734</td>
<td>C=O stretch</td>
<td>32</td>
</tr>
</tbody>
</table>
4.2.2. Other spectra.

4.2.2.1. Tryptophan.

The spectrum of tryptophan, a precursor of serotonin, was also measured to investigate the correspondence between neurotransmitters and some of their precursors and/or metabolites. This spectrum is given in Figure 4.12 (note the similarity to the serotonin spectrum of Figure 4.6).

![Raman spectrum of tryptophan (0.1 M) in physiological saline.](image)

Figure 4.12. The Raman spectrum of tryptophan (0.1 M) in physiological saline.

4.2.2.2. Cerebrospinal fluid.

CSF was obtained from the lateral ventricle of a male rat under anesthesia (sodium pentobarbitol and ketamine, see Chapter 10) during the implantation of an indwelling cannula as well as removed via the cannula 24 h and 72 h after surgery from the unanesthetized animal. The Raman spectra from these samples were measured and are shown in Figure 4.13. The spectrum taken during surgery differs from those taken after surgery, possibly due to sample contamination with traces of blood hence changing
the background fluorescence (e.g. compare to the spectrum of the cell culture medium containing 5% serum as shown in Figure 4.15).

![Raman spectrum](image)

Figure 4.13. The Raman spectrum of cerebrospinal fluid taken during surgery (solid line), 24 h later (dashed line), and 72 h later (bottom line, dotted) Traces displaced for clarity.

4.2.2.3. Anesthetic.

A Raman spectrum of the anesthetic (see previous paragraph) employed in the cannula implantation procedure mentioned above was also taken for comparison with the spectrum of CSF taken during surgery to determine whether the presence of the anesthetic could be detected in the CSF sample. No sign of the anesthetic was evident in the spectrum of CSF taken during surgery. No attempt was made to dilute the anesthetic to physiological concentrations. The spectrum of the anesthetic is shown in Figure 4.14.
4.2.3. Difference spectra.

The Raman spectra of two biological matrices were measured before and after having been spiked with acetylcholine. The spectrum taken before spiking was then subtracted from that taken after spiking to obtain an acetylcholine difference spectrum. The collection parameters were as described above.

4.2.3.1. Acetylcholine in DMEM.

A Raman spectrum was taken of Dulbecco's modified Eagle medium (DMEM) with 5% calf serum and 0.2% sodium azide (as preservative). DMEM is a complex cell culture medium consisting of several inorganic salts, amino acids, vitamins, and other nutrients. The Raman spectrum of DMEM, shown in Figure 4.15, exhibits a strong fluorescence band. The sample was subsequently spiked with acetylcholine to a concentration of approximately 0.5 M and the Raman spectrum measured again (also
shown in Figure 4.15). The difference spectrum of acetylcholine in DMEM was obtained by subtracting the two spectra and it is shown in Figure 4.15. Figure 4.15 (bottom line) clearly resembles the acetylcholine spectrum shown in Figure 4.2, especially the 722 cm$^{-1}$ and 646 cm$^{-1}$ peaks.

4.2.3.2. Acetylcholine in CSF.

The Raman spectrum of CSF taken during surgery is shown in Figure 4.13. This sample was subsequently spiked with acetylcholine to an approximate concentration of 0.2 M and then remeasured. The spectrum is shown in Figure 4.16 and the difference spectrum of acetylcholine in CSF is shown in the same figure. The difference spectrum can be compared to the spectrum of 0.5 M acetylcholine dissolved in physiological saline shown in Figure 4.2.
Figure 4.16. The Raman spectrum of cerebrospinal fluid taken during surgery (shown in Figure 4.13) spiked with acetylcholine to 0.2 M (top trace), and the difference spectrum (bottom trace).

4.3. FIBER-OPTIC PROBE RAMAN SPECTRA.

The spectra of four neurotransmitters were obtained with a fiber-optic probe. The spectra were excited with the argon ion laser producing 200 mW at 514.5 nm of which 60% was coupled into the probe, thus giving an effective power of 120 mW at the sample, and measured with the single monochromator's entrance slit set at 250 μm. The Raman shift was calculated from the pixel (or diode) number with the following formula:

\[ S = CWN - R(P - CP) \]  

where \( S \) is the Raman shift (cm\(^{-1}\)), \( CWN \) is the monochromator center wavenumber setting (cm\(^{-1}\)), \( P \) is the pixel number, \( CP \) is the number of the center pixel (530, see Chapter 3), and \( R \) is the wavenumber to pixel ratio (cm\(^{-1}\)/pixel) at the center wavenumber based on the instrument's dispersion (0.031 nm/pixel, see Chapter 3).

The probe used consisted of a small excitation fiber (core \( \phi = 100 \) μm) for the transmission of the excitation radiation and a larger collection fiber (core \( \phi = 300 \) μm) for the collection of Raman scattering. The tips of both fibers, obtained from Fiberguide
Industries (Superguide G for UV-Visible transmission), were cleaved perpendicular to
the fiber axis (courtesy of the Departments of Physics and Electrical Engineering, The
University of British Columbia) and aligned flush under a microscope. They were then
joined together with an epoxy resin. The collection fiber was then connected to the
spectrometer via a standard SMA termination and a fiber optic adapter. The excitation
frequency was coupled into the other cleaved end of the excitation fiber with a special
laser-to-fiber coupler (Newport Model F-915 with M-10X visible objective, UV
objective U-27X). The probes were then inserted into disposable glass pipettes for
protection before being immersed into small test tubes (φ = 5 mm, h = 50 mm)
containing the analyte(s). The development, design, and construction of fiber-optic
probes are discussed in detail in Part III.

4.3.1. Neurotransmitter spectra.

4.3.1.1. Acetylcholine.

The spectrum of acetylcholine in physiological saline is shown in Figure 4.17.

![Raman spectrum of acetylcholine](image)

**Figure 4.17. The Raman spectrum of acetylcholine dissolved in physiological saline (0.5
M) obtained with a fiber-optic probe.**
The spectrum consists 20 accumulations of 1 s each. The monochromator center wavenumber setting was at 788 cm\(^{-1}\) from the excitation line. A comparison with Figure 4.2. reveals a strong correspondence between the spectra and enables peak identification.

4.3.1.2. Dopamine.

The spectrum of 0.5 M dopamine in physiological saline, obtained with a fiber optic probe is shown in Figure 4.18. The monochromator center wavenumber setting was at 788 cm\(^{-1}\) from the excitation line. For this spectrum, 30 accumulations of 1 s each was used. A comparison to Figure 4.3 reveals a general correspondence.

![Figure 4.18. The Raman spectrum of dopamine dissolved in physiological saline (0.5 M) obtained with a fiber-optic probe.](image)

4.3.1.3. Serotonin.

The spectrum of 0.5 M serotonin in physiological saline, obtained with a fiber optic probe, is shown in Figure 4.19. The monochromator center wavenumber setting
was at 993 cm\(^{-1}\) from the excitation line. For this spectrum, 20 accumulations of 1 s each was used. Note the correspondence to Figure 4.6.

![Raman Spectrum of Serotonin](image)

**Figure 4.19.** The Raman spectrum of serotonin dissolved in physiological saline (0.1 M) obtained with a fiber-optic probe.

4.3.1.4. γ-Amino butyric acid.

The spectrum of 0.5 M γ-amino butyric acid in physiological saline, obtained with a fiber optic probe, is shown in Figure 4.20. The monochromator center wavenumber setting was at 1401 cm\(^{-1}\) from the excitation line. For this spectrum, 30 accumulations of 1 s each was used. This compound is a weak Raman scatterer and it is difficult to make any peak assignments with certainty from this spectrum (the peak near 1400 cm\(^{-1}\) corresponds to that in Figure 4.9). This figure demonstrates the need for a good rejection filter to reduce the very high background counts and a high throughput instrument to conserve all the collected Raman photons.
Figure 4.20. The Raman spectrum of \( \gamma \)-amino butyric acid dissolved in physiological saline (0.5 M) obtained with a fiber-optic probe.

4.3.1.5. Mixture spectra.

The spectrum of a mixture of 0.5 M acetylcholine and 0.5 M serotonin in water, obtained with a fiber optic probe, is shown in Figure 4.21. The monochromator center wavenumber setting was at 849 cm\(^{-1}\) from the excitation line. For this spectrum, 30 accumulations of 1 s each was used. The acetylcholine peak near 720 cm\(^{-1}\) is indicated. A comparison to Figures 4.17 and 4.19 reveal that both serotonin and acetylcholine peaks can be identified.
The tryptophan spectrum, shown in Figure 4.12, was very similar to that of serotonin shown in Figure 4.6. In addition, the spectra of epinephrine and norepinephrine, shown in Figures 4.4 and 4.5, were very similar as pointed out previously. These results suggested that good quality spectra of fairly high resolution were needed to differentiate between some compounds. It may therefore be fruitful to extend the spectral region to 4000 cm\(^{-1}\). For example, the bonded symmetric and antisymmetric NH\(_2\) stretches of primary amines, as well as the bonded NH stretch of secondary amines, occur at different points in the 3250-3400 cm\(^{-1}\) spectral region (Dollish et al., 1974) and are of moderate strength. Therefore, epinephrine, differing from norepinephrine in being a secondary amine and not a primary amine, may perhaps be more effectively distinguished based on information contained in this region (barring interference from the water bands).
However, in general, the 10 small-molecule neurotransmitters showed clearly distinguishable Raman spectra. These were also all different from the Raman spectrum of water. This finding suggested that aqueous neurotransmitter solutions could be differentiated and that the neurotransmitters could be identified on the basis of their Raman spectra. The same procedure should be applied to determine the differentiability of the neurotransmitters based on their resonance Raman spectra. Literature reports, however, indicate that chemical species can be differentiated on the basis of their resonance Raman spectra (Asher et al., 1986; Johnson et al., 1984).

The Raman spectra obtained with the fiber-optic probes were generally similar to those obtained with the scanning spectrometer. They differed, however, in some respects, such as showing different baselines, peak heights, and peak positions. These differences reflect most likely differences in sample holder geometry (e.g. producing different backgrounds), detectors (e.g. varying frequency responses producing different peak heights), and light coupling into the spectrometer. On the latter point, it should be noted that a custom-made optical fiber holder was used to couple the collection fiber of the probe to the spectrograph. This sometimes resulted in small misalignments of the input beam with the spectrograph optics, thus producing shifts in peak position. An aqueous standard solution with 3 or more prominent peaks of known Raman shift could be used to calibrate the instrument when commencing fiber-optic probe data collection to correct for such peak shifts.

The spectra of the CSF samples did not show any prominent Raman features, nor could any sign of the anesthetic be detected in the Raman spectrum of CSF obtained during surgery. However, the latter sample did show a different fluorescence background from CSF samples taken 24 h and 72 h after surgery, possibly due to contamination with traces of blood.

Taken together, these results indicated that Raman spectra could be obtained via optical fiber and that neurotransmitters could be differentiated based on Raman spectra
obtained in this way. The Raman spectra of samples of CSF and a binary neurotransmitter mixture were also measured successfully. It is important to realize that the spectra presented in this chapter were obtained with probes still under development. Substantially better spectra (resolution and signal-to-noise ratios) could be obtained with current probes and state-of-the-art instrumentation.

4.5. REFERENCES.


CHAPTER 5

5.1. INTRODUCTION
5.2. RESONANCE RAMAN SPECTRA
  5.2.1. Methyl orange at different visible frequencies
  5.2.2. Different concentrations of methyl orange
  5.2.3. Differentiation between methyl orange and chromate
5.3. UV ABSORPTION SPECTRA
  5.3.1. Methyl orange
  5.3.2. Dopamine and epinephrine
  5.3.3. Other neurotransmitters
  5.3.4. CSF
5.4. UV RESONANCE RAMAN SPECTRA
  5.4.2. Other
5.5. DISCUSSION
5.6. REFERENCES

5.1. INTRODUCTION

The determination of the viability of fiber-optic based resonance Raman spectroscopy for obtaining selectivity and sensitivity was of crucial importance in the development of this neuroprobe. The purpose of the work reported in Chapter 5 was to determine the viability of such an approach by demonstrating both resonance enhancement and resonance-based signal discrimination using spectra collected via optical fiber probe. Enhancement of the signal of a given analyte can be obtained by tuning to its resonant frequency(ies) and discrimination by tuning to the resonant frequency of a chromophore of a particular analyte to enhance its signal relative to the chromophore(s) of the same or other analyte(s) (see Figure 2.2.). It was furthermore necessary to demonstrate that UV resonance Raman spectra of neurotransmitters could be excited and collected with a fiber-optic probe. In order to determine the excitation frequencies for effective discrimination, the UV absorption spectra of some neurotransmitters were measured and others were obtained from the literature. Finally, the UV absorption spectrum of a CSF sample was obtained as part of the characterization of the intended measuring environment. The reader may wish to refer to Chapter 2 for the theoretical background to the present chapter.
5.2. RESONANCE RAMAN SPECTRA

The resonance Raman spectra of methyl orange and potassium chromate were measured with a fiber-optic probe. Both compounds were readily available, water soluble, and could be excited at resonance using visible and near UV cw radiation from the Ar$^+$ laser. Methyl orange, an acid-base indicator, shows resonance enhancement in the visible (see Long, 1977) and chromate in the near UV (Kiefer and Bernstein, 1972) regions. Using cw for resonance enhancement was less demanding than using tunable pulsed UV for excitation, and different probe designs (see Part III) could therefore be evaluated and problems resolved with comparative ease. For the initial feasibility studies reported in this chapter, the same probe as that described in Chapter 4 was used.

5.2.1. Different concentrations of methyl orange.

The Raman spectra of different concentrations of methyl orange ($1\times10^{-3}$, $1\times10^{-4}$, $1\times10^{-5}$, and $1\times10^{-6}$ M) were measured at 514.7 nm, 100 mW at the sample, single monochromator arrangement with entrance slit set at 250 μm, except for the two weakest concentrations (500 μm), and 989 ms integration. The spectra are shown in Figure 5.1. These spectra revealed, as is evident from Figure 5.2, that self absorption by the methyl orange attenuated the signal at the higher concentration. This experiment indicated that a methyl orange concentration of about $1\times10^{-5}$ M could be used for doing resonance Raman investigations since a stronger signal could be expected at resonance frequencies.
5.2.2. Methyl orange at different visible frequencies.

A methyl orange concentration of $1 \times 10^{-5} \text{ M}$ in water was excited with the
different visible frequencies of an Ar$^+$ laser (514.7, 496.5, 488, 476.5, 472.7, 465, and
457.9 nm) to determine the frequency dependence of the resonance enhancement. The
integration time used was 989 ms, power at the sample was 100 mW, and the single
monochromator's entrance slit was set at 250 μm. Figure 5.1 shows the Raman spectrum
of methyl orange excited with 457.9, 472.7, and 488 nm, respectively. The resonance
enhancement of the unresolved doublet near 1400 cm$^{-1}$ attained a maximum with
excitation at 472.7 nm. The enhancement at 488 nm is likely a pre-resonance effect.
Figure 5.2. Spectra of methyl orange (1x10^{-5} M) excited at different visible frequencies: (a) 488 nm; (b) 472.7 nm; and (c) 457.9 nm. Resonance enhancement occurred with excitation at 472.7 nm. The traces are offset for clarity.

5.2.3. Differentiation between methyl orange and chromate.

When excited near 370 nm, potassium chromate shows resonance enhancement (Kiefer and Bernstein, 1972) while methyl orange shows a maximum resonance enhancement near the Ar^{+}-line at 472.7 nm. Using the probe with 472.7 nm excitation, 130 mW at the sample, and a total exposure time of 30 s in a mixture of 20x10^{-6} methyl orange and 200x10^{-6} potassium chromate in water, the resonance enhanced methyl orange peak near 1400 cm^{-1} can clearly be seen while the potassium chromate (resonance near 850 cm^{-1}) cannot be detected and vice versa for excitation at 364 nm, 40 mW at the sample and 10 s integration time. These spectra, shown in Figure 5.3, demonstrated the usefulness of frequency tuning to discriminate between molecular species.
5.3. UV ABSORPTION SPECTRA.

The UV absorption spectra of some neurotransmitters were obtained to determine their optimum excitation frequencies as well as their molecular extinction coefficients in order to establish the optimum concentrations for fiber-optic probe detection. In addition, the UV absorption spectrum of cerebrospinal fluid was measured to determine how absorbing the biological matrix in which the neurotransmitters were imbedded was. The molecular extinction coefficients of methyl orange were also determined and could be compared to those of the neurotransmitters to give an indication of their limits of detection.

5.3.1. Methyl orange.

The absorbance of $1 \times 10^{-5} \text{ M}$ methyl orange was measured with a Varian DMS 200 UV visible spectrophotometer (Australia) using a path length of 1 cm and scanning
speed of 100 nm/min.). The absorbance at the Raman excitation frequency was 0.254 and at the Raman collection frequency less than 0.181.

5.3.2. Some aromatic neurotransmitters and amino acids.

The UV absorption spectra of dopamine and epinephrine were obtained as above. These spectra are shown in Figure 5.4. It is evident from Figure 5.4 that the optimum frequencies for resonance enhancement are near 280 nm and below 220 nm. The absorption spectra for the amino acid neurotransmitter precursors tryptophan and tyrosine are displayed in the same figure showing their maxima near 220 nm.

![Graph of UV absorption spectra](image)

Figure 5.4. The UV absorption spectra of $3.3 \times 10^{-5}$ M dopamine (solid line), $3.8 \times 10^{-5}$ M epinephrine (dotted line), $108 \times 10^{-5}$ M tryptophan (dashed line), and $110 \times 10^{-5}$ M tyrosine (dash/dot). All analytes were dissolved in water.
5.3.3. Other neurotransmitters.

The UV absorption spectra of most of the 10 small-molecule neurotransmitters were obtained from the literature (see Lang, 1975) and compiled in Figure 5.5. The figure shows that the aromatic neurotransmitters in general have two absorption maxima in the UV, but that the aliphatic neurotransmitters have monotonically increasing absorption curves. This indicated a possibility of differentiation between these two groups of compounds.

Figure 5.5. A compilation of the UV absorption spectra of some of the 10 small-molecule neurotransmitters: (a) dopamine, epinephrine, norepinephrine, serotonin; (b) histamine; (c) glutamate; (d) aspartate; (e) acetylcholine; (f) glycine.

5.3.4. Cerebrospinal fluid.

The UV absorption spectrum (190 to 400 nm) of cerebrospinal fluid obtained from the lateral ventricle of a male rat 72 h after surgery and diluted approximately 20 times with deionized water. The spectrum was obtained with a Hewlett Packard 8452A diode array spectrophotometer (distributor: Mandel Scientific Company, Guelph, Ont.)
using a path length of 1 mm and integration time of 1 s. The spectrum, shown in Figure 5.6, indicates that CSF is highly absorbing in the ultraviolet but exhibits no features other than the tail of a possible Gaussian band absorption profile below 200 nm.

![Absorption spectrum](image)

Figure 5.6. *The UV absorption spectrum of cerebrospinal fluid diluted approximately 20 times with deionized water.*

5.4. UV RESONANCE RAMAN SPECTRA.

The UV resonance Raman spectra of the aromatic neurotransmitters dopamine, epinephrine, and serotonin, the aliphatic neurotransmitter γ-amino butyric acid, as well as some neurotransmitter precursors (tyrosine, tryptophan) and a metabolite (melatonin) were measured with an angled fiber-optic probe (see Chapter 8). The doubled output from a tunable dye laser pumped with a Nd:YAG laser was used for excitation and the collected light frequency decoded with a single grating spectrograph (see Chapter 3).
5.4.1. Neurotransmitters.

The resonance Raman spectrum of dopamine dissolved in distilled water is shown in Figure 5.7. The absorbance of $3.3 \times 10^{-6} \text{ M}$ dopamine at 225 nm was measured as 0.293, which is equivalent to 0.879 at a concentration of $1 \times 10^{-5} \text{ M}$. The angled probe similar to the one used for this experiment showed a working curve optimum at a medium absorbance corresponding to a methyl orange concentration of $1 \times 10^{-5} \text{ M}$ (see Chapter 8 for working curves). Beer’s law states that:

$$A = \varepsilon bc$$  \hspace{1cm} (5.1)

where $A$ is the absorbance of the solution, $\varepsilon$ is the molar extinction coefficient ($\text{cm}^{-1} \text{ mol}^{-1} \text{ L}$), $b$ is the optical pathlength (cm), and $c$ is the analyte concentration (mol L$^{-1}$).

![Figure 5.7. The resonance Raman spectrum of $1 \times 10^{-3} \text{ M}$ dopamine (dotted line) dissolved in distilled water obtained at 227 nm, 20 Hz, 90 s exposure time, a monochromator setting of 235.3 nm and with $\approx 17 \text{ mJ/pulse}$. The spectrum of the solvent only (solid line) and obtained under the same conditions.](image)
It follows from Equation 5.1 that for equal absorbances:

\[(\varepsilon \beta c)_{\text{dopamine}} = (\varepsilon \beta c)_{\text{methyl orange}} \quad (5.2)\]

\[(0.879 \text{cm}^{-1} \text{M}^{-1} \times 1 \text{cm} \times \text{concentration M})_{\text{dopamine}} = (0.254 \text{cm}^{-1} \text{M}^{-1} \times 1 \text{cm} \times 1 \times 10^{-5} \text{M})_{\text{methyl orange}} \quad (5.3)\]

Therefore, a dopamine concentration of 2.89 x 10^{-6} M was estimated to give the maximum signal. However, an initial concentration of 1 x 10^{-3} M was used and successively diluted. In this manner, three resonance Raman spectra of dopamine dissolved in distilled water were obtained, the lowest being 200 x 10^{-6} M.

The ring stretch near 1617 cm\(^{-1}\) (see Figure 4.3) was resonance enhanced at 227 nm. Consequently, the resonant peak of dopamine was superimposed on a normal Raman peak of the solvent (water). Two water spectra were obtained independently under the same conditions as the three dopamine spectra mentioned above and all 5 spectra were smoothed with a 19-point fast Fourier transform. Thereafter, the maxima of the spectra were obtained to provide a mean and standard error for water (1641.0 ± 2.6) and dopamine (1629.0 ± 1.6), respectively. An independent t-test indicated that these values were significantly different (t = -4.167, p = 0.025).

The spectra for epinephrine, serotonin, and histamine are shown in Figures 5.8, 5.9, and 5.10 respectively.
Figure 5.8. The UV resonance Raman spectrum of $1 \times 10^{-3}$ M epinephrine (dotted line) dissolved in distilled water measured at 227 nm, 20 Hz, 40 s exposure time, a monochromator setting of 234.8 nm and with $\approx 17$ mJ/pulse. The spectrum of the solvent only taken under the same conditions (solid line).

Figure 5.9. The UV resonance Raman spectrum of $1 \times 10^{-4}$ M serotonin dissolved in distilled water obtained at 227 nm, 20 Hz, 90 s exposure time, a monochromator setting of 234.3 nm and with $\approx 17$ mJ/pulse.
Figure 5.10. The UV resonance Raman spectrum of 1x10^{-3} M histamine dissolved in distilled water obtained at 227 nm, 20 Hz, 90 s exposure time, a monochromator setting of 234.3 nm and with = 20 mJ/pulse.

The aliphatic (amino acid) neurotransmitter γ-amino butyric acid was not expected to show resonance at 227 nm. The spectrum of 1X10^{-3} M γ-amino butyric acid showed no features other than that produced by the solvent. When the solvent spectrum was subtracted, analyte Raman signals could still not be discerned, thus suggesting longer collection times and/or higher analyte concentrations and/or higher average laser powers were needed with the probe used here.

5.4.2. Other.

Tyrosine is a precursor to the monoamine neurotransmitters (e.g. dopamine), tryptophan a precursor to serotonin, an indole amine, and melatonin a metabolite of serotonin (see Ganong, 1991). As mentioned in the introduction to this chapter, an important issue is the ability to differentiate between related substances on the basis of
their spectra. The measurement of some precursor and metabolite spectra was expected to give an indication of the difficulty that could be encountered in this regard. The UV resonance Raman spectra of tyrosine, tryptophan, and melatonin are shown in Figure 5.11.

![Raman Spectra](image)

**Figure 5.11.** The UV resonance Raman spectrum of $50 \times 10^{-6}$ M melatonin (top), $820 \times 10^{-6}$ M tryptophan (center), and $100 \times 10^{-6}$ M tyrosine, all dissolved in distilled water and obtained at 227 nm, 20 Hz, 40 - 90 s exposure time, monochromator settings of 234.3 - 235.3 nm and with 18 - 24 mJ/pulse.

### 5.5. Discussion

The work reported in this chapter demonstrated that deep ultraviolet resonance Raman spectra could be excited and collected with fiber-optic probes. Furthermore, that components in a multicomponent sample could be differentially enhanced by tuning to a specific component's resonant frequency. This established the viability of using fiber-optic based tunable resonance Raman spectroscopy to increase the sensitivity of the method and to obtain selectivity for neurotransmitter applications. The use of ultraviolet
excitation at different frequencies to discriminate between different amino acid species in a mixture as well as different amino acid residues in a protein has also been reported in the literature (Johnson et al., 1984).

The absorption spectra obtained allowed the determination of the optimal excitation frequencies and were in general agreement with those presented in the literature for related substances (e.g. precursors in Rava and Spiro, 1985). For the biogenic amines, two maxima appear: one near 275 nm and a more pronounced one near 200 nm. Based on these measurements and the working curve for a 300x600 probe, the optimal concentrations of some neurotransmitters for detection with a fiber-optic probe could be calculated for excitation at 225 - 227 nm. This optimum concentration was approximately 3x10^{-6} M. However, due to the fact that a different probe was needed for pulsed ultraviolet resonance Raman work (see Chapters 7 and 8), the actual measured spectra ranged in concentration from 50x10^{-6} to 1000x10^{-6} M. Using shorter wavelength excitation could produce further resonance enhancement for some compounds, thus reducing the measurable concentrations. For instance, in the tryptophan and serotonin spectra the ring breathing band near 1000 cm\(^{-1}\) should be maximally enhanced and be the most pronounced in the spectrum. Rava and Spiro (1985) reported a tenfold enhancement in the benzene and pyrrole out-of-phase ring breathing band at 1016 cm\(^{-1}\) relative to the water band at 3400 cm\(^{-1}\) when exciting at 218 nm instead of 240 nm.

Two important issues brought to the fore by the work presented in this chapter are those of selectivity and sensitivity. It is clear from the spectra of dopamine (Figure 5.7), epinephrine (Figure 5.8), tyrosine (Figure 5.11), and norepinephrine, that the single resonant enhanced ring vibration peak near 1600 cm\(^{-1}\) in all of these spectra will render discrimination between these species difficult. A possible way to discriminate between such similar spectra has been demonstrated with the dopamine and water spectra, but there is no doubt that an improved system resolution (and an increased system dispersion) would aid in discrimination. Figure 5.12 shows Lorentzian curves fitted to
the measured spectra of dopamine, epinephrine and norepinephrine and illustrates the small shifts between the peaks of these three neurotransmitters. Given the importance of discriminating between neurotransmitters based on their Raman and resonance Raman spectra, the issue of discrimination receives additional attention in Chapters 13 and 16.

![Lorentzian curves fitted to the measured spectra of dopamine, epinephrine and norepinephrine indicating peak positions.](image)

**Figure 5.13. Lorentzian curves fitted to the measured spectra of dopamine, epinephrine and norepinephrine indicating peak positions.**

Although it is also clear from the resonance spectra that discrimination between neurotransmitters can be obtained (e.g. signals for dopamine but not γ-aminobutyric acid), the detection of aliphatic neurotransmitters now appears to be more problematic. Using higher frequencies for excitation will not only increase the Raman signals for the aliphatic neurotransmitters, but simultaneously move into a region where the analyte matrix becomes more absorbing, thus potentially negating any advantages accruing from such signal increases. However, this problem is mitigated by the fact that not all neurotransmitters are present in equal amounts throughout the brain. Furthermore, in
accordance with the axioms stated in Chapter 1, neurotransmitters are assumed not all released at the same rate and time. Finally, an increase in probe sensitivity, which is continually being improved, may make the detection of aliphatic neurotransmitters, nevertheless likely (see also Chapter 17). These neurotransmitters tend to be present in relatively high concentrations in cerebrospinal fluid (e.g. Roettger and Goldfinger, 1991).

5.6. REFERENCES.


PART III

Fiber-optics

CHAPTER 6

6.1. INTRODUCTION
6.2. FIBER-OPTICS
6.2.1. Light propagation
6.2.2. Light coupling
6.2.3. Light delivery efficiency
6.3. REMOTE RAMAN SPECTROSCOPY
6.3.1. In vivo Raman spectroscopy
6.4. REFERENCES

6.1. INTRODUCTION

The focus of this multidisciplinary thesis now shifts from spectroscopy to fiber-optics. The work reported in Part II indicated that the normal and resonance Raman spectra of neurotransmitters could be used for their identification and differentiation. These findings were extended to spectra obtained via optical fiber probes and the work reported in Part III details the development of these probes. Chapter 6 establishes the theoretical basis for Part III, while Chapters 7 and 8 contain experimental findings.

The work presented in Part III involved the characterization of the optical fibers for continuous wave and pulsed beam delivery and included investigation of the parameters affecting optical damage, transmission efficiency, and coupling efficiency. The effectiveness of different probe geometries in different sampling environments were investigated, using computer models and experimental measurements, in order to determine the most efficient probe geometry for in vivo work. Some of the characteristics of the in vivo environment were established in Part II. Part III also contains a description of the probe fabrication procedures and the present introductory and reference chapter which provides a brief introduction to fiber optics and a brief discussion of remote Raman spectroscopy.
6.2. Fiber-optics

Optical fibers are cylindrical dielectric waveguides made from a variety of light-transmitting materials (e.g. plastic, glasses, silica, and sapphire). Step index fibers have a central core surrounded by a concentric cladding with a refractive index (the ratio of the velocity of light in the medium to that in vacuo) lower than that of the core (e.g. Beck at al., 1993; Davis et al., 1995; Heiman et al, 1989). The core often consists of the same material as the cladding, but dopants such as the oxides of germanium or phosphorous are added to effect a variation in the refractive index. A jacketing material (e.g. acrylate, nylon) covers the cladding to protect the fiber. Core diameters of multimode fibers generally range from 100 to 1000 µm. Craig-Ryan and Ainslie (1991) give a description of several optical fiber fabrication methods. Manufacturing processes and the degree to which control could be exercised over the dopant concentrations in core and cladding can affect the fiber's light transmission properties.

6.2.1. Light propagation.

When a light ray moves from an optical material with high refractive index to an optical material with lower refractive index incident at a large angle relative to the boundary between the two materials, part of the ray is reflected back into the material with high refractive index, and the remainder of the ray is refracted or bent into the material with the lower refractive index. When this angle of incidence is reduced, more of the ray is reflected, less is refracted and the angle of refraction is reduced. If the angle of incidence is reduced further to reach the critical angle, the refracted ray becomes parallel to the boundary between the optical materials and its intensity reduces to zero. Therefore, at angles at or below the critical angle (β in Figure 6.1), all the power of the original ray is retained in the reflected ray. Light propagation in optical fibers depends on
this principle of total internal reflection based on Snell’s laws of reflection and refraction (e.g. Davis et al., 1995; Newport Catalog, 1993). For a silica fiber with core refractive index $n_1 = 1.472$ and cladding refractive index $n_2 = 1.458$, the critical angle with respect to the core-cladding interface is about $8^\circ$.

Depending on the coupling conditions, light can be propagated inside a fiber along a limited number of different characteristic trajectories called modes. For instance, a ray of light entering the fiber along the fiber axis may not bounce off the cladding at all on its path through the fiber. Conversely, a ray entering the fiber close to the boundary of the acceptance cone will frequently bounce and will take longer to traverse its path. A ray on a helical path would bounce even more and take even longer to complete its path. Those fibers with very small cores that allow only one mode of propagation are called single mode fibers while multi-mode fibers have larger cores and allow several modes of propagation.

6.2.2. Light coupling.

Before light can be propagated down an optical fiber, it has to be coupled from a source into the fiber. At the endface between the core and the air where light is coupled from the source into the fiber, there is another critical angle. Light entering the fiber from within the cone described by the critical half-angle, will be propagated along the fiber by total internal reflection as described above. Light not entering the fiber from within this cone will be coupled into the cladding and eventually lost. The numerical aperture (NA) of an optical fiber is defined as the sine of this critical half-angle ($\alpha$ in Figure 6.1). The NA depends on the refractive indices ($n_1, n_2$) of the materials involved and can be approximated by (Davis et al., 1995):

$$NA = (2n_1(n_1-n_2))^{0.5} \quad (6.1).$$
6.2.3. Light delivery efficiency.

Several factors pertaining to the fiber affect the efficiency with which light can be coupled into and propagated by an optical fiber. In addition, transmission properties are also strongly dependent on the output power and the beam qualities of the laser (Beck et al., 1993). Transmission curves for glass and silica are shown in Figure 6.2.

Losses occur at the input and output endfaces of the fiber due to Fresnel reflections from the endfaces. Specular reflections from dust and other irregularities of the fiber surface (e.g. due to polishing) will also contribute to losses and may cause surface damage (Allison et al., 1987). Lossless coupling of light into a fiber is possible when two conditions are met: the input beam diameter should be less than the fiber core diameter, and the convergence half-angle of the input beam should be less than the critical half-angle (Beck et al., 1993).

Light is lost in transmission at microbends in the fiber where the fiber core axis is misaligned with the rest of the fiber due to imperfect fabrication processes and/or
pressure against rough surfaces. Furthermore, fixed radius bends in the fiber also contribute to transmission losses (Davis et al., 1995). Additional attenuation occurs due to light absorption by the core material itself and/or impurities therein and due to Rayleigh scattering from small variations in the density of the core material (Davis et al., 1995). High intensity laser light, likely to occur in pulsed systems, can also give rise to stimulated Raman scattering by the core material. Stimulated Raman scattering is proportional to the square of the excitation frequency in the ultraviolet region and is responsible for additional absorption losses (Davis et al., 1995). The attenuation of light through a glass or silica optical fiber is given by the equation:

\[- \frac{\text{d}I}{\text{d}z} = (\alpha_0 + \Delta\alpha_d)I + \beta I^2\]

where I is the light intensity (W), z is the fiber length (m), \(\alpha_0\) is the linear attenuation coefficient (m\(^{-1}\)), \(\Delta\alpha_d\) is a length-dependent UV-induced attenuation term (m\(^{-1}\)), and \(\beta\) is the non-linear attenuation coefficient (m\(^{-1}\) W\(^{-1}\)) (e.g. Karlitschek et al., 1996).

Figure 6.2. Transmission curves of fused silica (a) and glass (b), redrawn from the Newport Catalog (1993).
6.3. REMOTE RAMAN SPECTROSCOPY.

The use of optical fibers for the relay of optical signals has already been mentioned in Chapter 1 with the discussion of optrodes (section 1.3.2.2). Several researchers have used optical fibers to excite and collect Raman and surface enhanced Raman scattering (e.g. Alarie et al., 1992; Vo-Dinh et al., 1990; Wang et al., 1992). Remote Raman spectroscopy offers the advantages of collecting Raman spectra in situ where the sample is a great distance from the spectrometer, in a hostile environment, or otherwise not easily accessible (e.g. Heiman et al., 1989). Remote Raman spectroscopy can be performed on a variety of solid, liquid, and gas samples, requires no alignment of the sample with input/output optics, can be performed with high efficiency, and often does not involve a focused beam, thus reducing the likelihood of sample damage (Schwab and McCreery, 1984). Finally, due to the small diameter of optical fibers, diminutive (less than 1 mm) optical fiber probes can be constructed (e.g. Schwab and McCreery, 1984; Zhu and Yappert, 1992).

6.3.1. In vivo Raman spectroscopy.

The advantages listed above render remote Raman spectroscopy eminently suitable for performing in vivo neurotransmitter analysis. Surface enhanced Raman scattering has been shown capable of detecting stimulated release neurotransmitter concentrations in vitro (McGlashen et al., 1990; Morris et al., 1990). Work on resonance Raman spectroscopy reported in Chapter 5 (and also Chapter 8) demonstrated that micromolar concentrations of the catecholamines can be detected with relatively simple optical fiber probes, low average power levels, and less-than-ideal equipment. If current state-of-the-art laser and spectrographic equipment, as well as an optimized probe, are used, detection limits within the range of stimulated release concentrations (e.g. Wood et al., 1992) are possible. Furthermore, ultraviolet resonance Raman spectroscopy does not
suffer from the kinetic and biocompatibility limitations associated with surface enhanced Raman scattering. This, in combination with continued advances in laser, fiber-optic, and detector technology should make the prospects for in vivo resonance Raman spectroscopy ever more accessible.

6.4. REFERENCES.


CHAPTER 7

7.1. INTRODUCTION

As mentioned in the preceding chapter, optical fibers have differences in light propagation abilities and mechanical attributes. Some fibers such as sapphire fibers have high laser damage thresholds but also high attenuation of deep UV wavelengths (Chang et al., 1995). Silica fibers show good transmission in the deep UV depending on the cladding material used (Krohn and McCann, 1995) and the OH⁻ content of the core (Taylor et al., 1988). In general, the different fiber characteristics depend on the core and cladding materials used, the dopants employed for refractive index adjustments, and the manufacturing processes used and so vary from manufacturer to manufacturer. Furthermore, as also mentioned in the preceding chapter, light coupling into optical fibers require proper alignment to ensure the matching of the light to the critical angle of the fiber's acceptance cone. Therefore, the fiber type and fiber quality as well as the fiber coupling characteristics dictate the beam delivery efficiency of the fiber (Beck et al., 1993).

The factors influencing the coupling efficiency in and the optical damage to optical fibers proposed for the development of fiber optic probes therefore had to be investigated. Furthermore, the optical fibers from different manufacturers had to be
characterized with regard to their transmission qualities to enable the selection of the most suitable fiber for building an efficient pulsed UV resonance Raman neuroprobe.

7.2. OPTICAL FIBER CHARACTERIZATION.

7.2.1. Input coupling

According to Beck et al., (1993) nearly lossless coupling of light into a fiber is possible when the input beam diameter is less than the fiber core diameter and the convergence half-angle of the input beam is less than the critical half-angle of the acceptance cone. According to De Hart (1992) the beam should never be focused inside the fiber (in order to avoid damage to the fiber). These requirements therefore establish the conditions for the efficient coupling of light into optical fibers.

7.2.1.1. Continuous wave excitation.

Continuous-wave light was coupled into fibers with a Newport fiber coupler and a microscope objective. For visible excitation a 0.25 N.A. microscope objective with focal length of 14.8 mm was used and for UV excitation a 0.13 N.A. objective with 5.77 mm focal length was used. Coupling efficiencies were regularly measured to be 60-80% for visible wavelengths and about 20% in the near UV. These could be improved with better positioning equipment and/or a higher quality laser beam.

7.2.1.2. Pulsed excitation.

The microscope objective could not be used for pulsed excitation because the short focal length (5.77 mm) caused an air plasma formation at higher peak energies (about 2 mJ/pulse) resulting in poor throughput of 9-14%. The effects of an air plasma formation on throughput are well demonstrated in Figure 3.2.
The extent to which the beam waist of a longer focal length (150 mm) lens induced losses was examined by measuring the pulse energy at various oscillator energy settings before and after the lens. The results, shown in Figure 7.1, indicated that no appreciable losses occurred. The use of a 150 mm focal length lens resulted in a coupling efficiency of 30-50% when focused in front of the fiber with neither significant under- nor over-illumination of the fiber face.

![Figure 7.1](image.png)

Figure 7.1. The pulse energy (mean ± S.E.M.; n = 10) as a function of oscillator setting when measured at 266 nm and 10 Hz before (solid line) and after (dotted line) a 150 mm focal length lens.

In order to determine the best position for placing the beam waist (i.e. verifying De Hart, 1992, for the current application), the throughput of a silica clad 15 cm section of 3M UER fiber of 300 μm diameter (3M Specialty Optical Fibers, West Haven, CT) was measured with the focal point alternately 24 mm in front of the fiber input face (position A) or 24 mm behind it (i.e. focusing inside the fiber - position B). The energy output
from the fiber was measured alternating between positions A and B as shown in Table 7.1. This procedure was repeated at oscillator energy settings of 620, 630, and 640 respectively. A 150 mm focal length lens was used for coupling light into the fiber.

At the 640 setting, all the readings at the B position were significantly elevated with respect to those of the A position (paired t-test on means 45.5 and 53.1, df = 5, p = 0.003). The same trends held for lower oscillator energies. The consistent readings obtained before and after focusing the light inside the fiber indicated that no serious bulk damage occurred to the fiber from placing the beam waist inside the fiber at the pulse energies used. In addition, it was noted that the fiber was axially aligned with the beam when in position A but slightly off-axis in position B. The results are shown in Table 7.1 and indicate that input coupling is optimal when the beam waist is placed inside the fiber (provided the pulse energies are low enough to avoid optical damage to the fiber).

<table>
<thead>
<tr>
<th>Oscillator (V)</th>
<th>Position</th>
<th>Output (μJ)</th>
<th>Std. error (μJ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>640</td>
<td>A</td>
<td>50.7</td>
<td>0.4</td>
</tr>
<tr>
<td>640</td>
<td>B</td>
<td>53.8</td>
<td>0.4</td>
</tr>
<tr>
<td>640</td>
<td>A</td>
<td>39.8</td>
<td>0.3</td>
</tr>
<tr>
<td>640</td>
<td>B</td>
<td>53.0</td>
<td>0.4</td>
</tr>
<tr>
<td>640</td>
<td>A</td>
<td>45.9</td>
<td>0.3</td>
</tr>
<tr>
<td>640</td>
<td>B</td>
<td>52.0</td>
<td>0.4</td>
</tr>
<tr>
<td>640</td>
<td>B</td>
<td>52.2</td>
<td>0.3</td>
</tr>
<tr>
<td>640</td>
<td>A</td>
<td>46.3</td>
<td>0.2</td>
</tr>
<tr>
<td>640</td>
<td>B</td>
<td>54.2</td>
<td>0.2</td>
</tr>
<tr>
<td>640</td>
<td>A</td>
<td>45.6</td>
<td>0.3</td>
</tr>
<tr>
<td>640</td>
<td>B</td>
<td>53.4</td>
<td>0.5</td>
</tr>
<tr>
<td>640</td>
<td>A</td>
<td>44.5</td>
<td>0.3</td>
</tr>
</tbody>
</table>

7.2.2. Optical damage

De Hart (1992) identifies six potential damage sites in an optical fiber: (i) bulk damage to the fiber if the beam waist is placed inside the core; (ii) surface damage to the fiber input face due to dust, embedded polishing grit, or fiber defects; (iii) damage to the
fiber cladding where some modes may be coupled out of the fiber at the first bounce; (iv) bulk damage to the fiber core at the first internal focal point, especially with pulsed light; (v) damage similar to (iii) caused by stress induced by the fiber support system (e.g. the chuck) near the input end; and (vi) damage to the output endface due to dust, grit, or defects. In addition to these, damage also occurs to the input surface when the focal point is outside the fiber but too close to the input face, especially during the formation of an air plasma (Allison et al., 1987), and due to self-focusing inside the fiber (Allison et al., 1985). Self-focusing could lead to higher localized fluences in a fiber and subsequent bulk damage. Self-focusing is proportional to the square of the wavelength; the critical power at 532 nm is about 1.5 MW (Allison et al., 1985) and at 225 nm about 250 kW. The energy obtained from the doubling unit which was coupled into the fiber-optic probes (see Figure 3.1, system C) rarely exceeded 0.1 mJ/pulse or 20 kW even when accounting for a pulse shortened to 5 ns by the dye laser and doubling crystal. It was therefore unlikely that self-focusing would cause fiber damage.

The modes of failure listed here are generally applicable to pulsed light sources but also, in cases of high average powers (> 50 W), to continuous wave sources.

7.2.2.1. Continuous wave light.

Because high average powers were not used (200 mW maximum), coupling the beam from the Ar\textsuperscript{+} laser into optical fibers never produced any damage.

7.2.2.2. Pulsed light.

Coupling pulsed light into optical fibers was more difficult and generated several modes of failure. To prevent damage due to focusing inside the fiber, the beam was focused in front of the fiber. At high fluences an air plasma was generated which absorbed part of the laser beam as mentioned above. Under these conditions, the fiber was observed to fracture approximately 1 cm from the input face which was reminiscent
of a mode of failure described by Allison et al. (1987). Pitting of the input end was also observed. Therefore, the external focus should be some distance away from the fiber endface yet have expanded to no more than about 80% of the core diameter (De Hart, 1992). The use of a longer focal length lens (150 mm) improved the coupling efficiency, prevented the formation of an air plasma, and allowed positioning of the fiber further from the focal point, thus avoiding damage to the fiber.

However, focusing the beam in front of the fiber does not produce as efficient coupling as focusing inside the fiber as described above. Therefore, focusing inside the fiber to improve the coupling efficiency while using low pulse energies and a long focal length lens appears desirable.

The use of a fiber-optic bundle to avoid high fluences was also investigated. The input face area of a single excitation fiber could be multiplied by using several excitation fibers of the same diameter at the cost of increasing the probe diameter by the diameter of a single excitation fiber. For instance, using 6 excitation fibers of 300 \( \mu \text{m} \) diameter surrounding a single collection fiber of the same diameter increases the input surface area 6 times but the total diameter only 1.5 times. The optical damage threshold of a 4x300 \( \mu \text{m} \) input bundle of 6 cm length was tested under increasing pulse energies and the results are shown in Table 7.2. No visible surface or transmission damage was noted up to the maximum pulse energies tested (12 mJ/pulse). In contrast, damage occurred to a 600 \( \mu \text{m} \) diameter fiber at pulse energies of about 5 mJ/pulse. Disadvantages of using bundles are that it is impossible to avoid coupling light into the cladding thus reducing throughput and risking absorption-induced damage to the probe, that the open spaces between the fibers will cause a further reduction in throughput and may make the bundle more susceptible to damage, especially with small diameter fibers.

It should also be noted that the power distribution of a Gaussian beam would cause fibers situated centrally in the bundle to receive a disproportionate amount of the pulse energy, that is, the input energy is not evenly distributed among the fibers in a
bundle. The beam diameter of a Gaussian beam is defined by $1/e^2$ or 2 standard deviations from the mean. The beam intensity has dropped to about 60% of the maximum value at 1 standard deviation from the mean and about 70% of the throughput in the area defined by 1 standard deviation. This suggests that coupling the center 25% of a Gaussian beam cross-sectional area into a bundle would yield a more even distribution of light amongst the fibers with about 30% light loss.

Table 7.2.
The effects of increasing pulse energies on the transmission through a 4x300 μm input bundle of 6 cm length. Excitation at 266 nm and 10 Hz.

<table>
<thead>
<tr>
<th>Oscillator (V)</th>
<th>Bundle output (mL)</th>
<th>Std. error(mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>590</td>
<td>0.131</td>
<td>0.005</td>
</tr>
<tr>
<td>600</td>
<td>0.207</td>
<td>0.005</td>
</tr>
<tr>
<td>610</td>
<td>0.329</td>
<td>0.009</td>
</tr>
<tr>
<td>620</td>
<td>0.445</td>
<td>0.019</td>
</tr>
<tr>
<td>600</td>
<td>0.213</td>
<td>0.007</td>
</tr>
<tr>
<td>630</td>
<td>0.651</td>
<td>-----</td>
</tr>
<tr>
<td>640</td>
<td>0.816</td>
<td>0.016</td>
</tr>
<tr>
<td>600</td>
<td>0.214</td>
<td>0.007</td>
</tr>
<tr>
<td>640</td>
<td>0.835</td>
<td>0.015</td>
</tr>
<tr>
<td>650</td>
<td>1.083</td>
<td>0.024</td>
</tr>
<tr>
<td>600</td>
<td>0.188</td>
<td>0.004</td>
</tr>
<tr>
<td>660</td>
<td>1.337</td>
<td>0.038</td>
</tr>
<tr>
<td>670</td>
<td>1.420</td>
<td>0.044</td>
</tr>
<tr>
<td>600</td>
<td>0.206</td>
<td>-----</td>
</tr>
</tbody>
</table>

Another way to increase the input area to facilitate light coupling and minimize the risk of damage associated with high fluences, is to use a tapered fiber. These fibers gradually taper from one diameter to another over a certain fiber length. This allows one to use a large input face area without sacrificing the probe diameter. A probe employing a tapered excitation fiber is described and evaluated in Chapter 8.
7.2.3. Light transmission

The quoted attenuation for UV transmission for fibers from different manufacturers vary considerably. For example, at 200 nm fiber from Polymicro Technologies has an attenuation of about 1 dB/m while fiber from Fiberguide Industries has an attenuation of about 2 dB/m. These values represent an attribute which could be termed 'static' attenuation. Static attenuation results from Rayleigh scattering in the fiber and absorption by dopants and impurities in the fiber. For use in the deep UV, an equally or more important 'dynamic' attenuation exists due to the formation of reversible and irreversible color centers in the fiber (e.g. Karlitschek et al., 1996; Klein et al., 1995; Toriya et al., 1995).

This type of attenuation, however, was not mentioned by the fiber manufacturers and its discovery was somewhat unexpected. There is some evidence that manufacturers are aware of this issue and that they have started to address it, driven by the need to find optical fibers suitable for the delivery of excimer wavelengths (e.g. Klein et al., 1995). Recently, Polymicro Technologies have developed an experimental fiber aimed at reducing dynamic attenuation and made some of it available for testing (see Figure 7.5). Another fiber manufacturer, 3M, is also actively engaged in developing optical fiber suitable for pulsed UV transmission.

7.2.3.1. Continuous wave excitation.

Continuous wave light transmission at 364 nm exhibited more static attenuation than at visible wavelengths but dynamic attenuation was not noticed.

7.2.3.2. Pulsed excitation.

Pulsed light transmission was severely hampered in the UV by both static and dynamic attenuation. The static attenuation was characterized by the manufacturer's specifications, but the dynamic attenuation was unknown until a transmission time-
course study revealed unexpected attenuation. Figure 7.2. shows the attenuation with
time for a 15 cm section of 300 μm diameter Fiberguide Industries fiber excited with a
pulsed Nd:YAG at 266 nm, 10 Hz repetition rate, and averaging 100 pulses per point.
The recovery exhibited by the fiber after a 5 minute rest period is also shown. It is
evident that complete recovery occurred for this fiber and under these conditions.

![Figure 7.2. The attenuation for a 15 cm section of 300 μm diameter fused silica fiber excited at 266 nm, 10 Hz, and averaging 100 pulses per point, as a function of time. The recovery exhibited by the fiber after a 5 minute rest period is also shown (last measurement).](image)

The attenuation was found to be a function of the pulse energies as can be seen
from Table 7.3. Higher pulse energies excited a greater attenuation resulting in a reduced
transmission efficiency. This was further confirmed by investigating the degree of
attenuation as a function of the distance between the input face and the beam waist. The
results (given in Figure 7.3) showed that the attenuation was less severe when the fiber
was a greater distance from the beam waist (thus subjected to decreased energy densities).

Table 7.3. 
The effects of increasing pulse energies on the transmission efficiency (average of 10 pulses) through a 15 cm section of 300 \( \mu m \) diameter Fiberguide Industries Superguide G fiber. Excitation at 266 nm and 10 Hz.

<table>
<thead>
<tr>
<th>Oscillator (V)</th>
<th>Input (mJ)</th>
<th>Output (mJ)</th>
<th>Efficiency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>570</td>
<td>0.248</td>
<td>0.126</td>
<td>50.8</td>
</tr>
<tr>
<td>580</td>
<td>0.571</td>
<td>0.255</td>
<td>44.7</td>
</tr>
<tr>
<td>600</td>
<td>1.821</td>
<td>0.495</td>
<td>27.2</td>
</tr>
</tbody>
</table>

![Graph showing pulse energy drop vs distance from beam waist](image)

Figure 7.3. The attenuation for a 1 m section of 600 \( \mu m \) diameter fused silica fiber excited at 266 nm, 10 Hz, and averaging 10 pulses per point as a function of the distance between the input face and the beam waist.

The cause of this dynamic behavior was not known and hypothesized to be due to temperature effects. To eliminate the possibility that misalignments due to mechanical distortions caused by temperature changes contributed to this phenomenon, the degree of dynamic attenuation was tested with either 2 mm or 6 mm of the bare fiber protruding beyond the supporting point of the fiber-holding chuck. As Figure 7.4 shows, the
dynamic attenuation was not caused by mechanical instability and was due to other causes. A literature search suggested that color center formation in the fibers (e.g. Klein et al., 1995; Kubo et al., 1994; Toriya et al., 1995) may have been the cause of dynamic attenuation.

Figure 7.4. The attenuation for a 15 cm section of 300 μm diameter fused silica fiber excited at 266 nm, 10 Hz, and averaging 100 pulses per point as a function of time. The fiber tip extended freely either 2 mm (top) or 6 mm (bottom) from its support. In neither case was the dynamic attenuation affected.

Subsequent testing at 225-235 nm and ~0.20 mJ/pulse also produced strong dynamic attenuation in low UV attenuation fibers from Polymicro Technologies, 3M, Oriel Corporation, and Fiberguide Industries, some of which are superimposed in Figure 7.2 for comparison, and appeared to vary based on the manufacturer. A direct comparison also revealed a dependence of dynamic attenuation on wavelength. Whether dynamic attenuation varied with respect to placing the beam inside or outside the fiber was not investigated.
Figure 7.5. The attenuation (mean ± S.E.M.; n = 10) for 15 cm sections of 300 μm fused silica fibers from Polymicro Technologies (dotted line), 3M (dashed line), Fiberguide Industries (dash/dot), and the new experimental fiber from Polymicro Technologies (200 μm, solid line) at 230 nm and 10 Hz as a function of time. The last point in each of the two top traces were obtained after a 5-minute recovery period.

It can be seen from Figure 7.5 that the experimental fiber, despite its smaller diameter, outperformed the other fibers both in terms of decay and recovery characteristics. The new fiber was further characterized with regard to recovery time, throughput and efficiency, as well as attenuation. A 20 cm segment of 200 μm diameter experimental fiber was tested with approximately 75 μJ input energy at 230 nm and 10 Hz to determine its recovery behavior. The results are shown in Figure 7.6 and indicated that a resting period of 1 minute produced nearly full transmission recovery.
Figure 7.6. The transmission efficiency (mean ± S.E.M.; n = 10) of an experimental fiber (20 cm section, 200 μm) as a function of recovery time at 230 nm and 10 Hz.

The efficiency of the same section of fiber was subsequently tested as a function of input energy. The input energy was incremented from about 80 μJ/pulse to about 250 mJ/pulse in 10 steps without any intervening recovery periods. The results, shown in Figure 7.7, indicated that the throughput efficiency of the fiber declined with increasing pulse energies (consistent with data given in Table 7.3 for a different fiber). Output energies gradually increased and eventually exceeded 50 μJ/pulse.

Although output energies in excess of 50 μJ/pulse could be achieved and sustained via a gradual increases in input energy, it could not be done with a single very large increment. Less large increments produced initial sharp declines in fiber throughput as shown in Figure 7.2 and, for comparison, again in Figure 7.8. Small increments did not seem to produce these declines at all.
Figure 7.7. The transmission efficiency (dotted line) and output pulse energies (solid line) of an experimental fiber (20 cm section, 200 μm) as a function of input energy at 230 nm and 10 Hz (mean ± S.E.M.; n = 10).

For example, when coupling about 203.43 ± 1.57 μJ/pulse (representing a 200 μJ/pulse increase) into a fresh 20 cm section of experimental fiber, the fiber output dropped to below the detection threshold after the first reading of 40.40 ± 0.35 mJ/pulse. When the experiment was terminated 5 minutes later, the output energy was still below the detection threshold. In contrast, an input of 217.50 ± 1.81 mJ/pulse (representing a 28 μJ/pulse increase), produced an output of 48.86 ± 0.16 mJ/pulse for at least 2 minutes when the previous input energy levels were attained via gradual increases. The moderately large pulse increase consisted of an input of 77.17 ± 1.15 mJ/pulse and produced a maximum decline of about 13%. These results indicated that a gradual step-wise increase in input energy was required to achieve high energy throughputs.
Figure 7.8. The output energy as a function of time for fibers tested under 3 conditions: increasing the input energy with small increments (e.g. oscillator voltage from 640-646; solid line); with a moderate increment (590-615 V; dashed line), and with a large increment (590-639 V; dotted line) to the final operating level. Data were scaled to the same initial values.

7.2.4. Beam profiles.

The emission profiles from optical fibers of varying diameter were measured as part of the optical fiber characterization. Light from a HeNe laser (5 mW) was coupled into an optical fiber of given diameter and the emitted light was projected onto a Reticon 4096 element diode array detector (EG&G, Princeton, NJ). The fiber, 133 mm from the array, was mounted onto a translation stage that could be adjusted to move the beam incrementally across a mask containing a 1 mm slit covering the array. The readout from the array was moved to a microcomputer using Computerscope software via a Reticon 1020 interface and a ISC-16 data acquisition board. A C program (by L.S. Greek) was used to convert the binary data to ASCII format. The results indicated that the emission
beam profiles from smaller diameter fibers approached Gaussian distributions, but larger diameter fibers showed increasingly bimodal beam profiles. The central section beam profile of a 300 μm fiber is shown in Figure 7.9. Figure 7.9 reveals that although the beam profile of a 300 μm fiber can still be approximated by a Gaussian distribution, some bimodality is already evident. Although it was known that the different lasers used exhibited radically different beam profiles, time did not permit investigating the dependencies of the emission profile on the profile of the input beam and on other coupling parameters. Such dependencies could reasonably be expected to exist.

![Figure 7.9](image)

Figure 7.9. The central section emission beam profile of a 300 μm fused silica fiber with a Gaussian fit (dashed line) superimposed for comparison.

7.3. DISCUSSION.

The central issues regarding the use of optical fibers for UV resonance Raman spectroscopy are those of physical damage to fibers, light coupling into the fibers, and light propagation by the fibers.
The findings reported here concerning fiber damage, light coupling, and dynamic and static attenuation, were in general agreement with those found in the literature (Allison et al., 1987; Klein et al., 1995; Kubo et al., 1994; Toriya et al., 1995; Taylor et al., 1988). However, due to variations in fiber attributes based on the manufacturer and due to the absence of critical information on the supplier's specification sheets and in the literature, individual fiber characteristics had to be determined. From the reports mentioned above, it was learned that the dynamic attenuation was due to color center formation in the fiber. Therefore, fibers with a weak tendency to form color centers should be used as excitation fibers when fabricating an optical fiber probe.

Physical damage was shown to occur at high pulse energies and be aggravated by short focal length lenses, consequently, these are best avoided. Lower pulse energies will allow the beam to be focused inside the fiber to improve light coupling into the fiber as shown in Table 7.1. The use of a long focal length lens would avoid damage at first bounce and first internal focus due to modal spreading. The best method for coupling light into an optical fiber may be to use a Galilean telescope to reduce the input beam diameter to somewhat less than that of the fiber core. With this arrangement, no focusing of the beam would occur anywhere in the optical path.

The transmission of light through a fiber was given by the Equation 6.2 where the linear attenuation has previously been termed 'static' and the UV-induced attenuation 'dynamic'. The dynamic attenuation was found to be a function of the pulse energies as can be seen from Table 7.3, number of transmitted pulses, light frequency, and laser repetition rate. Equation 6.2 could be restated as:

\[- \frac{dI}{dz} = (\alpha_0 + \Delta\alpha_d(p, \nu, R))I + \beta(\nu)I^2\]  

(7.1)

where the symbols are as defined before and \(p\) is the number of pulses transmitted, \(\nu\) the light frequency (Hz), and \(R\) the laser repetition frequency (Hz). The transmission coefficients are now being determined for the new experimental fibers from Polymicro Technologies.
Using lower pulse energies would therefore result in less dynamic attenuation and greater average throughput. Taken together, low pulse energies will avoid fiber damage, enable efficient light coupling into the fiber, and ensure more efficient light propagation by the fiber. Energy delivery could be maximized by using the new fibers and by gradually increasing the input energy to avoid catastrophic failure due to the sudden transmission of high energy pulses.

Additional methods to increase the average power delivered to the sample should be investigated or other means found to deliver high energy pulses to the sample. Increasing the pulse repetition rate or extending high energy pulses could be used to increase the average power. Since the laser repetition rate of the laser used here cannot be increased beyond 20 Hz, the need to find a suitable pulse extender becomes all the more important given the many problems it would ameliorate. Two recently developed resonance Raman instruments boost average power along these very lines: one employs a continuous wave (infinitely extended pulse) intracavity doubled argon laser (FRED) as a UV source (Russell et al., 1995) and the other a tripled/quadrupled high repetition rate Ti:sapphire laser (Manoharan et al., 1994). Some researchers have also had success in decreasing dynamic attenuation thus boosting fiber throughput by annealing the fiber before use (Toriya et al., 1995).

Concerning alternative light guides, hollow, small-bore light pipes with aluminized reflecting surfaces could potentially be used. Short sections (3-5 mm) of fused silica or sapphire optical fiber could be made into windows for the hollow light guides for use in solutions. These windows could be shaped (e.g. lens or angle), anti-reflection coated, and provided with reflecting surfaces if required. Silver coated small-bore glass waveguides have recently been developed for light delivery from CO$_2$ and Er:YAG lasers (Matsuura et al., 1995).
The work reported in this chapter provided the information regarding fiber damage, light transmission characteristics, and input coupling necessary to proceed with the design and fabrication of actual optical fiber probes.

7.4. REFERENCES.


8.1. INTRODUCTION

The aim of the interdisciplinary work undertaken and reported in this thesis was to conceptualize, design, and lay the foundations for the development of a brain sensor capable of meeting the general requirements of sensitivity, speed, and versatility while operating non-invasively in an aqueous in situ environment. Therefore, the basic goals consisting of the design and development of a suitable fiber-optic probe for in vivo operation and the determination of the conditions for its optimum use (see Chapter 1) were of pivotal importance to this project.

The present chapter discusses the investigations concerning the optimal geometry of a fiber-optic probe for resonance Raman work in vivo. These investigations consisted of computer simulations of probe performance, researching different fabrication procedures, and evaluations of the probe in vitro. For the theoretical background, the reader may wish to refer to Chapter 6.
8.2. FIBER-OPTIC PROBE DESIGN.

Criteria important in the design of optical fiber probes can sometimes be investigated with the aid of simulations and has been done for single (e.g. Zhu and Yappert, 1992a) and double fiber sensors (e.g. Zhu and Yappert, 1992b; Schwab and McCreery, 1984). Such simulations can then be experimentally verified and used to optimize the probe design. Some design considerations and two simulations are discussed in the next sections: an intensity profile and collection efficiency simulation and a more fine-grained probe simulation (the latter generated by L.S. Greek and evaluated by us, see Greek et al., 1996).

8.2.1. Design considerations and simulations.

When using optical fibers for Raman spectroscopy, Raman and luminescence signals generated in the optical fiber core or cladding should be considered. Prominent Stokes shifted silica Raman features which can interfere with analyte signals occur at 1535, 817, and 516 cm\(^{-1}\) and are pronounced in long fibers (Kercel et al., 1990), while the anti-Stokes band near 500 cm\(^{-1}\) has also been observed (Gambling and Poole, 1988). In addition, strong Rayleigh scattering by the fiber may completely obscure the much weaker Raman signals generated by analytes (Heiman et al., 1989). Although the use of a single fiber for both excitation and collection is efficient because of a complete overlap of excitation and collection volumes the interference caused by Raman and Rayleigh scattering from the fiber itself requires a dual or multiple fiber arrangement (Myrick and Angel, 1990).

In a dual fiber geometry, one fiber is used for delivery of the excitation radiation, while another fiber is used to collect the scattered radiation from the sample. Myrick et al. (1990) found the collection efficiency in a very dilute rhodamine 6G solution (~30 pM) to be optimal for small angles between collection and excitation fibers and poor for right
angle collection. The same advantage also benefits single excitation, multiple collection fiber arrangements. Such dual (Heiman et al., 1989) and multiple fiber (Schwab and McCreery, 1984; Wang et al., 1992) probes have found applications in remote Raman spectroscopy. A forward scattering geometry is also efficient, but requires filters to remove the laser line (Myrick et al., 1990) and Raman lines produced by the fiber (Ma and Li, 1994) which could be impractical where very small probes are needed. For these reasons, it was decided to use separate excitation and collection fibers.

As part of the characterization of the intended measuring environment, the normal Raman (Chapter 4) and UV absorption (Chapter 5) spectra of cerebrospinal fluid samples were obtained. Although normally a clear fluid, the UV absorption spectrum of CSF (Figure 5.6) indicated that it was highly absorbing in the region of interest. It was therefore necessary to allow for the effects of a highly absorbing environment in the probe design.

8.2.1.1. Model parameters.

A knowledge of the excitation intensity profile was required in order to aid in the determination of the optimum probe geometry. In order to obtain this profile, a coarse-grained simulation was generated. The detailed simulation available (Greek et al., 1996) did not accommodate differences in probe tip geometry and did not allow for a visualization of the excitation light intensity distribution in the analysis volume. Parameters that could be varied in the simpler model included excitation and collection fiber diameters, separation distance between these fibers, numerical aperture, sample absorptivity, intensity gradations of the excitation light, and fiber tip geometry. The source code for this program (in Basic 4.0, Microsoft Corporation, Redmond, WA) is given in Appendix A. A gaussian beam profile with centred mean and 3 standard deviations across the fiber radius was assumed for the exiting light, but could be changed if required. This assumption was justified by the results obtained from measuring the
emission profiles from optical fibers as discussed in Chapter 7. Figure 8.1 shows the simulated intensity distribution of light emerging from the excitation fiber of a front-casting probe.

![Image of simulated intensity distribution](image)

**Figure 8.1. The simulated intensity distribution of light emerging from the excitation fiber of a front-casting probe.**

The profile simulation program was used to calculate the efficiencies of fiber optic probes with flush or 45° angled excitation tips in samples with different absorbances. This calculation was not integrated over the entire analysis volume, but consisted only of the central section through the analysis volume. In addition, the maximum sample depth on which the calculations were based, was only 533 μm. It was assumed that the collection fiber collected photons as long as they were scattered from the overlapping region between the excitation and collection fiber cones. This was an oversimplification which tended to overestimate the collection efficiency of probes where the fibers were aligned flush and overestimated to a lesser degree the efficiency of probes.
with angled excitation fibers. Furthermore, the likelihood of a photon being collected by the collection fiber depended on the intensity of the light in the volume element where it was scattered from and on the distance to the collection fiber endface. Excitation and scattered photons were in a 1:1 ratio. However, scattering, though in direct proportion to the excitation light intensity, has a very much smaller probability. For these reasons, the results obtained with the coarse-grained simulation were mostly of qualitative value.

8.2.1.2. Excitation and collection fiber sizes.

The effects of varying sizes of excitation and collection fibers were investigated with the model. The results are given in Table 8.1.

<table>
<thead>
<tr>
<th>Excitation ( \phi ) (( \mu \text{m} ))</th>
<th>Collection ( \phi ) (( \mu \text{m} ))</th>
<th>Separation (( \mu \text{m} ))</th>
<th>Efficiency(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>100</td>
<td>20</td>
<td>17.2</td>
</tr>
<tr>
<td>100</td>
<td>200</td>
<td>30</td>
<td>13.4</td>
</tr>
<tr>
<td>100</td>
<td>300</td>
<td>40</td>
<td>10.5</td>
</tr>
<tr>
<td>100</td>
<td>100</td>
<td>20</td>
<td>17.2</td>
</tr>
<tr>
<td>200</td>
<td>100</td>
<td>30</td>
<td>6.4</td>
</tr>
<tr>
<td>300</td>
<td>100</td>
<td>40</td>
<td>3.4</td>
</tr>
</tbody>
</table>

The results show that for a given diameter collection fiber, smaller excitation fibers are more effective. This is consistent with the fact that with smaller excitation fibers all the light energy gets "compressed" into a volume closer to the collection fiber which improves collection efficiency. Using larger diameter collection fibers given a certain excitation fiber diameter leads to small decreases in collection efficiency. This makes little conceptual sense because with larger diameter collection fibers, scattered light ought to be collected more easily. The latter result therefore reflects only the effects of increased separation distances between the fibers.
8.2.1.3. Fiber tip modifications.

Due to internal filtering in samples with pronounced molar absorptivity, it is important for the excitation volume to be in close confinement to the collection fiber endface. This can be achieved by side-casting the excitation light across the collection surface. Figure 8.2 shows the intensity distribution of light emerging from the excitation fiber of a side-casting probe.

Figure 8.2. *The simulated intensity distribution of light emerging from the excitation fiber of a side-casting probe.*

Although the total analysis volume of the side-casting probe is smaller than that of the front-casting probe, a comparison of Figures 8.1 and 8.2 reveals that those areas of the analysis volume of the side-casting probe excited by the most intense light fall within the acceptance cone of the collection fiber in contrast to the other arrangement where it falls mostly outside of the collection cone. A comparison of Figures 8.1 and 8.2 further reveals
that the average distance from a volume element in the analysis volume of the side-casting probe to the collection face is shorter than the corresponding distance in the front-casting probe. Both considerations favor the side-casting probe in absorbing solutions and possibly in arrangements with restricted sample volumes. Table 8.2 shows the relative collection efficiencies of front and side-casting probes in samples with different absorbances.

Table 8.2.

<table>
<thead>
<tr>
<th>Excitation φ (μm)</th>
<th>Collection φ (μm)</th>
<th>Absorbance (%)</th>
<th>Efficiency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>100</td>
<td>0</td>
<td>1.55</td>
</tr>
<tr>
<td>100</td>
<td>100</td>
<td>50</td>
<td>0.84</td>
</tr>
<tr>
<td>100</td>
<td>100</td>
<td>90</td>
<td>0.00</td>
</tr>
<tr>
<td>Front-casting probe</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>100</td>
<td>0</td>
<td>38.86</td>
</tr>
<tr>
<td>100</td>
<td>100</td>
<td>50</td>
<td>41.34</td>
</tr>
<tr>
<td>100</td>
<td>100</td>
<td>90</td>
<td>33.77</td>
</tr>
<tr>
<td>Side-casting probe</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

An inspection of Table 8.2 confirms that the side-casting probe is much more efficient than the front-casting probe in samples with limited probe depth, especially absorbing samples.

The model was also used to investigate the positioning of the collection fiber endface relative to the excitation fiber tip. These results are shown in Table 8.3 and indicated that the collection efficiency of the probe increased as the collection fiber endface was moved away from the excitation fiber tip. Moving the endface of the collection fiber away caused more overlap between excitation and collection volumes. An optimum could not be attained with the current model, but is expected to occur at a distance where the entire output ‘endface’ (the fiber surface area over which the excitation fiber emits light) of the excitation fiber falls within the collection cone of the collection fiber.
Table 8.3.
The effects of probe geometry on the relative collection efficiencies of dual side-casting fiber probes for maximum probe depths of 500 μm. The separation distances between fibers were 20 μm and that between collection fiber endface and excitation fiber apex was varied as indicated.

<table>
<thead>
<tr>
<th>Excitation φ (μm)</th>
<th>Collection φ (μm)</th>
<th>Distance (μm)</th>
<th>Efficiency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>100</td>
<td>100</td>
<td>38.86</td>
</tr>
<tr>
<td>100</td>
<td>100</td>
<td>120</td>
<td>41.17</td>
</tr>
<tr>
<td>100</td>
<td>100</td>
<td>140</td>
<td>43.02</td>
</tr>
<tr>
<td>100</td>
<td>100</td>
<td>160</td>
<td>46.08</td>
</tr>
</tbody>
</table>

8.2.1.4. Bundles.

For a given diameter excitation fiber, the collection efficiency of a probe can be increased by increasing the collection area of the probe. This can be accomplished by simply using a larger diameter collection fiber. An alternative method is to distribute the same area evenly around the excitation fiber by using several smaller diameter collection fibers or a hollow capillary instead of one larger fiber. The alternative method has the benefit of distributing the collection area closer to the high intensity excitation, thus improving collection efficiency. Thus, using four 200 μm collection fibers should be four times better than one 200 μm collection fiber as well as better than using a single 400 μm collection fiber (use values from Table 8.1 for comparison).

However, by using an angled excitation fiber, the collection area does not need to be distributed around the excitation fiber since all the excitation light can now be directed towards the collection fiber. Furthermore, if several excitation fibers are used and oriented to cast the excitation energy across a single central collection fiber, two advantages accrue: (i) more energy can be transmitted to the sample without risking damage to the probe and (ii) a high collection efficiency can be assured because the single collection fiber now substitutes for several collection fibers.
8.3. FIBER-OPTIC PROBE FABRICATION.

This section provides a detailed description of the probe fabrication processes used. Probes using bundles of fibers with modified tips were the most delicate and time-consuming to assemble. This work required the use of a microscope and some photographs were taken during the fabrication process. Probes are designated as: 
((number of fibers or other descriptor) excitation diameter in μm)x((number of fibers or other descriptor) collection fiber diameter μm), e.g. ((10)100)x(400) is a probe with 10 excitation fibers of 100 μm each and one collection fiber of 400 μm.

8.3.1. End-face preparation.

The quality of the fiber endfaces are crucial determinants of probe efficiency, especially if used for coupling high intensity excitation light into the fiber (Allison et al., 1987; Beck et al., 1993; De Hart, 1992). Good quality endfaces can be obtained by polishing or cleaving the fiber after stripping about 5 cm of jacketing material. Polishing was performed by hand using a Newpor optical fiber chuck (FPH-DJ) to hold the fiber, gently applying pressure to the fiber tip while executing figures-of-eight on successively finer emery grit (600, 1200, and 4000). Cleaving was performed with a Newport fiber-optic cleaver (F-BK2) for 100 μm fibers and with a diamond scribe for larger fibers. The surfaces were then inspected under a microscope with 40x magnification for imperfections. If a bundle was to be assembled, the requisite number of fibers, plus a few additional ones, were cut to the desired length, the endfaces prepared, and more jacketing stripped if needed to yield about 5 cm of bare fiber. Figure 8.3 shows two 300 μm fibers with polished endfaces before being assembled into a probe.
If a UV reflective angled endface were to be made, the chuck was held at the required angle while polishing. After polishing, about 1 cm of the angled tip was covered with positive photoresist (S1400-27) and baked in an oven at 95°C for 20 minutes to harden the photoresist. The fiber was removed from the oven and the angled facet lightly sanded with lapping paper to remove the photoresist on the angled surface only. The surface was cleaned with deionized water to remove dust and grit. The angled tip was then inserted into a well in a vacuum chamber (CHA Industries Model SE-600-RAP, Menlo Park, CA) and positioned close to a vacuum deposition thickness gauge. The vacuum chamber was subsequently closed and aluminum deposited on the fiber by vacuum evaporation (~2x10^{-6} Torr) to a thickness of 3/4 $\lambda$ (for 250 nm the thickness was 187.5 nm) for optimum reflectance. After aluminum deposition, the fiber was removed and the tip inserted into warm acetone (~ 30°C) to remove the remaining
photoresist from the protected areas of the fiber tip. An alternative method consisted of using aluminum pigmented acrylic paint and a very fine (#1) paintbrush to coat the endface. The latter method however, produced more diffuse than specular reflection.

8.3.2. Probe assembly.

8.3.2.1. Dual fiber probes.

A collar made from a 0.5 mm section of 400 μm diameter silastic tubing was fitted around the end of one fiber by stretching it with a piece of wire (with ~300 μm diameter obtained from a stranded electrical wire) or a pair of small forceps. By repeating the process, a second fiber could be accommodated inside the same collar as shown in Figure 8.4.

Figure 8.4. A silastic collar being stretched with fine wire to accommodate two optical fibers.
When the wire was removed, the elastic tension in the silastic collar held the two fibers together about 1 mm from the probe tip. By gently rotating and pulling the longer fiber while holding the other fixed (e.g. with a piece of adhesive tape), the two fiber endfaces were aligned with relative accuracy as shown in Figure 8.5. The fibers were then permanently glued together.

![Image of aligned optical fibers](image.png)

Figure 8.5. *The aligned endfaces of two optical fibers of 300 μm diameter being held in place by the silastic collar.*

8.3.2.2. Multi-fiber probes.

Although multi-fiber probes with several excitation and several collection fibers could be made, it was easier to use either a single excitation fiber with multiple collection fibers or *vice versa*. The probe-end geometry of such probes always consisted of the single fiber (whether excitation or collection) in the center with the other fibers arranged around its perimeter. In order to obtain this geometry, the following procedure was followed. All the fibers were gently tied together, about 1 cm from their endfaces, with a
thread. A silastic collar was then stretched in three directions with three separate pieces of wire making a triangular opening through which the bundle was fitted. Upon release the collar would hold all these fibers together. For example, the 8 angled excitation fibers (200 μm) and one flat end collection fiber (300 μm) shown in Figure 8.6 were fed through a silastic collar and held together during the assembly process.

![Figure 8.6.](image)

Figure 8.6. *End view of an optical fiber probe consisting of a central collection fiber (300 μm) surrounded by 8 angled excitation fibers (200 μm each).*

At this stage, the fibers were not in their correct positions and considerable care and effort were required to gently prod the central fiber with a sharp blade to the center of the bundle. An alternate method used consisted of inserting the central fiber through the collar first and then feeding the surrounding fibers one by one through the collar into position around it. The fibers were now aligned as described above. The individual surrounding fibers were labeled to facilitate the process. Where modified tips were used, fibers had to be aligned with respect to length as well as oriented with respect to angled
surface. Light from a small HeNe laser was sequentially coupled into the one end of each fiber and the orientation adjusted until the deflected beam crossed the central fiber. When the alignment was satisfactory, the fibers were glued together. Figure 8.7. shows a side view of the probe with 8 angled excitation fibers and one central collection fiber depicted in Figure 8.6.

After the glue had completely cured, the probe tip was inserted into a glass pipette for protection. For measuring \textit{in vitro} spectra, the probe was pushed forward until the tip protruded a little (~1 mm) from the pipette and for measuring \textit{in vivo} spectra, the probe was removed entirely from the pipette.
If the multiple fibers were to be used as collection fibers, no further assembly was required. For the purposes of coupling them to the spectrograph, a special chuck was machined in-house allowing the fibers to be lined up in front of the entrance slit. If multiple collection fibers were to be used as excitation fibers, their input ends had to be formed into a bundle. For this purpose, the fibers were packed as tightly as possible into a short section (~ 5 cm) of glass capillary tubing with internal diameter just large enough to accommodate them and with the bare fibers extending about 2 mm beyond the capillary. Often a suitable diameter tubing was not available and a larger diameter had to be used requiring one or more spacers. Figure 8.8 shows a spacer in the input bundle of the probe with 8 angled excitation fibers and one collection fiber reflected from a mirror held at 45°. Clearly an additional spacer was required but not provided for fear of risking damage to the probe.

![Figure 8.8](image.png)

Figure 8.8. *End view of the input bundle of the optical fiber probe consisting of a central collection fiber surrounded by 8 angled excitation fibers (200 µm each) shown in Figure 8.6. The single lighter fiber functions as a spacer to improve alignment of the fibers inside the glass collar.*
8.3.3. Glue.

Initially, an epoxy resin was used to glue the fibers together in the fabrication process. This proved unsatisfactory due to the large diameter probe tips that resulted from the very viscous resin. Furthermore, the biocompatibility of the substance was not known. It was therefore decided to use medical grade silicone rubber (Medical Adhesive Silicone Type A, Dow Corning, Midland, Michigan) which was certified biocompatible for implantations of less than 30 days. Silicone Type A could be mixed with cyclohexane to form an arbitrarily thin suspension. This gave a considerable degree of control over the thickness of the applied glue layer and hence over the final probe diameter. It was necessary, however, to increase the setting time of the glue suspension beyond the approximately 5 minutes available to allow for the correction of small alignment and orientation disturbances occurring during the gluing process. This was accomplished by adding tetradecane to the suspension to slow the vulcanization rate of the silicone rubber. Figure 8.9 shows a dual front-casting probe glued beyond the silastic collar with a suspension by volume of 2:40:1 silicone rubber, cyclohexane, and tetradecane, respectively.

Figure 8.9. A dual front-casting probe glued beyond the silastic collar with a suspension of silicone rubber, cyclohexane, and tetradecane, to produce a thin film of glue.
8.4. FIBER-OPTIC PROBE EVALUATION.

The probes were evaluated in different solutions to determine their efficiencies. Initially, benzene was used as a standard because it was a strong Raman scatterer. The intensities of the benzene spectra obtained with the different probes were used for comparison. However, benzene was soon abandoned in favor of KNO$_3$ which did not dissolve the probe glues, was water soluble, and was nevertheless a relatively strong Raman scatterer in the spectral region of interest. Subsequently, working curves for a number of probes were established using methyl orange and KNO$_3$ dissolved in water.

8.4.1. Working curves.

A mixture of a fixed concentration of 0.153 M KNO$_3$ and concentrations of methyl orange varying from 0-400 µM was used to determine the working curves for several front-casting probes and one side-casting probe. For each methyl orange concentration, 8 spectra of 5 s integration each, covering the range 1500 cm$^{-1}$ to 1000 cm$^{-1}$, were collected with a specific probe in situ. The 473 nm line of the Ar$^+$ laser was used for excitation and the McPherson single monochromator for frequency decoding. The average height of the 1400 cm$^{-1}$ methyl orange peaks and the 1050 cm$^{-1}$ KNO$_3$ peaks from the 8 spectra were plotted against methyl orange concentration to establish the working curves. A program (written by L. S. Greek) was used to calculate the peak heights from the spectra. Details can be found in Greek et al. (1996).

8.4.1.2. Excitation and collection fiber sizes.

Figure 8.10 shows the working curves of 3 dual fiber probes with varying collection fiber diameters. The excitation fiber diameter was 300 µm and the collection fibers were 300, 600, and 1000 µm, respectively. The working curves were normalized with respect to output power. It is evident from Figure 8.10 that the different probes achieved their maximum efficiencies at different methyl orange concentrations. This
finding is useful in that it allows one to select the probe dimensions most suitable for a given analyte concentration (range). Overall, the 300x600 probe produced the strongest signals over the widest range. Thus, even though the 300x300 probe reached its maximum efficiency at a methyl orange concentration of 50 μM, it was still less efficient than the 300x600 probe. The optimum geometry represented by the 300x600 probe possibly reflected the fact that an increased collection area, leading to an increased collection efficiency, was countered by an increased separation distance between the probes, leading to a decreased collection efficiency. It was not possible, however, to rule out the contributions of other factors, such as endface quality, to these results. A more rigorous test would have required comparing the averages of several probes of a given geometry, which time and resources did not permit.

![Graph showing working curves of 3 front-casting probes with 300 μm excitation fiber diameters and 300 (dotted line), 600 (dashed line), and 1000 μm (solid line), collection fiber diameters, respectively. The working curves were normalized with respect to output power.](image)

**Figure 8.10.** The working curves of 3 front-casting probes with 300 μm excitation fiber diameters and 300 (dotted line), 600 (dashed line), and 1000 μm (solid line), collection fiber diameters, respectively. The working curves were normalized with respect to output power.
The working curves of two probes with the same collection fiber diameter (1000 μm) and varying excitation fiber diameters, 300 μm and 1000 μm, respectively, are shown in Figure 8.11. The optima of the two probes occurred at virtually the same methyl orange concentration. The increased separation distance between the fibers in the 1000x1000 probe should result in its reaching an optimum efficiency at a lower methyl orange concentration than the 300x1000 probe. However, the magnitude of the effect is hard to assess. On the other hand, the effect of a smaller excitation fiber is clear and in agreement with the qualitative predictions from the crude simulation: a smaller excitation fiber increases the probe's efficiency by exciting a smaller volume closer to the collection fiber, but more intensely, as discussed above.

**Figure 8.11.** The working curves of two probes with the same collection fiber diameter (1000 μm) and varying excitation fiber diameters, 300 μm (dotted line) and 1000 μm (solid line), respectively, normalized with respect to output power.
8.4.1.3. Fiber tip modifications.

The effects of fiber tip modifications were investigated with a 200x200 front-casting probe and a 200x200 side-casting probe with aluminum pigmented acrylic paint applied to the angled facet of the excitation fiber to provide a reflective surface. The working curves for these two probes are shown in Figure 8.12. An inspection of Figure 8.12 shows that the side-casting probe reached its maximum efficiency at a higher methyl orange concentration than the front-casting probe. The better overall efficiency of the side-casting probe was predicted from the crude simulation.

![Graph](image)

Figure 8.12. The working curves of a 200x200 front-casting probe (dotted line) and a 200x200 side-casting probe (solid line) with aluminum pigmented acrylic paint applied to the angled facet of the excitation fiber to provide a reflective surface. The working curves were normalized with respect to output power.

8.4.1.4. Bundles.

Two probes with fiber-optic bundles were made and tested. Previous results and simulations suggested that small excitation fibers and larger collection fibers were required
required for optimal efficiency. Therefore, the probes were assembled using the smallest
diameter excitation fibers generally used and intermediate diameter collection fibers. One
probe had an input bundle of 10 excitation fibers of 100 \( \mu m \) diameter each and a 400
\( \mu m \) collection fiber and is shown in Figure 8.13, while the other probe consisted of a 300
\( \mu m \) to 100 \( \mu m \) tapered excitation fiber and 4 collection fibers of 200 \( \mu m \) diameter each
as shown in Figure 8.14.

![Figure 8.13. A probe consisting of an input bundle of 10 excitation fibers of 100 \( \mu m \) diameter each and a 400 \( \mu m \) collection fiber. Some excitation fibers were misaligned.](image)

![Figure 8.14. A probe consisting of a 300 \( \mu m \) to 100 \( \mu m \) tapered excitation fiber surrounded by 4 collection fibers of 200 \( \mu m \) diameter each.](image)
The total collection area of the probes shown in Figures 8.13 and 8.14 was the same and, for reasons of symmetry, the effective excitation diameters as well when normalized for output power. A comparison of the working curves of the two probes would therefore reveal the most effective probe geometry. The working curves for these two probes are represented in Figure 8.15.

![Graph showing working curves of two probes](image)

**Figure 8.15.** *The working curves of the 2 probes shown in Figures 8.13 (input bundle, dotted line) and 8.14 (tapered input, solid line), respectively. The working curves were normalized with respect to output power.*

It is evident from the figure that the ((taper)300 to 100)x((4)200) probe was far superior to the ((10)100)x(400) probe, especially for highly absorbing analytes. Because the collection area of the tapered probe was distributed around the excitation cone, light scattered from the entire excitation volume could be collected. Where the excitation fibers were surrounding the collection area, however, light scattered from part of the excitation volume of each fiber could not be collected. An additional advantage of the tapered
excitation fiber probe was the greater ease with which light could be coupled into the fiber. Unfortunately, the long tapered section (2 m), precluded its use in the deep UV due to the increased attenuation in this region.

8.4.1.5. A side-casting bundle.

A probe consisting of one excitation fiber of 300 μm diameter and 8 angled collection fibers of 200 μm diameter was built and is shown in Figures 8.6 and 8.7. The probe could also be used with the angled fibers as excitation fibers, and it was tested in the latter configuration. This configuration was preferred, because the excitation radiation is cast across the surface of the collection fiber, thus minimizing the potential for tissue damage with in vivo use. The angled facets were given a reflective coating of aluminum by vacuum deposition. Because solarizing-resistant fiber was not available when this probe was fabricated, its performance deteriorated rapidly. A probe of this design, but using solarizing-resistant excitation fiber(s), will be evaluated in future. A comparison of side-casting and side-collecting probes would also be informative and remains to be done.

8.5. DISCUSSION.

Although the beam profile simulation contained several approximations, it was useful for predicting qualitative effects, could accommodate different probe tip geometries, and could show the exit beam intensity profile. On the other hand, the more detailed simulation produced excellent quantitative results as shown in Figure 8.16.

The detailed probe simulation was sufficiently general in nature to be used with any analyte provided the probe geometry was known (or selected), the molar absorptivities of the analyte and medium at the excitation and collection frequencies were known, and the excitation fiber beam profile and the collection fiber beam profile (or estimates thereof) were known (Greek et al., 1996). A drawback of this model was its
inability to simulate probes with angled fibers and it is hoped that this capability will be added.

Figure 8.16. The working curve (closed symbols) and detailed simulated working curve (open symbols) of a probe with 300 \( \mu \text{m} \) excitation fiber diameter and 600 \( \mu \text{m} \) collection fiber diameter. The working curves were normalized with respect to output power.

Regarding the probe fabrication process, one concern is the biocompatibility of silastic suspensions in cyclohexane with tetradecane added. The cured silastic should be analyzed for traces of cyclohexane and tetradecane to determine if harmful levels of these substances are present and could be released from the silastic matrix \textit{in vivo}. If present, behavioral studies should ensue to determine whether adverse effects result from the implantation of probes made with such silastic suspensions.

The results from investigating different probe geometries experimentally and by simulation yielded important information for optimizing the design of a probe for use in an \textit{in vivo} environment. The results from the simulation (Table 8.1) indicated that for a given
diameter collection fiber a small excitation fiber produced the best efficiency. This finding was confirmed by experiment (see Figure 8.11). Experiment also indicated that for a given diameter excitation fiber, an optimum collection fiber diameter existed (see Figure 8.10). Simulation showed that the maximum efficiency in absorbing solutions was obtained when the excitation and collection volumes were in close spatial proximity and had a high degree of overlap (see Figure 8.2 and Table 8.2). These findings were confirmed by the measured working curves for front- and side-casting probes using visible and near-UV excitation (Figure 8.12) and could be extended to the deep UV as shown in Figure 8.17 below (provided an appropriate non-solarizing excitation fiber was used).

Figure 8.17. The working curve for a side-collection probe using pulsed ultraviolet resonance Raman spectra of tryptophan dissolved in water excited at 227 nm, 20 Hz, and 300 μW average power. The tryptophan signal (solid line) shows a maximum near 80 μM while the water signal (dotted line) declines monotonically.
Taken together, the following requirements for constructing an efficient probe were established:

(i) small excitation fiber(s);
(ii) larger collection fiber(s);
(iii) confinement of the excitation volume close to the collection area; and
(iv) a collection area that can collect light from the entire excitation volume.

From the preceding information it was deduced that a probe with several angled excitation fibers and a single central collection fiber (e.g. Figure 8.6) would be the most suitable for use in a highly absorbing in vivo environment. As mentioned above, this geometry was preferred to using the central fiber as an excitation fiber, due to the negligible likelihood of damaging brain tissue by direct exposure to high energy light. A possible additional advantage is the formation of a resonance cavity resulting in moderate gains in signal intensity. The use of an input bundle would also avoid high fluences and subsequent energy-related damage and attenuation as mentioned before.

Drawbacks to using the angled fiber for excitation were the possibilities of incurring damage to the reflective coating with the higher excitation light intensities and of coupling light directly from the excitation fiber into the collection fiber. The former could be overcome with the use of pulse extender, and the latter with careful placement of the collection fiber, by shaping the angled tip so as not to cast light directly into the collection fiber, and by using a suitable filter to reject the excitation light.

Further consideration of the working curves of different probe geometries suggest that the efficiency of a probe identical to that depicted in Figure 8.6 could be enhanced by using the central fiber and 4 of the angled fibers as collection fibers. One of the angled fibers should be an excitation fiber and the fiber diametrically opposite should be a reflection fiber to reflect the light emitted from the excitation fiber back across the cavity. The reflection fiber should have a flat facet polished on its side perpendicular to the
surface of the central collection fiber. Due to the use of a single excitation fiber, high pulse
energies cannot be used, except in conjunction with a pulse extender.

The models and empirical investigations described in this chapter have provided an
exceptional understanding of the factors influencing probe performance. Integrating these
models would produce even greater advantages and allow the design of optimal probes for
neurotransmitter measurement under various conditions.

The work presented in this chapter demonstrated, both by experimentation and
simulation, that highly sensitive optical fiber probes of relatively small diameter could be
designed and fabricated for use under a variety of conditions, thus paving the way for
commencing probe-based *in vitro* and *in vivo* studies.

8.6. REFERENCES.


Beck, T.; Reng, N.; Richter, K. 1993. Fiber type and quality dictate beam delivery
characteristics. Laser Focus World: October, 111-115.

De Hart, T. 1992. Where and why optical fibers fail... and how to prevent it. Photonics


Rational design of fiber-optic probes for visible and pulsed-ultra-violet resonance
Raman spectroscopy. Applied Optics, in press.

for spectroscopy. Raman Scattering, Luminescence, and Spectroscopic

Kercel, S. W.; Roberts, M. J.; Garrison, A. A. 1990. Recent advances in the development
of a fiber-optic based instrument for on-line Raman analysis. Raman and

Ma, J.; Li, Y. S. 1994. Optical-fiber Raman probe with low background interference by


9.1. INTRODUCTION.

Part IV represents yet another important perspective from which the development of a neuroprobe should be viewed, namely that of its implantation and performance in, and its effects on, an actual living brain. To this end it was necessary to develop a probe of dimensions and materials suitable for in vivo use, and to investigate the possible consequences of in vivo use on the probe and on the brain as described in the next chapter.

The purpose of the research in the present chapter was to determine whether the probe was ready for in vivo testing by first testing it in vitro with samples mimicking in vivo conditions. To find such samples, the work in this chapter describes how neurotransmitter secretions were obtained from cell cultures and measured with high performance liquid chromatography (HPLC). The HPLC measurements were intended to determine whether the samples were of suitable complexity and physiological relevance to mimic the extracellular fluid and to serve as controls for eventual probe measurements of these samples. The reader may wish to refer to Chapter 1 which serves as a general introduction and a reference chapter for this section of the thesis.
9.2. PC-12 CELLS.

Rat PC 12 pheochromocytoma cells originated from an adrenal medullary tumor, respond to nerve growth factor, secrete a variety of neurotransmitters, and multiply logarithmically under routine culture conditions with a 48-72 h doubling time (see Greene and Tischler, 1982). Of the neurotransmitters, PC12 cells synthesize, store, and secrete large quantities of dopamine and norepinephrine, but also smaller amounts of acetylcholine and γ-amino-butyric acid (see Greene and Tischler, 1982). Mixtures of a small number of neurotransmitters at physiologically relevant concentrations can therefore be obtained from secretions of PC12 cells.

9.3. MEASUREMENT OF NEUROTRANSMITTER SECRETIONS.

A PC12 cell line was cultured in 4 wells of a 6-well plate (Nunclon 152795 from GibcoBRL, Burlington, Ont.) in a DMEM base with 10% fetal bovine serum and 5% horse serum. After several weeks, cells in 2 of the wells were used for counting under a microscope employing Trypan-blue stain to distinguish non-viable cells. Approximately 2x10^6 cells were counted per well. The cells in the remaining two wells were depolarized to obtain secretions for measurement.


Cells were washed with phosphate buffer solution before depolarization. At different intervals after polarization, samples of 1 ml were drawn from each of the wells, filtered through a 0.22 μm sterile filter into sample tubes and deposited on ice for transportation purposes and stored at -70°C until analysis.

Samples were taken immediately after washing with phosphate buffer solution, and again 15 minutes later, immediately prior to depolarization. This would reveal whether washing caused neurotransmitter release by virtue of mechanically stimulating the cells. Another pair of samples were taken 5 minutes after depolarization with 6 mM
KCl and 2 mM BaCl₂ in phosphate buffer solution at pH 7.2 (Greene and Rein, 1977) and again 10, 20, and 30 minutes after depolarization to yield 12 samples in total. Samples were taken at different intervals after depolarization to determine when the neurotransmitter concentrations reached their maxima.

9.3.2. High performance liquid chromatography.

Within 1-2 h after samples were taken, 5 μl from the samples were injected into a liquid chromatograph (Waters HPLC with 501 HPLC pump, U6K injector, 460 electrochemical detector, Maxima software 3.3, and Beckman ultrasphere column (5 μm, 15 cm, 2 mm i.d.)) for analysis. The instrument was calibrated with 4 μM each of norepinephrine, epinephrine, and dopamine, and 2 μM of serotonin. The results are shown in Figure 9.1.

Figure 9.1. Secretions of neurotransmitters (dopamine, top; serotonin, center solid line; epinephrine, dotted line; norepinephrine, dashed line) from rat PC12 cells, measured with a high performance liquid chromatograph, as a function of time.
It is evident from Figure 9.1 that much larger amounts of dopamine were secreted than of serotonin, norepinephrine, or epinephrine. It is also evident from the figure that washing the cells with phosphate buffer caused the secretion of dopamine and serotonin. The decline in the concentrations of these two neurotransmitters over the 15 minute period prior to depolarization may reflect the presence of uptake mechanisms similar to that for norepinephrine which is known to exist in these cells (e.g. Greene and Rein, 1977).

An inspection of Figure 9.1 further reveals that the secretion of serotonin was nearing a plateau after 30 minutes of depolarization, but that the concentration of dopamine was still increasing at this point. A longer time-course study should be undertaken to find the dopamine maximum. However, the existing samples were suitable for detection of dopamine at the micromolar-level with the resonance Raman neuroprobe. The changing neurotransmitter levels also made the application of difference spectroscopy to identify the components of the mixture possible.

9.3.3. Resonance Raman spectroscopy.

At the time of performing resonance Raman spectroscopy on cell secretions, the detection limit of the probe was approximately 10 µM for tryptophan and approximately 30 µM for tyrosine given 90 s of integration time. In order to increase the concentration of neurotransmitters in the cellular secretions to a level guaranteed detectable with the probe in a reasonable period of time (about 3 minutes), it was decided to grow more cells and induce them to secrete in the same sample volume. The results reported above showed that 1x10^6 cells produced a dopamine concentration of 3 µM in 1 ml of sample volume. A hundred-fold increase in dopamine concentration would therefore require about 1x10^8 cells secreting neurotransmitters into 1 ml of depolarizing solution. Cells were grown in 2 T-flasks (Nunclon 156502, GibcoBRL) with a surface area 17.5 times
that of a single well (of a 6-well plate) in the medium as described above, but also containing the anti-microbial agents penicillin, amphotericin B, and streptomycin (GibcoBRL #15240). When confluent, the medium was removed, 5 ml trypsin solution (0.25% trypsin-EDTA GibcoBRL 25200-056) was added (and the flask returned to the incubator for 5 minutes) in order to detach the cells from the flask bottom. The cell suspension from both flasks were then pooled in a 50 ml centrifuge tube, the flasks rinsed with 5 ml of fresh medium each and combined with the suspension to give 20 ml total of cell suspension. The cells in the suspension were declumped by pipetting them a few times back-and-forth in the centrifuge tube. From this suspension, 1 ml was taken and diluted 1:9 with medium. Trypan blue (50 ml) was then added to 50 ml of the diluted cell suspension and the mixture used to count the cells in a hemocytometer (Hausser Scientific Fuchs Rosenthal Ultra Plane 1/10 mm², 2/10 mm deep, from PGC Scientifics, Gaithersburg, MD.). The number of viable and non-viable cells counted in this manner are given in Table 9.1 below.

<table>
<thead>
<tr>
<th>Area</th>
<th>Viable</th>
<th>Non-viable</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>36</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>36</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>45</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>50</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>54</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>56</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>29</td>
<td>2</td>
</tr>
</tbody>
</table>

Average: 43.7  5.4

Given a multiplication factor of $1 \times 10^4$ for the hemocytometer, there were 4.37x10⁵ cells/ml in the diluted sample, 4.37x10⁶ cells/ml in the original sample which now consisted of 19 ml, thus giving a total of 8.30x10⁷ cells. This figure is a lower
bound on the number of cells, since a few cell clumps were present and these were counted as 1 cell.

The cells were separated from the medium by centrifuge (1500 r.p.m. for 5 minutes), washed with 5 ml physiological saline, centrifuged, the supernatant removed, and the washing sequence repeated. The cells (0.1 ml) were resuspended in 1 ml of depolarizing solution (6 mM KCl and 2 mM BaCl₂ in physiological saline) for 30 minutes, centrifuged (10,000 r.p.m. for 5 minutes), and approximately 1 ml of supernatant removed. The supernatant (sample A) was filtered with a 0.2 μm filter (Gelman Sciences 4192, Ann Arbor, MI.) and 0.4 ml thereof (sample B) ultrafiltered (Millipore Ultrafree UFC3LGC25 with 10,000 NMWL, Nepean, Ont.) for 80 minutes at 5,000 r.p.m. An aliquot of 100 μl was taken from each of samples A and B and frozen at -70 °C until HPLC analysis could be done 5 days later. The remaining samples were put on ice pending resonance Raman spectroscopy.

The HPLC analysis revealed much higher neurotransmitter levels in the samples compared to those obtained earlier. The results are shown in Table 9.2.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Norepinephrine</th>
<th>Epinephrine</th>
<th>Dopamine</th>
<th>Serotonin</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2.84x10⁻⁷</td>
<td>3.60x10⁻⁷</td>
<td>6.41x10⁻⁵</td>
<td>1.55x10⁻⁶</td>
</tr>
<tr>
<td>B</td>
<td>2.27x10⁻⁷</td>
<td>2.91x10⁻⁷</td>
<td>5.45x10⁻⁵</td>
<td>1.28x10⁻⁶</td>
</tr>
</tbody>
</table>

A chromatogram of the results (sample B) are shown in Figure 9.2 and on an expanded scale in Figure 9.3. It is evident from both figures that unknown compounds were present in the samples. The origin of these compounds was most likely the cell culture medium, but they could also originate from the cells themselves, either by secretion or through damage sustained during washing and/or depolarization.
Figure 9.2. A high performance liquid chromatograph trace of neurotransmitter secretions from rat PC12 cells taken after 30 minutes of depolarization. Sample A (1.5 µl injected, solid line) was filtered with a 0.2 µm filter and sample B (3.0 µl injected, dotted line) with an ultrafilter.

Figure 9.3. A high performance liquid chromatograph trace of neurotransmitter secretions from rat PC12 cells taken after 30 minutes of depolarization, as in Figure 9.2, but enlarged 4 times.
The spectra of samples A and B were measured with the probe using excitation at 227 nm (optimal for dopamine) and 218 nm (possibly optimal for serotonin) and average power at the sample of approximately 200 µW. Monochromator settings were 234.4 nm 224.8 nm, respectively, the entrance slit open to 300 µm, and a 90 s integration time was used. Spectra of pure ethanol (for calibration) and depolarizing solution (for background correction) were first obtained at both frequencies. The resonance Raman spectra obtained from sample B are shown in Figure 9.4. Spectra collected at 227 nm and 218 nm from sample B are shown in Figure 9.5 after subtraction of the solvent spectrum and after baseline correction using the method discussed in Chapter 12.

![Resonance Raman spectra of neurotransmitter secretions from rat PC12 cells taken after 30 minutes of depolarization (dotted line) and depolarizing agent (solid line, scaled down by a factor of 2 for comparison). Spectra were obtained with 227 nm excitation.](image-url)
Figure 9.5. Resonance Raman spectra of neurotransmitter secretions from rat PC12 cells taken after 30 minutes of depolarization: 8 spectra obtained with 227 nm excitation added together (top line); 3 spectra obtained with 218 nm excitation added together (bottom line). Backgrounds were removed.

In the spectra above, the peak near 1615 cm$^{-1}$ corresponded with that of dopamine. It is also clear that a number of other peaks were present and these could not be identified.

9.4. DISCUSSION.

The experiments reported in this chapter indicated that secretions from PC12 cell cultures could be used to test the performance of the probe in vitro to determine its readiness for in vivo use. Although resonance Raman signals corresponding to that of dopamine and serotonin, both shown to be present in the samples by HPLC analysis, were obtained with the probe, the presence of unidentified components in the secretion samples precluded the unambiguous assignment of these peaks. In order to control for the
possibility of unknown components originating from the cell culture medium, the
experimental procedure was modified. The results of the modified experiment,
representing a more comprehensive application of the neuroprobe method including the
use of neural networks to identify secretion components, are reported in Chapter 16.

Experimental conditions did not permit longer integration times or the collection
of additional spectra for superaddition, however, either of these strategies could be used
to substantially enhance the resolution of these measurements, notwithstanding
limitations due to pulse length, repetition rate, or spectrograph efficiency.

9.5. REFERENCES.

cell line of nerve growth factor (NGF) responsive pheochromocytoma cells Brain
Research 129: 247-263.

Greene, L. A.; Tischler, A. S. 1982. PC12 pheochromocytoma cultures in
CHAPTER 10

10.1. INTRODUCTION

It would be futile to develop a neuroprobe, capable of meeting the stringent criteria of in vivo measurements delineated in Chapter 1, that is difficult to use in the brain due to its awkward geometry or extreme dimensions, or that produces excessive brain damage upon implantation or operation, or that produces behavioral interference, or that cannot be used long enough to obtain meaningful data. Therefore the objectives of these studies were to develop and test a prototype brain/probe interfacing assembly to ensure a functional match between brain and probe, to test the operation of the probe under in vivo conditions on some basic performance parameters, and to determine whether the probe produced any behavioral effects.

10.2. THE BRAIN-PROBE INTERFACE

10.2.1. Interface assembly design

Manipulation of the central nervous system is often effected by electrical or pharmacological means. Cranial assemblies with indwelling cannulae have been developed for use with removable electrodes (e.g. Schulze and Gorzalka, 1994) and/or
infusion needles at single (Gray and Gorzalka, 1979) or multiple sites (Schulze and Gorzalka, 1990) to permit such central nervous system manipulations. Therefore, a solution to the basic brain-probe interface already existed if these cranial assemblies could be utilized. Given that the fiber-optic probes developed and tested in Chapter 8 were of the same general dimensions (~ 500 µm) as the electrodes/infusion needles used with the cranial assemblies mentioned above, these cranial assemblies were readily adapted for use with the fiber-optic probes.

For fitting the probe, a probe-and-sleeve system was designed as shown in Figure 10.1.

![Diagram](image)

Figure 10.1. A schematic diagram of the protective sleeve designed for inserting the fiber-optic probe into the centrally implanted guide-cannula.

The sensing tip of the fiber-optic probe was inserted into a 12 mm stainless steel tube of 22 gauge until the probe tip was in the center of the bevel. The probe was fixed in place at the dorsal end of the tube with a drop of cyanoacrylate glue (LePage 8, UBC
The tube acted as a protective sleeve against damage of the probe during insertion into the cannula. A 19 gauge collar on the tube allowed it to be inserted into the cannula to the correct depth, i.e. just below the ventral tip of the implanted cannula.

10.2.1. Placement coordinates.

Surgeries were performed in order to fit experimental rats with indwelling cannulae through which a fiber-optic resonance Raman neuroprobe could be inserted to measure neurotransmitter levels. These were preceded with trial surgeries to assess the proper brain coordinates for cannula placement.

The initial rostro-caudal coordinates relative to bregma (a position on the skull where 4 cranial bones fuse) were -0.8 mm, medio-lateral relative to the midline -1.6 mm, and dorso-ventral relative to the top of the skull -3.3 mm. These coordinates for the left lateral ventricle were obtained from the atlas of Paxinos and Watson (1982) which is calibrated for male rats weighing 250-350 g. Male rats of the Wistar strain derived from stock obtained from Charles River Canada Incorporated (Montreal, Que.), were used for the trial cannulations. These rats were larger than those referenced in the atlas and ranged between 600-700 g each. Therefore, coordinates appropriate for cannula placement in the left ventricle of these animals had to be extrapolated. The rats were anesthetized with 0.75 ml/kg sodium pentobarbital (Somnotol) and 0.33 ml/kg ketamine (Ketalean), both from MTC Pharmaceuticals, Cambridge, Ontario, given intraperitoneally. During anesthesia each eye was protected with a drop of mineral oil.

After surgery the animals were perfused with 20 ml of physiological saline followed by 20 ml of phosphate buffered formalin, decapitated, the cannula removed, and the brains removed and stored in formalin overnight. The brains were then quick frozen, sectioned on a microtome in 30 μm sections, and collected on microscope slides before
staining with cresyl violet (Nissl stain). The stained slides were inspected under a microscope to determine the location of the cannula tip.

After 3 trial surgeries, the final anterior-posterior, medial-lateral, and dorsal-ventral coordinates for cannula tip placement in the left lateral ventricle of 600-700 g male Wistar rats were established as -0.3, -1.4, and -4.4 mm, respectively. Figure 10.1 shows a histological slide from a trial surgery with the cannula tract entering the left lateral ventricle. Of special importance in Figure 10.2 are the relative dimensions of the cannula tract (i.e. probe dimensions) and large brain structures such as the ventricles.

![Histological slide showing cannula tract and ventricle](image)

Figure 10.2. A digitized histological slide from a trial implantation into the lateral ventricle of a male rat showing the cannula tract entering the left lateral ventricle. The slide was digitally enhanced to demonstrate the relative scales of the cannula tract (about 1,100 \( \mu \)m outside diameter) and the ventricles.

10.2.2. Cannulations.

Another 3 male rats of the same population as above were provided with cranial assemblies containing chronic indwelling cannulae according to the procedures described by Gray and Gorzalka (1979) using 19 gauge stainless steel cannulae of 9 mm length. The
animals were concurrently fitted with restraining devices as described by Schulze and Gorzalka (1990). The restraining devices were provided to prevent freely-moving animals from putting strain on the probe during measurement.

After surgery the animals were individually housed in recovery cages overnight before being returned to their home wire mesh cages. Food and water were available ad libitum, a reverse day-night cycle followed, and the colony room maintained at 21°C and 50% relative humidity.

10.3. IN VIVO PROBE EVALUATION.

The functionality of the probe was tested by collecting Rayleigh scattering from the excitation volume in vivo. Rayleigh scattering is about 8 orders of magnitude stronger than Raman scattering and an inability to detect Rayleigh scattering in vivo would indicate some serious problem with the probe or its insertion.

10.3.1. Rayleigh scattering in the lateral ventricle.

One of the male rats fitted with an indwelling cannula was anesthetized as described above and placed on the optical table containing the laser equipment and probe as in Figure 10.3. Because thermoregulation in anesthetized animals is impaired, the rat was placed on a bed of paper towels to prevent hypothermia by insulating the animal from the metal table top. The 473 line of the Ar⁺ laser was coupled into a 200x200 probe. The input power was 200 mW and the output power 50 mW, the monochromator set to 500.8 nm center position and the entrance slit at 200 μm. Spectra taken consisted of 5 accumulations of 0.989 ms integration each and 10 sequential acquisitions.
The performance of the probe was checked by immersing it in 0.153 M KNO₃ which served as the standard solution. After having been rinsed with deionized water, the probe was inserted into the protective tube, until the tip became just visible upon exit, before insertion of the assembly into the guide cannula (see Figure 10.2).

Spectra were collected from the left lateral ventricle of the rat with the laser output power turned down to 60 mW and power at the sample estimated at 15 mW (25% coupling efficiency) to reduce the likelihood of causing brain lesions in the animal.

The spectra taken with the monochromator setting at 1187, 1080, and 896 cm⁻¹, respectively, are shown in Figure 10.4. The slope of the Rayleigh line clearly increased with monochromator settings approaching the excitation frequency, as would be expected, indicating effective detection of Rayleigh scattering in vivo. A high background...
is furthermore evident from Figure 10.4 and possibly some fluorescence in the 896 cm⁻¹ trace. Taken together, these results suggested that the cerebrospinal fluid was neither excessively absorbing at this wavelength, nor very fluorescent. The high background was most likely due to back reflection from the ventricular walls.

![Figure 10.4. Rayleigh scattering obtained in vivo from an anesthetized rat with 473 nm excitation and different monochromator settings: 896 (top), 1080 (middle), and 1187 cm⁻¹ (bottom).](image)

10.3.2. Probe fouling.

The protein content of cerebrospinal fluid is 0.035% by weight (Rowland, 1985), a value much lower than that of serum. In order to determine the extent to which protein fouling of the probe may occur in cerebrospinal fluid, two spectra at a monochromator setting of 1187 cm⁻¹ were taken approximately 25 minutes apart. These spectra are shown in Figure 10.5 and revealed a reduction in signal intensity. The powers at the sample corresponding to these spectra were 15 mW and 10 mW, respectively, while the
laser power remained stable at 60 mW. These results indicated that some fouling of the probe sensing tip did occur after approximately 30 minutes in vivo, but the probe remained operative. Inspecting the probes under magnification after different periods of insertion time would be a worthwhile future task.

![Figure 10.5. Two spectra obtained in vivo with 473 nm excitation and the same monochromator setting: (top) initial spectrum and (bottom) 25 minutes later.](image)

An attempt was made to conservatively quantify the possible lifetime of the probe in vivo by using the data from diodes 200-300. The mean ± S.E.M. for the earlier spectrum was 18606 ± 30 counts and for the one 25 minutes later 15335 ± 26 counts. With the background level being approximately 250 counts, and requiring a signal-to-noise ratio of 3 for the detection limit, a minimum signal intensity of about 250 + 3x30 (i.e. 340) counts was needed. Assuming a linear fouling rate, a maximum lifetime of about 2 h and 20 minutes in vivo was predicted. Given the possibility of a non-linear
fouling rate, it may still be reasonable to expect the lifetime of the probe to exceed 60 minutes *in vivo*.

Upon removing the probe, the obturator was replaced in the cannula thereby closing the entrance to the cannula. The animal was returned to a recovery cage and then its home cage. A cursory inspection revealed no motor deficits in the animal after recovery from anesthesia.

10.4. **POSSIBLE BIOLOGICAL EFFECTS OF *IN VIVO* SPECTROSCOPY.**

The ultimate aim of this project was to create a probe that could be used to correlate behavior with changing neurotransmitter levels. Given the novelty of this technique, it was not known whether it would produce any behavioral or cognitive deficits. It was therefore important to show that no behavioral deficits were produced by this procedure.

10.4.1. *In vivo* pulsed UV probe operation

A probe consisting of 5 side-scattering excitation fibers of 200 µm diameter surrounding 1 collection fiber of the same diameter was constructed from low UV attenuation fused silica fiber from Fiberguide Industries. The probe was then inserted into the brain of an anesthetized rat via the indwelling cannula to the depth of the cannula. Light at 230 nm, 10 Hz, and 20-140 µJ/pulse was coupled into the excitation bundle and delivered to the brain.

The probe performance was verified by closing the spectrometer slits to 5 µm and detecting the Rayleigh scattered light at the beginning and at the end of the experiment. The experiment consisted of collecting spectra with consecutive monochromator settings at 1000, 2000, 3000, and 4000 cm⁻¹ respectively. The 1176 accumulated spectra of approximately 30 ns duration each, collected at each monochromator setting, were processed using 500 autoaccumulations with a 50 diode normally displaced bandwidth (see Chapter 12). Autoaccumulation revealed some structure common to all spectra and it
was estimated with the 4000 cm\(^{-1}\) spectrum and removed from the other spectra. The resulting spectra are shown in Figure 10.6 and the Rayleigh wing can be discerned in the top spectrum (center position 1000 cm\(^{-1}\)).

At the end of the experiment (after 33 minutes) the probe was removed from the brain, the obdurator replaced into the cannula, and the animal returned to a recovery cage and after overnight recovery, to its home cage.

![Figure 10.6. Spectra obtained in vivo from an anesthetized rat with pulsed 230 nm excitation and 3 different monochromator settings after processing: 1000 cm\(^{-1}\) (top), 2000 cm\(^{-1}\) (middle), and 3000 cm\(^{-1}\) (bottom). Spectra are vertically displaced for ease of viewing.](image)

### 10.4.2 Motor deficits.

Alteration of the ongoing activity of central neurochemical systems often produces behavioral effects. It is known that activity in certain neurotransmitter systems affects certain behaviors. For instance, modulation of the serotonergic system has effects
on sexual behavior (Gorzalka et al., 1990), pain perception, and sleep; of the cholinergic
system on drinking; and of the dopaminergic system on locomotor activity - to name but
a few (see Grupp and Kalant, 1985).

Therefore, if the operation of the resonance Raman neuroprobe in vivo produced
pronounced disturbances in any neurochemical system, it is likely to be detected with
some behavioral screening test. Given that at least 4 of the 10 small-molecule
neurotransmitters, acetylcholine, dopamine, norepinephrine (see Grupp and Kalant,
1985), and serotonin (Green and Backus, 1990) affect locomotion, it seemed essential to
determine if motor deficits occurred. Therefore, the open field assay was used to detect
hypo- or hyperactivity (e.g. Kohlert and Bloch, 1996). Changes in the dopaminergic
system can also produce stereotypy often consisting of persistent licking and gnawing
(Kilbey and Sannerud, 1985). Grooming behavior was therefore also monitored to detect
the induction of stereotypy. In addition, rearing behaviors were recorded.

Two animals were tested in an open field measuring 1 m x 1 m x 30 cm and
divided into 9 equal squares. The number of times an animal crossed from one square into
another was recorded over a 30 minute period while simultaneously monitoring the
number of grooming bouts and number of times the animal reared up against the wall of
the open field. Animals were habituated to the open field for 10 minutes per day for 4
days prior to testing. The animals were tested on 4 consecutive days prior to in vivo
spectroscopy. One of the animals died shortly after receiving the anesthetic and the
remaining animal was tested again on 4 consecutive days after in vivo spectroscopy.
Dependent t-tests were performed on the scores for each behavior before and after the
spectroscopy. No significant differences were found in any of the behavioral measures
taken before and after probe operation. Therefore, probe operation produced no
discernible motor effects. The results are shown in Table 10.1.
Table 10.1

The means, standard errors of the means, and probabilities ($\alpha = 0.05$) of 3 different behavioral measures obtained on 4 consecutive days before and after in vivo operation of the UV resonance Raman neuroprobe.

<table>
<thead>
<tr>
<th>Behavior</th>
<th>Mean</th>
<th>S.E.M.</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crossing</td>
<td>4.250</td>
<td>0.629</td>
<td>0.141</td>
</tr>
<tr>
<td></td>
<td>3.000</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Grooming</td>
<td>4.250</td>
<td>0.479</td>
<td>0.638</td>
</tr>
<tr>
<td></td>
<td>5.500</td>
<td>2.5331</td>
<td></td>
</tr>
<tr>
<td>Rearing</td>
<td>1.000</td>
<td>0.408</td>
<td>0.761</td>
</tr>
<tr>
<td></td>
<td>0.750</td>
<td>0.479</td>
<td></td>
</tr>
</tbody>
</table>

10.4.3. Brain lesions.

The brain of the remaining animal was recovered, sectioned, and stained as described above and the sections examined for brain damage due to laser irradiation. No damage attributable to laser irradiation was observed in the sections, although some scar tissue surrounding the cannula was evident. The scar tissue was likely due to the long interval since cannulation (9-10 months).

10.5. DISCUSSION.

The brain-probe interface assembly designed for this neuroprobe seemed to function adequately. A minor problem arose with the tendency of the probes to break at the level of entry into the protective sleeve. This could be overcome by attaching a polyethylene sleeve of about 10 mm to the stainless steel sleeve to provide more support for the fiber. Provided the probe dimensions could be kept similar to those tested, no unusual problems are foreseen inserting or removing the probe.

It was also possible to detect Rayleigh scattering with a fiber-optic probe from within the living brain using cw visible excitation at 472.7 nm. Furthermore, the 896 cm$^{-1}$ trace in Figure 10.4 exhibited a number of discrete features, which could have been Raman lines. The laser was detuned to check for the presence of plasma lines among these features and the one near 660 cm$^{-1}$ was found to be due to the Ar$^+$ emission at 488
nm. However, the presence of a high background and the absence of much fluorescence, combined with the presence of possible Raman lines, suggested that it was worthwhile to further investigate *in vivo* Raman spectroscopy. This could be done with the aid of an interference filter to remove plasma lines and a Raman notch filter to remove Rayleigh scattering.

The extent to which probe fouling may occur *in vivo* has also been estimated from the collected data shown in Figure 10.4. This estimate was 60 minutes or more, a period of time sufficient for many behavioral measurements. Polymer coatings (e.g. McGlashen et al., 1990; Morris et al., 1990; Rosenberg et al., 1994) can also be employed to reduce protein fouling, but they will restrict the number of compounds that can be measured and may furthermore introduce kinetic effects. These coatings may also become opaque in the deep UV and would therefore not be suitable for pulsed UV resonance Raman systems. The best way to deal with probe fouling may be to periodically remove the probe and clean it in a suitable aqueous solution.

Both animals tested with functional probes (one with cw visible and the other with pulsed deep UV radiation) showed no noticeable behavioral deficits after probe testing. It is known that brain lesions in some brain areas, e.g. the striatum and cortex, produce elevated open field activity (Robinson and Coyle, 1979). However, no elevated open field activity was noticed. The side-casting probe design, lack of observed brain damage, and lack of observed behavioral effects all argue against probe-induced damage. Although to obtain greater confidence in this neuroprobe a more extensive battery of behavioral and cognitive tests on more animals should ultimately be undertaken, these findings do allow development to proceed with cautious optimism.
10.6. REFERENCES.


PART V
Signal processing

CHAPTER 11

11.1. INTRODUCTION

The general approach taken in the development of this resonance Raman neuroprobe was that all data available in situ be collected and mathematically post processed for species identification and quantification (as stated in Chapter 1). Had selectivity been sought at the measurement level, much useful data would not be collected at all. Most attention was focused on attaining the required sensitivity from a probe small enough for in vivo use and little design attention (beyond using tunable excitation) was given to obtaining selectivity in situ. The advantage of this approach was that data could be selected after collection for further processing as needed. The disadvantage was that more post-processing of the measurements was required (Erickson et al., 1992). Therefore, it was essential to investigate different signal processing methods to generate post-collection selectivity, to resolve mixtures, and, due to the weakness of Raman signals, to detect weak signals in noisy backgrounds. In Part V of this thesis, the
emphasis is shifted to this fourth and final discipline involved in the development of the resonance Raman neuroprobe. In this chapter, the theoretical basis for Part V is established, while the remaining chapters of Part V contain some methods for signal recovery and enhancement, as well as neurotransmitter classification and quantification.

11.2. SIGNAL RECOVERY

The data measured by a spectroscopist can be modeled as:

\[ D_i = \Sigma B_{ik}(f_k + b_k) + \sigma_i \]

\[ = B^*(f_i + b_i) + \sigma_i \]  \hspace{1cm} (11.1)

where the \( D_i \) are the measured data points, the \( f_k \) represent the data points of the underlying spectrum and the \( b_k \) those of the background convoluted (*) with the instrumental response or blurring function \( B \) (of bandwidth \( k \)), and the \( \sigma_i \) the standard deviation of the noise. The basic problem confronting the spectroscopist is how to recover, identify, and quantify the underlying spectrum \( f_i \) from the imperfect, incomplete, and contaminated measured data.

An inspection of Equation 11.1 suggests a 3-step approach to recover the perfect underlying spectrum \( f_i \): (i) removal of the noise \( \sigma_i \); (ii) deconvolution (deblurring) of \( \Sigma B_{ik}(f_k + b_k) \) to obtain \( (f_i + b_i) \); and (iii) removal of the background \( b_i \). Each of these will be briefly discussed below.

11.2.1. Noise removal

Methods for removing noise from a signal depend greatly on the character of the noise. Perhaps the easiest way to reduce the random noise content of a signal is to allow the noise from different measurements to partially cancel itself. For independent noise, the variance of a sum equals the sum of the variances (e.g. Glass and Hopkins, 1984). Therefore, by accumulating \( n \) measurements, Equation 11.1 becomes:

\[ nD_i = nB^*(f_i + b_i) + (n\sigma_i^2)^{0.5} \]  \hspace{1cm} (11.2)
An inspection of Equation 11.2 reveals that the signal has increased by $n$ times and the noise by $n^{0.5}$ times, thus giving an improvement in the signal to noise ratio of $n^{0.5}$ (e.g. Demas and Demas, 1990).

A similar approach can be applied to successive frequency points in a spectrum by selecting a window of small size relative to the size of the spectrum and by replacing the center value of the window by the average of all the points contained within the window. The window is then moved one point ahead and the procedure repeated until the entire spectrum is processed. The rationale is that the average of a small number of adjacent points is a better measure of the signal than any single point (e.g. Skoog, 1985). There is however, a trade-off between noise reduction and loss of resolution which depends on the window size (e.g. Demas and Demas, 1990).

The moving window (also moving average or boxcar) method can be seen as a linear filter of the spectrum in that the average obtained from the data in the window consists of a simple linear or evenly weighted combination of points (e.g. Skoog, 1985). More complex polynomial relationships can be used to weight the data before obtaining the average and one of the most popular methods is the Savitzky-Golay method. This filter uses a window with an odd number of points and a least squares polynomial is fitted through them after which the center point in the window is replaced with the value of the polynomial at that point (e.g. Demas and Demas, 1990).

The filtering of the measured data $D_i$ of Equation 11.1 can also be seen as a convolution of the measurements with a (finite impulse response) filter function to give the noise filtered data:

$$y_i = D_i * h_i$$
$$= (B*(f_i + b_i) + \sigma_i) * h_i$$
$$= (B*(f_i + b_i))$$

(11.3)

For removing completely random noise, a matched filter can be constructed by assigning $h_i = c(f_i + b_i)$ where $c$ is a constant and $(f_i + b_i)$ represents the reversed pure signal
plus background uncorrupted by noise. The details can be found in Erickson et al., (1992). Matched filtering is inadequate for removing cyclic noise, and for this purpose the Kalman innovation filter is used. The measured data are "whitened" by orthogonalizing the cyclic noise and the rest of the data and applying an extension of the matched filter to the whitened data. Both methods thus require a priori knowledge of the signal (and background).

Another approach for processing multivariate measurements contaminated by noise and interferents is principal component regression (Erickson et al., 1992). Singular value decomposition (e.g. Scharf, 1991) and principal component analysis (Karhunen and Joutsensalo, 1995) can be used to partition the data space into a signal subspace and a noise subspace. If the higher order principal components represent only the noise in a system, their removal reduces the overall noise of the system as well as the dimensionality of the problem (Erickson et al., 1992). Where the pure signals are well known and the system is well understood, matched and Kalman innovation filtering produce superior results, however, principal component methods can estimate a larger range of properties at the cost of somewhat poorer noise reduction (Erickson et al., 1992).

11.2.2. Deconvolution

There are several methods to deconvolute a signal, for example, Fourier deconvolution (e.g. Press et al., 1994), maximum entropy deconvolution (Burch et al., 1983), and simulated annealing deconvolution (Ferry and Jacobsen, 1995). The Fourier and maximum entropy methods are briefly discussed below.

Because the Fourier transform of a convoluted signal is equal to the product of the Fourier transforms of the underlying spectrum with the blurring function, deconvolution by means of division in the Fourier domain and obtaining the inverse Fourier transform of the result is trivial. After noise removal, Equation 11.1 becomes:
\[ D_i = B*(f_i + b_i) \]  
Taking the Fourier transform of Equation 11.4 gives:
\[ F(B*(f_i + b_i)) = F(B)xF(f_i + b_i) \]  
Then:
\[ F(B*(f_i + b_i))/F(B) = F(f_i + b_i) \]  
and \((f_i + b_i)\) is obtained by taking the inverse Fourier transform.

However, the effectiveness of this process depends on the amount of noise in the input data (Equation 11.4 assumes no noise) and the accuracy with which the blurring function is known (Press et al., 1994).

Deblurring can also be accomplished with the maximum entropy method (Burch et al., 1983; Davies et al., 1991). This method estimates the underlying spectrum by minimizing a multidimensional function that is the sum of a smoothing function, a constrained goodness-of-fit function, and a constrained constraining function. The smoothing function used is the negative entropy (-S):
\[ -S = \sum p_i \ln(p_i) \]  
where \(p_i\) is the normalized spectral intensity at frequency \(i\). The goodness-of-fit function is the \(\chi^2\) function and the constraining function \(C\) is a function that relates the totals of the recovered and measured signals. The multidimensional function minimized is:
\[ M = -S + \lambda \chi^2 + \mu C \]  
where \(\lambda\) and \(\mu\) are Lagrange multipliers. Because an infinite number of estimates of \(f_k\) may be compatible with the measured data \(D_i\), that estimate with the least structure (or maximum entropy) should be selected to avoid an over-interpretation of the data. A knowledge of the blurring function or an estimate thereof is required.

11.2.3. Baseline removal

Raman signals are generally weak and often superimposed on a large fluorescence and/or Rayleigh background, hence the term \(b_k\) in Equation 11.1 is often significant. A
shifted excitation technique has been described for reducing the background due to fluorescence (Shreve et al., 1992). A dye laser is used to excite a Raman spectrum of a particular analyte at a certain wavelength and again at a wavelength red-shifted by 10 cm\(^{-1}\). The resulting spectra are then subtracted to reduce the fluorescence background. Alternatively, experimental or digital abscissa shifting could be used to achieve the same effect (Mossier-Boss et al., 1995). Maximum entropy has also been used to remove the fluorescence background by separate reconstructions of the Raman and background signals (Durman and Wood, 1988).

As mentioned above, singular value decomposition (e.g. Scharf, 1991) and principal component analysis (Karhunen and Joutsensalo, 1995) can be used to partition the data space into a signal subspace and a noise subspace. Iwata and Koshoubu (1994) applied singular value decomposition to a cyclic matrix constructed in the frequency domain from a symmetrically folded partial spectrum to separate the signal from the noise and background. It seems that it ought to be possible to construct separate signal, noise, and background subspaces with singular value decomposition and principal component analysis.

The easiest way of removing a large background is perhaps to fit linear segments to points chosen along the baseline and then removing all points below this segmented baseline. However, baselines are often irregular and non-linear and a better method consists of fitting a polynomial (e.g. Shreve et al., 1992) or exponential function to the baseline and subtracting the value of the polynomial from the corresponding value of the spectrum.

11.3. SIGNAL IDENTIFICATION AND QUANTIFICATION

When \( f_i \) is isolated from the equation

\[
D_i = B^* (f_i + b) + \sigma_i
\]  

(11.1)
by sequential noise removal, deconvolution, and background removal, it remains to be identified and quantified. For identification and quantification, several methods are possible, two of which will be introduced below.

11.3.1. Classical least squares

In practice, none of the methods employed to isolate \( f_1 \) are perfect and the value obtained for \( f_1 \) can be considered only as an estimate. This estimate can then be compared to a library of known spectra and identified on the basis of least squares. In the classical least squares model, the spectrum of a mixture is seen as a linear combination of the spectra of individual components of the mixture multiplied by the concentration of that component. For instance, if these components were A, B, and C, with spectra \( A, B, C \) and concentrations \([A], [B], \text{and} [C]\), respectively, then the spectrum of the mixture \( X \) can be considered to be given by the equation:

\[
X_i = [A]A_i + [B]B_i + [C]C_i
\]  

(11.9)

where \( i \) represents a particular point in the spectrum (e.g. a wavelength or wavenumber) and ranges across the spectrum. The use of linear or matrix algebra allows one to generate a mixture spectrum from the linear combination of the constituent spectra multiplied by their respective (but unknown) concentrations. This calculated mixture spectrum is then compared to the measured mixture spectrum and the concentration values ([A], [B], and [C]) are adjusted until a least squares fit is produced, thus giving the concentration estimates for the mixture components. This method requires a library of known spectra and a well-defined system to produce good results (see Erickson et al., 1992, and references therein).

11.3.2. Artificial neural networks

Trained artificial neural networks have the advantages of being very fast at spectral identification (Tanabe et al., 1992), and also noise and distortion tolerant (Lin et
al., 1992). These advantages are important in applications where real-time mixture identification is required and where poor or distorted spectra are to be processed.

In contrast, the classical least squares method needs to be well-defined (Erickson et al., 1992) and may be more susceptible to noise. Furthermore, it is not clear how sensitive this method is to spectral distortions. Such distortions can be the result of several factors, e.g. spectral blurring due to instrumental characteristics, accumulated error from difference spectra where the analytes of interest are present in a complex matrix, calibration errors generating spectral shifts, or the result of spectral data processing.

For these reasons, artificial neural networks were investigated for the purpose of spectral identification and analyte quantification.

11.3.2.1. Architecture

An artificial neural network consists of a number of interconnected computational elements called nodes. Such a network is not an amorphous collection of randomly interconnected nodes, but has a well-defined architecture. Most neural networks consist of three layers of nodes: the input; hidden; and output layers. In general, every node in the input layer is connected to every node in the hidden layer. Hidden and output layers are likewise fully connected. Furthermore, every connection has a certain weight or strength. Figure 11.1 shows an example of a neural network producing an output from experimentally obtained data (e.g. the Raman spectrum of dopamine).
Figure 11.1. This figure shows an artificial neural network with 250 input nodes, 30 hidden nodes, and 10 output nodes. Each output node corresponds to one of 10 neurotransmitters. Also shown are data from a dopamine spectrum being given as input to the network and a network output classifying the input as dopamine.

Data are introduced to the network at the input layer. These data are propagated, via the connections between the input and hidden nodes, to the nodes of the hidden layer. Here the data are combined and transformed before being relayed to the output nodes. At the output nodes, the data from the hidden nodes are further combined and transformed to produce the network output. From input to output then, the data go through a sequence of stepwise combinations and transformations, details of which will be provided in subsequent sections.

Every connection in the network is weighted to indicate the strength of the connection. Data propagated along a connection is multiplied by the connection weight. Because the connection strengths (or connection weights) between nodes are initially set to small random values, the output from an untrained network is nonsensical. In essence,
training consists of providing the network with a number of training "examples" and the corresponding "responses" required of the network. "Learning" occurs by using the differences between the network output and the desired output (this difference is called the output error) to adjust the connection weights between the nodes until the output error is minimized and an optimal set of weights found. In backpropagation networks, this adjustment occurs from the output layer backwards.

The examples in the training set are repeatedly presented to the network in a random order. Training is periodically stopped and the network is tested with a testing set of examples to cross-validate its performance. The examples in the testing set are often fewer in number and generally similar, but not identical, to those in the training set. A network is considered optimally trained, not when the output error generated by the training set reaches a minimum, but when the testing set output error reaches a minimum. This procedure prevents overlearning (the network learns the training set so well that nothing else is recognized) and improves generalizing (Schoner, 1992). Training is ideally terminated when the network generalizes well to new data and the cross-validation method is a way to estimate when this occurs.

Determining the appropriate number of hidden nodes requires some investigation. The number of training samples serves as an upper limit on the number of hidden nodes (see Hush and Horne, 1993). Others suggest using the larger of the number of input or output nodes or half their sum (California Scientific Software, 1992). Using half their sum, provided it is less than the number of training examples, is a good starting value for the number of hidden nodes. If the network does not train, the number of hidden nodes should be increased because there are probably too few connections in the network. On the other hand, networks with a large number of connections train more readily but often perform poorly on the testing set. Networks with many connections also take longer to train because of the need to adjust them all. If the network trains readily, the number of hidden nodes should be reduced until the minimum required to allow training is found.
Reducing the number of nodes will also reduce the number of connections. The number of input nodes can most easily be reduced by increasing the sampling frequency.

11.3.2.2. Layer computation

The computations performed by a neural network are equivalent to a number of matrix operations (Jordan, 1988). The matrix of input layer values is multiplied by the matrix of weights (between input and hidden layers) to give a matrix of hidden layer values. The matrix of hidden layer values are transformed with a transfer function before being multiplied with another matrix of weights (between hidden and output layers) to give the matrix of output layer values. The latter matrix values are transformed again with a transfer function to give the final output of the network. The ability of networks to perform non-linear computations are due to the use of non-linear transfer functions, mostly for the hidden and output layers. The most appropriate transfer function(s) for a particular application has to be determined by trial and error, however, the sigmoidal transfer function is frequently used and a good initial choice. The entire computational process can be schematically represented as follows:

\[
\begin{align*}
\text{[Input layer values]} \\
\text{TRANSFER 1} \\
\text{[transformed values 1]} \times \text{[weights 1]} \\
= \text{[hidden layer values]} \\
\text{TRANSFER 2} \\
\text{[transformed values 2]} \times \text{[weights 2]} \\
= \text{[output layer values]} \\
\text{TRANSFER 3} \\
\text{[final network output]}
\end{align*}
\]

Above, \textit{weights 1} refers to the connection weights between the input and hidden layers and \textit{weights 2} to those between the hidden and output layers. These weights are
initially randomly assigned and adjusted during training until good values are found. Also above, TRANSFER 1 refers to the transfer function between input and hidden layers, TRANSFER 2 to that between hidden and output layers, and TRANSFER 3 to that of the output layer. The transfer function between input and hidden layers is almost always a linear function causing [input layer values] to be identical with [transformed values 1]. A more detailed numerical example showing the correspondence to a neural network follows.

\[
\begin{bmatrix}
3 & 1 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
4 \\
\end{bmatrix}
\text{Input nodes.}
\]

\[
\begin{bmatrix}
10 & 11 \\
11 & 01 \\
11 & 01 \\
10 & 11
\end{bmatrix}
\begin{bmatrix}
8 \\
4 \\
4 \\
8
\end{bmatrix}
\text{Connection weights between 4 input and 2 hidden nodes.}
\]

\[
\begin{bmatrix}
14
\end{bmatrix}
\text{Hidden nodes.}
\]

\[
\begin{bmatrix}
F(x)
\end{bmatrix}
\text{Hidden layer transfer function.}
\]

\[
\begin{bmatrix}
11 & 1 & 11 \\
10 & 0 & 01 \\
10 & 0 & 01
\end{bmatrix}
\begin{bmatrix}
6 \\
2 \\
2
\end{bmatrix}
\text{Connection weights between 2 hidden and 3 output nodes.}
\]

\[
\begin{bmatrix}
2 \\
2 \\
2
\end{bmatrix}
\text{Output nodes.}
\]

\[
\begin{bmatrix}
G(x)
\end{bmatrix}
\text{Output layer transfer function.}
\]

\[
\begin{bmatrix}
8 \\
8 \\
8
\end{bmatrix}
\text{Output.}
\]

In the above sequence, matrix multiplication accounts for the combination of data and the transfer functions for the transformation of data. For the sake of completeness, an input layer transfer function should be added, as well as a fixed bias. The bias is produced by an additional input node that is connected with weighted connections to all the nodes in all the layers. The output from this node is always 1. The bias acts as a threshold value above which a particular node becomes activated.
A network with three layers performs two sequential mappings of the input data. The input pattern \([3 \ 1 \ 0 \ 1]\) is first mapped onto the hidden layer pattern \([2 \ 8]\) which is then mapped onto the output pattern \([8 \ 8 \ 8]\). The mapping at the hidden layer forms an internal representation of the input pattern (Rumelhart et al., 1988b). Multiple hidden layers enable a network to perform more complex mappings between input and output, as do some nonlinear transfer functions (Lee, 1992).

A neural network has the ability to map any arbitrary set of vectors onto any other arbitrary set of vectors. What is interesting, however, is that in the process of mapping a subset of the domain (the training data set) onto a subset of the range (the desired network outputs), a well-trained network manages to map the entire domain onto the entire range. In this case, new data will generate the correct output when presented to the network and the network is said to be capable of generalizing from the training data set.

The process is rather analogous to polynomial curve (or surface) fitting (Baldi and Hornik, 1995). In general, when given a number of points, \(n\), a polynomial of order \(n - 1\) can be found to fit exactly through all the points. Such a polynomial succeeds in mapping the input data (a subset of the total input data set) onto the required output data (a subset of the total output data set), but often not in mapping the total input data set onto the total output data set, as is desirable. Often, though, when the order of the polynomial is reduced, such a mapping is achieved. This happens because the number of parameters in the polynomial equation that can be adjusted (free parameters) to optimize the fit, are reduced. Hence the polynomial is no longer fitted to the data points themselves, but to those aspects of the data that are held in common, namely, their trends. Analogously, a reduction in the number of free network parameters, that can produce overfitting, can be attained by reducing the total number of connections between nodes: reducing the number of hidden layers, nodes per hidden layer, or connections between any of the nodes. In practice, however, not too much is known about overfitting and its onset in terms of network parameters and training time (Baldi and Hornik, 1995).
11.3.2.3. Node computation

Every node in a network is a computational unit (Rumelhart et al., 1988a) and some data processing occurs at every node. A hidden layer node (e.g. hidden layer node 1), as shown in Figure 11.2, performs the following computations: the data transferred from every input node are combined (usually by addition) into a single value. The output from input node 1 (the value 3), multiplied by the connection strength between input node 1 and hidden node 1 (the value 0), is added to the outputs from input nodes 2, 3, and 4 multiplied by their respective connection weights ([3 x 0] + [1 x 1] + [0 x 1] + [1 x 0] = [1]). This value (1) is passed through the transfer function $F(x) = 2x$ to determine the hidden node's output $F(1) = 2 \times 1 = 2$. The product of the hidden node output and the connection strength between the hidden node and a node in the next layer is then passed on to that next node. At this next node, all the inputs are again combined and the process repeated.

In essence then, a node simply performs a computation on the combined inputs from the previous layer, and sends the result to every node of the next layer (Rumelhart et al., 1988a). Although the neural network structure described here allows all the nodes in a particular layer to perform their computations in parallel, hence parallel distributed processing (Rumelhart et al., 1988a), such computations are serially implemented on computers, except those with specialized hardware.
Figure 11.2. This figure shows (from left to right) the computations performed at hidden layer node 1. The data transferred from each of the 4 input nodes, multiplied by the respective connection weights (in small rectangles), are combined into a single value at this hidden node. This value is passed through the transfer function $F(x)$ to determine the hidden node's output. As before, this output is multiplied with the connection weights (in small rectangles) and transferred to the 3 output nodes.

The connection weights are adjusted, based on the output error, as follows (indices omitted):

$$\Delta w = \eta \delta o,$$

where $\Delta w$ is the amount by which to adjust the weight $w$. The term $\delta$ is proportional to the output error $(d - o)$ where $d$ is the desired output, and $o$ is the generated output of the node. The output error can be scaled by the learning rate constant $\eta$. Since learning is proportional to the size of the output error, scaling modifies the rate of learning. The term $\delta$ is given by:

$$\delta = (d - o)g'(\Sigma i)$$

(11.11),
where $G'$ is the derivative of the transfer function employed by the node with respect to the total input to that node, and $\Sigma i$ is the sum of all the inputs $i$ to the node. Due to the use of a derivative, weights can be adjusted in a systematic manner (that amounts to performing a steepest descent on a surface in weight space) as opposed to changing the weights in a random manner until a good set of weights are found.

For a node other than an output node, the term $\delta$ is modified to incorporate the effect of the previous layer. Thus for a hidden layer node:

$$\delta = f'(\Sigma j)\Sigma[(d - o)G'(\Sigma i)w]$$

where $\Sigma j$ is the sum of the inputs to that node and $f'$ is the derivative of the transfer function employed by the node with respect to the total input to that node. A derivation of the generalized delta learning rule can be found in Rumelhart et al. (1988b).

11.3.2.4. Weight optimization

Finding the optimum weights for the connections of a neural network through training constitutes an optimization problem in multidimensions, the number of dimensions being the number of connection weights to be optimized. Hence, aside from the various backpropagation methods, a whole array of multidimensional optimization methods can be brought to bear on the problem of optimal network training. These optimization methods represent different training methods.

For instance, small networks may be amenable to optimization by the simplex method of Nelder and Mead (1965), albeit a rather slow method. Faster convergence can be obtained with the use of second-order gradient descent methods and they have been applied to neural networks (Hagan and Menhaj, 1994; see also Baldi, 1995). These gradient descent methods include the accurate but computationally expensive Kalman filtering and its variants (e.g. Blank and Brown, 1994). Simulated annealing methods (Berthiau et al., 1994; Bilgen and Hung, 1994), which are often slow but more likely to converge to the desired global optimum, have also found neural network training
applications. Another method applied to weight optimization is stochastic minimization which does not require gradient information and can be applied to networks with non-continuous transfer functions (Brunelli, 1994)

11.3.2.5. Biological neural networks

The computational units of biological neural networks are nerve cells or neurons and were described in Chapter 1 and shown in Figure 1.1. The summation of postsynaptic potentials caused by signals from many different input neurons is analogous to the summation performed by an artificial neural network node on its inputs while the cellular threshold is represented by the artificial neural network bias. Connections between neurons are often established and then selectively eliminated during development (Ivy and Killackey, 1982). Artificial neural networks too can be pruned by eliminating some (mostly weak) connections during training (e.g. Borggaard and Thodberg, 1992). In maintained connections, the connection strengths may correspond to the amount of neurotransmitter released in the synaptic cleft. The gated ion channels have non-linear characteristics that contribute to cellular computations (Koch et al., 1983; Bialek et al., 1991) and possibly correspond to the transfer functions. It is much less clear how to draw an analogy between the backpropagation of the network output error and neural function. It has been found, however, that stimulated dopaminergic neurons of the substantia nigra secrete dopamine dendritically to affect afferent neurons (Cheramy et al., 1981), thus potentially 'backpropagating' signals.
Neurons are often highly organized into structures in many brain areas (see Chapter 1). The layered architectures of these structures invite a comparison with artificial neural networks which are in a general sense modeled after them. For instance, the organization of cells in the primate retina, shown in Figure 11.3, indicates a 3-layered structure similar to the backpropagation artificial neural network discussed in the previous section. The input layer consists of the rod and cone receptor cells, the hidden layer consists of the bipolar cells, and the output layer of the ganglion cells. The axons of the ganglion cells form the optic nerve which transmits visual information via the thalamus to the visual cortex.

Figure 11.3. A simplified rendition of the primate retina showing a 3-layered structure similar to the backpropagation artificial neural network discussed in the text. The input layer consists of the rod receptor cells, the hidden layer consists of the bipolar cells, and the output layer of the ganglion cells. The axons of the ganglion cells form the optic nerve.
An examination of Figure 11.3 reveals two novel features. First, nodes in the input and output layers of this natural neural network are often connected to other nodes in the same layer via horizontal and amacrine cells, respectively. Second, not all the nodes in one layer are connected to all the nodes in the next layer. The receptors are, in general, connected only to a few neighboring bipolar cells. Artificial neural networks with such local connections, which process only a local region of the input, often generalize better than networks with exhaustive connections (see Hush and Horne, 1993; Schulze et al., 1995). In this respect, a network with local connections represents a cost advantage: fewer nodes are required and need to be maintained in order to adequately process a wide range of information. Perhaps local connections also occur in biological networks as a consequence of these advantages.

Although the retinal circuitry is not a passive relay system, but a sophisticated signal processing system (see Celesia, 1988) that recognizes certain patterns (e.g. on-center, off-surround), the extent to which pattern recognition at the retinal level is dependent on 'training' is unknown.

11.3.2.6. Modifications

Network modifications can have profound effects on the network's operation. The network architecture can be modified, for example, by adding one or more hidden layers to generate a network with multiple hidden layers and an increased ability to perform complex mappings between input and output patterns (see Lippmann, 1987). The addition of direct connections between input and output nodes often improves the flexibility of the network and increases its training speed (Borggaard and Thodberg, 1992). Local connections can be employed between any number of interconnected layers. Networks with partial (but not necessarily local) connections also show an increase in generalization, more so when the connections are fixed before training (quenched dilution) than when pruned during training (annealed dilution) (Campbell, 1992). Given
all these possibilities, network optimization is a daunting goal. The minimum net size (Kamruzzaman et al., 1992) and the optimum network architecture (Kendall and Hall, 1992) to perform certain required tasks have been thoroughly investigated, and work on these issues remains an active area of research.

Analyses of biological networks can also contribute to the design of new network structures. For instance, the structure of the cerebellum has inspired the development of associative-memory models (see Kanerva, 1992) and tectal maps the development of models of sensory-motor control (Spence et al., 1989). An analysis of the visual cortex is also informative. In the cat visual cortex the input from the lateral geniculate nucleus of the thalamus is relayed to layer 4 of the cortex with a subsidiary input to layer 6. Layer 6 also provides an output from the primary visual cortex (Kelly, 1985b). This arrangement resembles the direct connection network mentioned above. Layer 6 furthermore returns a connection to layer 4 (e.g. Bolz et al., 1989), thus recirculating information to the input layer, an organization similar to dynamic artificial networks with output feedback (see Hush and Horne, 1993). In addition, layers other than layer 6 act as output layers. Layer 5, for instance, provides an output to the superior colliculus (Kelly, 1985b). Given the fact that information contained at hidden nodes are internal representations of the input information (Rumelhart et al., 1988b) and constitutes a compression of the input data when there are fewer hidden than input nodes, such output may serve as useful information to other systems.

Localized learning can also be attained with the use of Gaussian transfer functions to improve network training speed and adaptability (Bakshi and Stephanopoulos, 1993). In general, various transfer functions could be used in the different layers. Some transfer functions allow a network to perform more complex mappings between input and output (Lee, 1992). Sigmoidal hidden layer transfer functions allow some networks to generalize better (Schulze et al., 1994), while sine hidden layer transfer functions improve training speed (Zhang and Wang, 1992) and recognition accuracy (Schulze et
al., 1994). It is also possible to split a layer into sections using different transfer functions to improve the network performance (Schulze et al., 1995). This approach could be extended to every node having its own transfer function or an extreme where every node would have several transfer functions, one for every output connection.

The operations performed by the individual nodes can be manipulated. The inputs to a node are combined, usually by addition (Rumelhart et al., 1988a), but sometimes also by multiplication (e.g. Fahner et al., 1992), Boolean logic (see Martland, 1989), probabilistic logic based on random access memory (Kan and Aleksandr, 1989), and fuzzy logic (e.g. d'Alche-Buc et al., 1992). Each of these networks is likely to display its own characteristic "behavior".

The choice of training data sets can also have a strong influence on the network performance. In order for a network to generalize well from the training data set, the training data set has to be sufficiently large relative to the network size (Baum and Haussler, 1989) and appropriately chosen (Ogawa and Yamasaki, 1992) so as to representatively sample the entire domain.

Finally, the choice of output coding can be used to help optimize network performance. This optimum is attained when the number of output classes form a set of equidistant points. Therefore, encoding the output to minimize the number of output nodes and hence reduce the number of weights would be counterproductive (Thomas and Mitchie, 1994).

11.3.3. Water as an internal standard

Measuring analytes in aqueous solution provides the opportunity for using water as an internal standard. For instance, Rava and Spiro (1985) reported a tenfold enhancement in the benzene and pyrrole out-of-phase ring breathing band at 1016 cm⁻¹ relative to the water band near 3650 cm⁻¹ when exciting at 218 nm instead of 240 nm. In addition to the water band near 3650 cm⁻¹ (symmetrical stretch), another exists near
3750 cm\(^{-1}\) (anti-symmetrical stretch) and a third near 1600 cm\(^{-1}\) (deformation or scissors-vibration) that could be used for quantification purposes (e.g. see Figures 4.1-4.11). The availability of 3 water bands offers some leeway in the event that one of the bands be obscured by a vibrational band(s) of one of the analytes. For instance, the water band near 1600 cm\(^{-1}\) can be obscured by a ring stretch vibration as shown in Figures 4.3-4.5. Calibration curves can be determined for each neurotransmitter to aid in quantification, giving the ratios of the most intense vibrational band(s) relative to those of water.

11.4. REFERENCES


12.1. INTRODUCTION

In Chapter 11, the 3-step approach to recover the perfect underlying analyte spectrum from the measured data modeled by:

\[ D_i = B^*(f_i + b_i) + \sigma_i \]  

was described as consisting of noise removal, deconvolution, and baseline removal. Methods to accomplish this were briefly introduced. The objectives of this chapter were to apply some of these methods and their extensions to the enhancement of experimentally measured Raman signals. These methods were linear filtering and autoaccumulation for noise reduction, maximum entropy processing for noise reduction and deconvolution, and curve fitting for baseline correction. Applications involving neural networks are discussed in subsequent chapters.

12.2. NOISE REMOVAL

The first step in recovering the signal was noise removal. The noise present in the measured Raman spectra was found to be Gaussian distributed noise, hence relevant noise reduction methods were employed.
12.2.1. Filtering.

The linear moving average filter was often used to smooth Raman spectra. Because the size of the window could be manipulated, the amount of smoothing could be controlled. Moving average smoothing, however, resulted in the loss of fine spectral detail, reduced and broadened peaks, increased baselines, and unsmoothed endpoints. Figure 12.1 demonstrates this with a section of raw dopamine Raman spectrum processed with a 10-point and 50-point moving average filter (zero-order Savitzky-Golay filter). Filtering is mostly of cosmetic advantage, except when the signal and noise are significantly different in a known way and the noise can be selectively suppressed (Demas and Demas, 1990). Due to the signal distortions introduced by filtering, least squares data fitting should not be applied to filtered data (Demas and Demas, 1990).

Figure 12.1. The effects of filtering with a zero-order Savitzky-Golay filter with a 10 (dotted line) and 50 (dashed line) point spread on a section of a dopamine Raman spectrum.
12.2.2. Autoaccumulation.

As described in Chapter 11, signal averaging can be used to increase the signal-to-noise ratio (SNR) of a spectrum in proportion to the square root of the number of accumulations (Demas and Demas, 1990). To prevent signal distortions, this requires that all the signal copies must be aligned (Yaroslavsky and Eden, 1994) and have no overlap (Woldorff, 1993). Often though, accumulation is not possible due to the uniqueness of an event or measurement. Consequently, single trial events have been analyzed by several authors: to extract very low SNR events using an outlier processing method with time-varying influence functions (Mason et al., 1994); to optimize averaging by statistical verification based on single trial analysis (Papakostopoulos and Gogolitsyn, 1994); and using an autoregressive algorithm with exogenous inputs to improve the SNR of brain potentials (Chiarenza, 1994).

In the case of a unique event, however, the accumulation of several copies of the same spectrum, each slightly misaligned with respect to the other, may be advantageous. Consider the following signal $b_k(x)$, which is a copy of the signal $s(x)$ displaced by an unknown arbitrary wavenumber shift $x_k$:

$$ b_k(x) = s(x-x_k) + n_k, \quad k = 1, 2, ..., k. $$

where $n_k$ is additive Gaussian noise. The best estimate of the shift $x_k$ is provided by the maximum of the cross-correlation between $b_k(x)$ and $s(x)$ according to the method of Yaroslavsky and Eden (1994). If the shift is within the signal cross correlation interval, some degree of coherent accumulation will occur (see Yaroslavsky and Eden, 1994). By extension, if the displacements are generated to be within the signal, but outside of the noise autocorrelation intervals, some coherent accumulation of the signal but not the noise, will occur using the same spectrum for displacement and accumulation.

The acetylcholine spectrum shown in Figure 4.2 was displaced randomly in either direction by a preselected amount (bandpass) according to a uniform or normal distribution (bandwidth equaled 6 standard deviations) and 1000 such displaced copies
averaged in order to improve the SNR. The noise characteristics were obtained from a region of the spectrum between 1900 cm\(^{-1}\) and 1975 cm\(^{-1}\). The results are shown in Table 12.1.

### Table 12.1.

The characteristics of signal, noise, and approximate SNR from an acetylcholine spectrum autoaccumulated with varying bandwidths and using uniform or normal distributions. Abbreviations: N = normal distribution; U = uniform distribution; n = number of samples; sd = standard deviation.

<table>
<thead>
<tr>
<th></th>
<th>Ach</th>
<th>50N</th>
<th>50U</th>
<th>30N</th>
<th>30U</th>
<th>10N</th>
<th>10U</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>75</td>
</tr>
<tr>
<td>mean</td>
<td>0.35</td>
<td>0.35</td>
<td>0.35</td>
<td>0.35</td>
<td>0.35</td>
<td>0.35</td>
<td>0.35</td>
</tr>
<tr>
<td>sd</td>
<td>0.028</td>
<td>0.008</td>
<td>0.004</td>
<td>0.011</td>
<td>0.006</td>
<td>0.018</td>
<td>0.013</td>
</tr>
<tr>
<td>peak</td>
<td>7.65</td>
<td>5.30</td>
<td>3.17</td>
<td>6.24</td>
<td>4.04</td>
<td>7.35</td>
<td>6.51</td>
</tr>
<tr>
<td>base</td>
<td>1.56</td>
<td>1.58</td>
<td>1.70</td>
<td>1.54</td>
<td>1.60</td>
<td>1.52</td>
<td>1.54</td>
</tr>
<tr>
<td>SNR</td>
<td>218</td>
<td>465</td>
<td>367</td>
<td>427</td>
<td>407</td>
<td>324</td>
<td>382</td>
</tr>
</tbody>
</table>

It is evident from Table 12.1 that all 6 manipulations produced a SNR improvement. However, the greatest enhancement was provided by a series of 50-point normally displaced accumulations. The uniformly displaced accumulations produced the greatest noise reduction, but at the cost of much broadened and reduced peaks. The results of autoaccumulation using a 50-point bandwidth with both normally and uniformly distributed displacements are shown for comparison in Figure 12.2 with the original spectrum.

Due to the symmetrical nature of the displacements, no shift in peak position was generally observed. In order to compare these results to an existing method, the same spectrum was filtered with a zero-order Savitzky-Golay filter. The results are shown in Table 12.2.
Figure 12.2. The effects of 50-point normally (dotted line) and uniformly (dashed line) displaced autoaccumulations on an acetylcholine spectrum.

Table 12.2. The characteristics of signal, noise, and approximate SNR from an acetylcholine spectrum processed with autoaccumulation and varying bandwidths of a zero-order Savitzky-Golay filter. Abbreviations: SG = Savitzky-Golay filter; others as in Table 12.1.

<table>
<thead>
<tr>
<th></th>
<th>Ach</th>
<th>50N</th>
<th>50SG</th>
<th>30SG</th>
<th>10SG</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>75</td>
</tr>
<tr>
<td>mean</td>
<td>0.35</td>
<td>0.35</td>
<td>0.35</td>
<td>0.35</td>
<td>0.35</td>
</tr>
<tr>
<td>sd</td>
<td>0.028</td>
<td>0.008</td>
<td>0.004</td>
<td>0.006</td>
<td>0.013</td>
</tr>
<tr>
<td>peak</td>
<td>7.65</td>
<td>5.30</td>
<td>3.13</td>
<td>4.08</td>
<td>6.60</td>
</tr>
<tr>
<td>base</td>
<td>1.56</td>
<td>1.58</td>
<td>1.70</td>
<td>1.62</td>
<td>1.55</td>
</tr>
<tr>
<td>SNR</td>
<td>218</td>
<td>465</td>
<td>357</td>
<td>409</td>
<td>388</td>
</tr>
</tbody>
</table>

A comparison of Tables 12.1 and 12.2 reveals that the Savitzky-Golay filter gave a performance very similar to that of uniformly displaced accumulations at low bandwidth. However, the uniformly distributed accumulations method performed better than the Savitzky-Golay filter with increasing bandwidth. Furthermore, both methods showed an optimum near a bandwidth of 30 cm\(^{-1}\). This probably reflects the fact that the autocorrelation of the noise was non-significant after a lag of 1 cm\(^{-1}\), while the
autocorrelation of the signal only became non-significant after a lag of 9 cm\(^{-1}\). Thus two signals with means displaced by 18 cm\(^{-1}\) could still be coherently accumulated (which corresponds to a bandwidth of 36 cm\(^{-1}\)).

Theoretically, 1000 accumulations of aligned spectra should yield a SNR improvement by a factor of 30. Only a factor of 2 improvement was obtained. Although the noise level was reduced by a factor of 3.5, a concomitant SNR improvement was not realized because the 'spreading' effect of the displaced accumulations caused peak broadening and peak height reduction. It seems reasonable that deconvolution would lead to a further SNR improvement, although an attempt to deconvolute the autoaccumulated spectra using Fourier transforms was unsuccessful. This approach, however, will continue to receive attention.

12.2.3. Moving product processing

Instead of a moving average, it ought to be possible to use a moving product to perform smoothing. In particular, if the baseline of a spectrum is adjusted such that its value is less than or equal to 1, the value of Raman signals should be more than 1. Therefore, if, in a manner analogous to that of the linear moving average, the center value in a moving window is replaced by the product of the values in the window, signals would in general be amplified while the baseline (and noise) would be suppressed. This procedure has been applied to a Gaussian signal with a baseline level of 1 arbitrary unit, mean at 75 cm\(^{-1}\), standard deviation of 14 cm\(^{-1}\), and maximum peak height of 1 arbitrary unit above the baseline. To this random Gaussian noise was added to give a SNR of 1. The raw and processed spectra are shown in Figure 12.3.
Figure 12.3. The effects of using a moving product to smooth a spectrum with SNR of 1. A Gaussian signal after Gaussian noise addition (top). The effects of smoothing with a 6 (dotted line), 20 (dashed line) and 25-point (solid line) moving product on the noisy spectrum. Original signal (before noise addition) shown at the bottom. Spectra were normalized (using factors of 4, 300, and 10,000 respectively) and offset for comparison.

An inspection of Figure 12.3 reveals that the amount of amplification depended on the size of the moving product window. A window size comparable to twice the standard deviation of the signal (25 cm$^{-1}$ versus 28 cm$^{-1}$) produced an enormously amplified signal while suppressing the baseline and noise considerably. A SNR of nearly 200 was obtained using the last 86 data points of Figure 12.3 to determine the noise standard deviation (signal height 10,220 and noise standard deviation 52). The same spectrum produced a SNR of 3 when processed with the autoaccumulation method (1000 accumulations, normally displaced, 50 cm$^{-1}$ bandwidth). These results suggested that moving product data processing has the potential to suppress noise and amplify signals in low SNR spectra. It is not clear to what extent the peak shape gets distorted by this
method. Further investigations, and extensions analogous to Savitzky-Golay filtering, are required and may prove worthwhile.

12.3. DECONVOLUTION.

Deconvolution was the second step in recovering the signal. Some methods such as maximum entropy processing can combine noise removal and deconvolution.

12.3.1. Maximum entropy.

The maximum entropy method discussed in Chapter 11 minimizes the function $M$:

$$M = -S + \lambda \chi^2 + \mu C$$  \hspace{1cm} (11.8)

The entire spectrum is considered to be a probability distribution with entropy $S$. However, $S$ is invariant under a complete reordering of the data points which contradicts the view of a spectrum as an ordered sequence. Replacing the entropy in Equation 11.8 with the two-point entropy (see Equation 11.7):

$$-S_2 = \sum (p_i \ln(p_i) + p_j \ln(p_j))$$  \hspace{1cm} (12.2)

immediately constrains the entropy function since it is now highly dependent on the ordering of the data points. This obviates the need for the constrained constraining function $C$ in Equation 11.8 and the equation reduces to two functions and one Lagrange multiplier:

$$M = -S_2 + \lambda \chi^2$$  \hspace{1cm} (12.3)

Minimizing $M$ is now computationally easier. This two-point maximum entropy method requires estimates of the noise statistics and the blurring function, but no other information. A complete derivation and application can be found in Greek et al., (1995).
12.4. BASELINE REMOVAL.

Strictly speaking, baseline removal is the third step after noise removal and deblurring. This was not always done, and baseline removal was often the first or second step in spectral processing. Where polynomial curve fitting was used to remove the baseline, the raw spectra were used.

12.5.1. Curve fitting.

Raman spectral baselines are often gently curving due to broadband fluorescence and/or gently sloping due to the Rayleigh wing. For these reasons, segmental linear baseline-fitting introduces artifacts in the corrected baseline. The Fourier transform of the raw spectrum could also be obtained and the low frequency components deleted. This would reduce the baseline, but often introduces undulations in the remaining baseline due to the cyclic nature of the Fourier components.

Polynomial baseline fitting was found to be more effective for baseline removal. Initially, 8 to 10 approximately evenly spaced points covering the whole spectrum were selected and a third order polynomial fitted to these points (e.g. Schulze et al., 1994). This gave an improved fit, but the points still had to be selected from a great number of possibilities, thus discarding much information about the baseline. Ideally, one would wish to fit the polynomial to the raw data without interference by the Raman signals. Shreve et al. (1992) used a polynomial method to remove the baseline from a shifted excitation Raman difference spectrum. The peaks in this difference spectrum were, however, symmetrical above and below the baseline due to the shifted excitation and a polynomial could be fitted without undue interference by the signal.

One possible approach to this problem is to fit a polynomial to the raw spectrum and replace parts of the spectrum in excess of 3 times the standard deviation of the noise with the value of the polynomial at that point. This procedure is then iterated until no points are being replaced. The polynomial so obtained is then used as an estimate of the
baseline and subtracted from the original spectrum. This procedure ensures that as many as possible of the baseline points are used in its estimation, thus producing an improved estimate. Figure 12.4 shows two iterations of the procedure with the original and corrected spectra.

![Figure 12.4](image)

Figure 12.4. First stage in the baseline correction of a dopamine spectrum showing a polynomial fit (dashed line) and the removed outliers (Raman signals, dotted lines). The baseline-corrected spectrum is shown at the bottom.

The spectrum used was that of dopamine (Figure 4.3) and the standard deviation of the mean was calculated from a part of the spectrum between 610 and 700 cm$^{-1}$ with the SYSTAT statistical program. The number of points used was 91 and the standard deviation calculated as 0.136. For curve fitting, a singular value decomposition method implemented in C by M. D. Lacasse, downloaded from the internet, was used.

It should be noted that the spectral fit improved little with increasing polynomial order. It is reasonable, however, to assume a better fit with increasing order. The
question remains as to what order polynomial to use to remove the baseline but not the entire spectrum. For these reasons, a seventh-order polynomial was seldom exceeded. It would be interesting to develop a means of quantification of baseline correction efficiency. The effect of polynomial order on baseline fitting is shown in Figure 12.5.

![Figure 12.5. The effects of polynomial order on baseline fitting, displaced vertically for ease of viewing: 3rd order (top), 5th order (middle), and 6th order (bottom).](image)

**Figure 12.5.** The effects of polynomial order on baseline fitting, displaced vertically for ease of viewing: 3rd order (top), 5th order (middle), and 6th order (bottom).

12.5. **Discussion.**

According to the model formulated by Equation 11.1, baseline removal should be the last step in the recovery of the signal. However, due to the least squares polynomial curve-fitting procedure used for baseline correction, the raw spectra had to be used to avoid applying the curve-fitting procedures to spectra distorted by noise-removal (Demas and Demas, 1990). This conflict could be resolved by employing noise-removal techniques that do not distort the signal (if these exist).
However, Equation 11.1 could be restated as:

\[ D_i = B^*f_i + B^*b_i + \sigma_i \]  

(12.4)

It is clear from Equation 12.4 that either the noise or the convolved background can now be removed as a first step towards signal recovery. This makes it possible to apply the polynomial curve-fitting method to the raw data to remove the convolved background followed by noise removal. This procedure (3rd order polynomial background correction followed by 1000 normally displaced accumulations with 50 cm\(^{-1}\) bandwidth) was used to recover the signal from an acetylcholine spectrum (0.01 M in physiological saline). Figure 12.6 shows the noticeable improvement in the signal-to-noise ratio.

![Figure 12.6. The recovered signal (solid line) from a 0.01 M acetylcholine in physiological saline Raman spectrum (dotted line) after a 3rd order polynomial background removal and 1000 autoaccumulations.](image)

The SNR of Raman signals can be expressed as:

\[ \text{SNR} = \frac{n_R}{\sqrt{\left[ (n_R + n_D + n_F + n_{SC}) + \xi^2(n_R + n_F + n_{SC})^2 \right]^{0.5}}} \]

(12.5)
where \( n_R \) is the number of Raman counts, \( n_{SC} \) the number of counts due to elastic scattering, \( n_F \) the number of counts from fluorescence, \( n_D \) the number of dark current counts, and \( \xi \) the source flicker factor (Ingle and Crouch, 1988). It is evident from Equation 12.5 that the SNR can be improved by increasing \( n_R \) relative to \( n_D, n_F, \) and \( n_{SC} \). Dark current reduction can be achieved with detector cooling (see Chapter 3), Rayleigh scattering reduction with a filter, and both fluorescence rejection and Raman signal enhancement with resonant excitation.

However, due to the inherently weak nature of Raman signals, digital methods of signal recovery with SNR enhancement will remain important. Some methods of realizing this have been applied and discussed in this chapter. Of these, the two-point maximum entropy method seemed the most promising. It would be useful to find an estimate of the maximum SNR obtainable from information theory, given certain spectral parameters, to determine to what degree the entropy method could be improved upon. The objective should be to obtain the maximum SNR possible from spectra of short integration times to achieve the aims of real-time neurotransmitter measurement in vivo.

12.6. REFERENCES.


CHAPTER 13

13.1. INTRODUCTION

Artificial neural networks (ANNs) are useful methods for the investigation of various spectra, e.g. electrical brain signals (Gevins and Morgan, 1988), colorimetric spectra (Lin et al., 1992), and nuclear magnetic resonance spectra (Anker and Jurs, 1992). Tanabe et al. (1992) reported that a neural network can be used on a personal computer to rapidly identify infrared spectra. Given the success of ANNs as pattern classifiers, they have here for the first time been applied to the identification of the 10 small-molecule neurotransmitter Raman spectra. The objectives of the work reported in this chapter were to design and train an ANN for real time identification and quantification of Raman and, by extension, resonance Raman signals obtained via a fiber-optic probe, and to investigate the abilities of ANNs to identify Raman spectra with eventual in vivo applications in mind.

13.2. SIGNAL CLASSIFICATION WITH NEURAL NETWORKS

It is known that the size of the intermediate or hidden layer (Rumelhart et al., 1988; Tanabe et al., 1992; Liu et al., 1993), the number of hidden layers (Hush and Horne, 1993; Liu et al., 1993), and the transfer functions used for the layers (Zhang and Wang, 1992) can affect the training and execution time of the network. The effects of the ratio of the number of hidden nodes to input nodes, the transfer functions for the layers,
different input features, as well as the number of input layer nodes, on the network performance were investigated as a first step in the optimization of the network. 

The Raman spectra of the small-molecule neurotransmitters, consisting of the region 2000 cm$^{-1}$ to 500 cm$^{-1}$ from the excitation frequency, served as data sets for network training and testing. This range is considered to be quite effective for the identification of infrared (Tanabe et al., 1992) and Raman spectra. These neurotransmitters were dissolved in physiological saline (0.9% NaCl) to concentrations between 0.1 M and 0.5 M. Raman scattering was excited with the 488 nm line from an Ar$^+$-laser operating at 200 mW and measured with a scanning spectrometer (JASCO, Tokyo, Japan) at a rate of 120 cm$^{-1}$/min. The spectra were digitized at 1 cm$^{-1}$ intervals resulting in 1501 points per spectrum with approximate signal-to-noise ratios (maximum peak height/noise standard deviation) between 20 and 50.

Spectra were smoothed with a 5-point moving average and baselines were corrected with a (3rd order) polynomial method (see Chapter 12). Before being used as training data sets, the peak heights of these spectra were also normalized to the value of the highest peak. Smoothing a spectrum with a moving average is a form of noise filtering. Using smoothed spectra therefore allowed the investigation of the effects of noise and other spectral distortions.

Fully-connected, three-layer backpropagation artificial neural networks (NeuralWare Incorporated, Pittsburgh, Pennsylvania, U.S.A.) with 10 output nodes, one corresponding to each of the 10 small-molecule neurotransmitters, were used. The performance of the networks was determined by the root mean square (RMS) error of the output vector:

$$\text{RMS} = \left( \Sigma (O_i - D_i)^2 / N \right)^{1/2}$$

where $O_i$ is the actual output at node $i$, $D_i$ is the desired output at node $i$, and $N$ the number of output nodes.
The networks were trained with a data set that consisted of the 10 neurotransmitter spectra, sampled from the spectra at uniform intervals of 10 cm\(^{-1}\), or, in the case of investigating the effects of input features on the network performance, of the means of 10 cm\(^{-1}\) intervals. Every network was trained with 5 different weight initializations and all networks were trained to the same criterion for comparison. Two typical learning curves are shown in Figure 13.1. Connection weights were randomly initialized between -0.1 and +0.1, or 0.0 and +0.1 for a sine transfer function before each training session. The momentum term and learning rate coefficients were both set at 0.2. Except where otherwise mentioned, sigmoid transfer functions were used.

![Typical learning curves for neural networks using sine transfer functions (solid line) and sigmoid transfer functions (dotted line).](image)

The same neurotransmitter spectra, but differently manipulated, were used for testing. Five sets of testing data were used. The first testing set (SM5) consisted of the
training data set, that is, spectra with corrected baselines and smoothed with a 5-point moving average. The second testing set (SM10) consisted of the same spectra, but smoothed with a 10-point moving average. The third testing set (SH5) contained spectral data that were shifted towards the exciting frequency by 5 cm⁻¹, and the fourth testing set (SH10) consisted of spectral data that were shifted in the same sense, but by 10 cm⁻¹. The final testing set (RAW) contained data that were neither smoothed nor shifted, but that had their baselines corrected.

13.2.1. Layer sizes.

As can be seen from Figure 13.2, network accuracy improved with an increase in the number of input nodes for the smoothed and raw, but not the shifted data sets. This suggested that networks became more sensitive to spectral shifts when more input nodes were used. The networks tested had the same ratio of input to hidden nodes, sigmoid transfer functions, and spectral data were uniformly sampled. The networks had the following configurations (number of input nodes-number of hidden nodes-number of output nodes): 136-22-10; 150-24-10; 250-40-10; and reached the same criterion after (mean ± S.E.M.) 5,454 ± 88; 5,096 ± 96; and 3,663 ± 59 sequential data presentations, respectively. Training presentations required decreased as the number of input nodes increased.
Figure 13.2. Performance of neural networks with the same ratio of hidden to input layer nodes but different numbers of input layer nodes (IN136: 136 input nodes, IN150: 150 input nodes, IN250: 250 input nodes). Testing sets consisted of the baseline-corrected Raman spectra of neurotransmitters manipulated as follows: smoothed with a 10-point moving average (SM10); smoothed with a 5-point moving average and also used for training (SM5); not processed (RAW); shifted towards the exciting frequency by 5 cm⁻¹ (SH5); and shifted towards the exciting frequency by 10 cm⁻¹ (SH10).

The network with the lowest ratio of hidden to input nodes performed the least well on most of the data sets. In agreement with the report by Tanabe et al. (1992), networks with a ratio of 0.16 were found to perform well, as shown in Figure 13.3. However, the network with the 0.16 ratio performed less well on the data set smoothed with a 10-point moving average and the set shifted by 10 cm⁻¹. This result implied that a network using a 0.16 ratio became more sensitive to the loss of detail that occurred with a 10-point smoothing.
Figure 13.3. Performance of neural networks with the same number of input nodes but different ratios of hidden to input nodes: ratio = 0.10 (RATIO10); ratio = 0.16 (RATIO16); and ratio = 0.25 (RATIO25). Testing sets as for Figure 13.2.

13.2.2. Transfer functions.

Different transfer functions were tested with a 150-24-10 network configuration and uniformly sampled data. Sine-sigmoid networks (SNSG), which use a sine and a sigmoid transfer function for the hidden and output layers, respectively, have been reported as showing fast learning (Zhang and Wang, 1992) and with the spectral data used here trained 3-5 times faster than the equivalent sigmoid-sigmoid networks (SGSG). For example, the SNSG and SGSG networks reached the training criterion after 4,272 ± 64 and 16,144 ± 211 trials, respectively. It is worthwhile noting that the SNSG network, while recognizing the spectral data of the training set as well as the SGSG network, performed less well on all the other data sets except where the spectra were shifted by 10 cm⁻¹. This result indicates that the SNSG network is generally more discriminating than
the SGSG network, or, alternatively, the SGSG network is more robust to some spectral
distortions. The concept of 'overtraining' has been used to describe the tendencies of
networks to lose their ability to generalize from training to testing data sets as they
become more highly trained (e.g. Anker and Jurs, 1992; Hush and Horne, 1993). The
present results indicate that networks with different transfer functions 'overtrain' to
different degrees even when they are trained to the same objective root mean square
(RMS) error criterion. Consequently, the phenomenon of overtraining could perhaps be
reinterpreted as the extent to which similarly trained networks possess the abilities of
discrimination and generalization.

Discriminating networks are here defined as networks being capable of generating
RMS errors much larger than the training criterion in response to small variations in
spectral data. Such networks could be useful to discriminate between highly similar
spectra. Generalizing networks, on the other hand, produce RMS errors not appreciably
larger than the training criterion when correctly identifying spectral data containing mild
to moderate distortions. The latter networks could be useful when dealing with data
subject to calibration errors or other instrumental distortions.

13.2.3. Generalization and discrimination.

In order to investigate the different discriminating and generalization abilities of
networks with different transfer functions, and to provide a basis for comparison, the
RMS error was normalized by defining a discrimination index:

\[
DI = \frac{([\text{RMS error of testing data}]-[\text{RMS error of training data}])}{[\text{RMS error of training
data}]} \quad (13.2).
\]

The DI expresses the difference between the testing and training RMS errors as a
proportion of the training RMS error, hence it indicates the degree to which a network
finds spectra 'dissimilar'. The degree to which a SGSG network found spectra more similar (i.e. generalization) could be increased by less 'overtraining' through the use of a more relaxed criterion (higher RMS error). Given that networks with sine transfer functions train much faster than those with sigmoid transfer functions, one could be tempted to substitute a SNSG network trained to the more relaxed criterion for a SGSG network trained to the same criterion. However, as the data in Figure 13.4 show, networks with different transfer functions performed differently on some types of spectral data. In particular, a less trained SNSG network tended to perform like a highly trained SGSG network on smoothed data (SM10) and like a less trained SGSG network on raw and shifted data (RAW, SH5, SH10). Hence, these networks are in general not interchangeable.

Figure 13.4. This figure shows the consequences of a relaxation of the training criterion on SGSG networks from a low root mean square (RMS) error value (SGSGL) to a higher RMS error value (SGSGH). The performance of a SNSG network with the more relaxed criterion (SNSGH) is shown for comparison. The discrimination index is defined in the text and the testing sets in the caption of Figure 13.2. SNSG networks use sine and sigmoid transfer functions for the hidden and output layers, respectively, and SGSG networks use only sigmoid transfer functions.
The SNSG network consistently generated the highest DIs, except for data that were shifted by 10 cm\(^{-1}\). The SNSG network was more sensitive to finer detail, a conclusion that is supported by the particular sensitivity shown to data smoothed with a 10-point moving average which resulted in the reduction of relative peak heights. Although the SNSG network was also sensitive to uniform shifts in spectral peaks, this sensitivity was similar to that of the SGSG network (SH10). These results are shown in Figure 13.5 and they suggest that SNSG networks should be used where rapid learning and fine discrimination are essential and that SGSG networks should be used where slight spectral distortions are a problem. For instance, the system described by Tanabe et al. (1992) could possibly benefit from using SGSG networks for the initial classification of spectra into subcategories and SNSG networks for making the finer discriminations and final identification.

![Graph](image)

**Figure 13.5.** Performance of neural networks using different transfer functions: a sigmoid function (SGSG) and a sine function (SNSG). Testing sets are as given in Figure 13.2.
It is often important to be able to discriminate between different but closely related compounds. If one is not interested in the precursors or metabolites of neurotransmitters (or other bioactive substances), a network must be able to reject the spectra of these closely related compounds that are not of interest. At the same time, the network should be able to discriminate between and correctly identify closely related compounds of interest such as the neurotransmitters. The preceding results indicated that a SNSG network would perform better on such a discriminatory task than a SGSG network.

The Raman spectrum of 5-hydroxytryptophan, a precursor of serotonin, was used to test the abilities of a SNSG and a SGSG network to correctly identify closely related compounds. This was repeated with the Raman spectra of epinephrine and norepinephrine, two closely related catecholamine neurotransmitters. Both the SNSG and SGSG networks had the same 150-24-10 configuration, were trained with the same Raman spectra until the same RMS error criterion (0.02) was reached. As before, the SNSG network trained about 4 times faster than the SGSG network and as predicted, the network using a sine transfer function produced the higher DIs, as illustrated in Figure 13.6, indicating that this type of network is superior at tasks requiring fine discrimination.
Figure 13.6. Abilities of two neural networks (SNSG, SGSG) with different transfer functions (sine, sigmoid) to discriminate between various spectra: 5-hydroxytryptophan (TRP, not in the training set) and serotonin (in the training set) as well as between epinephrine (ENP) and norepinephrine (NNP) and vice versa (both in the training set).

It is often important to be able to discern changes that occur in analytes(s) of interest. Difference spectroscopy is a commonly used method to detect such changes. When two samples are collected at different times for analysis, changes in the latter sample relative to the former can be detected: the Raman spectrum of the former sample is subtracted from the Raman spectrum of the latter sample and the resulting spectrum can be given as input to a neural network for identification. Given the expected need to perform difference spectroscopy on the data collected with this neuroprobe (see Chapter 1), the viability of this approach, using the networks trained to perform the discriminations described in the preceding section, was investigated. The Raman
spectrum of Dulbecco's modified Eagle medium (DMEM) with 5% calf serum and 0.2% sodium azide was measured. Following this, acetylcholine was added to the DMEM to give an approximate concentration of acetylcholine in DMEM of 0.5 M and then determined the Raman spectrum of the admixture. The Raman spectrum of the DMEM was then subtracted from the Raman spectrum of the acetylcholine-spiked DMEM and the difference spectrum given as input to a neural network trained to identify the individual spectra of aqueous solutions of pure small-molecule neurotransmitters ('clean' neurotransmitters).

This procedure was repeated, but instead of DMEM, cerebrospinal fluid (CSF), obtained from the left lateral cerebral ventricle of a male rat under sodium pentobarbital anesthesia, was spiked with acetylcholine to give a concentration of acetylcholine in CSF of about 0.2 M. The SGSG network readily recognized these difference spectra as acetylcholine spectra, but not with the same degree of confidence that a clean acetylcholine spectrum was recognized. Furthermore, the DIs of the SGSG network for the acetylcholine-spiked biological matrices were lower than those produced by the SNSG network (see Figure 13.7). These results confirm the observations in the preceding sections that SGSG networks were more robust and less susceptible to slight and moderate spectral distortions, while SNSG networks were more susceptible to such distortions due to their better discriminating abilities.
13.2.4. Input features.

The performance of networks on input vectors consisting of the uniform sampling (every 10th point in order to reduce the size of the network but to remain within a range where the spectra were adequately represented) or of interval means (10-point intervals) of the training data sets were evaluated. Two 150-24-10 networks with sigmoid-sigmoid and sine-sigmoid transfer functions, respectively, were used. As expected, it was found that the sine-sigmoid network, using either uniformly sampled or interval means data, gave the better DIs. However, using interval means as input reduced the DIs for both SNSG and SGSG networks. The lowest DI values were calculated for the SGSG network using interval means as input. This indicates that such a method of data processing could be used if additional generality or robustness is required for a SGSG network. The results
of investigating network performance with different input features are shown in Figure 13.8.

Figure 13.8. Performance of neural networks using different transfer functions: sine (SNSG) or sigmoid (SGSG); and different input features: interval means (IM) or uniform sampling (US). Testing sets as for Figure 13.2.

13.3. DISCUSSION.

The Raman spectra of aqueous solutions of biological samples (the 10 small-molecule neurotransmitters) can be accurately identified with a neural network. A network using a sine transfer function between hidden and output layers and a sigmoid transfer function for the output layer showed the most rapid learning, a finding in agreement with other reports in the literature. Of particular interest was the finding that the choice of transfer function influenced the degree of discrimination shown by
networks. The SNSG network was more sensitive to fine detail and in general more discriminating. The SGSG network was less sensitive to fine detail and more tolerant to slight and moderate spectral distortions. The network size also had a bearing on identification accuracy. Accuracy generally increased with the amount of input information. Test data with a resolution of 10 cm\(^{-1}\) proved adequate for the identification of the neurotransmitter spectra investigated. Furthermore, a ratio of the number of hidden layer to input layer nodes of 0.16 represented an optimum trade-off between accuracy and network size. Also, the input features used for network training and testing influenced the discriminating ability of the network: uniformly sampled data lead to better discrimination between spectra while data based on interval means lead to better generalization.

The performance of SNSG and SGSG networks on spectra of varying similarity was tested. SNSG networks were better able to make fine discriminations between related spectra (e.g. a neurotransmitter and related precursor) but less able to generalize to the more dissimilar spectra obtained by difference spectroscopy. On the difference spectra used here, the SGSG networks performed better than the SNSG networks. This is most likely due to the fact that the complex biological matrices in which the analytes were embedded produced some artifacts resulting in spectral distortions. SGSG networks were more impervious to spectral distortions and hence performed better than SNSG networks on the difference spectra.

These findings indicate that networks with different transfer functions should be used to accomplish different ends: for accurate discrimination sine functions should be employed while sigmoid functions lead to better generalization. Therefore, data that are subject to distortions should be processed with SGSG networks. Figure 13.9 summarizes many of the findings presented here. It shows the different general trends which are likely to be observed when networks with different characteristics are presented with spectra manipulated in various ways. These trends are shown as a continuum for ease of
interpretation and the figure is not intended to imply overt or covert relationships between the different types of spectral distortion (e.g. between loss of detail and loss of position). The DIs of networks with different discriminating abilities will show inverted parabola-like shapes with walls of different curvature when plotted against different types of spectral changes (e.g. loss of position and loss of detail). The dashed line in Figure 13.9 represents a network capable of better generalization while the dotted line represents a network capable of better discrimination on the same data sets. The distance between the minimum and the abscissa is influenced by the amount of noise present in the data: the more noise added to the data sets, the further the minimum is removed from the abscissa. Hence, the solid line in Figure 13.9 illustrates the performance of a discriminating network and the dotted line gives the results of the same network when Gaussian noise is added to all the testing data sets. In Figure 13.9, the point '0' on the abscissa represents the training data set.

It is not clear why these differences in generalization are observed between networks employing sine and sigmoid transfer functions, but they are most likely due to the different error surfaces associated with the abilities of networks with different transfer functions to form complex mappings of the input (Lee, 1992).
Figure 13.9. A schematic illustration of the discriminating abilities of ideal networks on hypothetical data sets. The solid line represents the performance of a discriminating network on noiseless data while the dotted line represents the performance of the same network on data with an arbitrary noise content. The dashed line represents the performance of a network with a better ability to generalize between data sets while the dotted line represents a network with a better ability to discriminate between data sets. The negative numbers on the abscissa represent an arbitrary increase in the loss of spectral detail (e.g. increasing the degree of smoothing) and the positive numbers an arbitrary increase in the loss of spectral peak positions (e.g. increasing the extent of spectral shifts). The '0' point on the abscissa represents the spectral characteristics of the training data set. The discrimination index is given in arbitrary units and defined in the text. The figure is intended to show general trends only.

Although there are other methods for abscissa error correction in a Raman spectrum (Shen et al., 1992), the findings reported here indicate that the SGSG network, which was more sensitive to spectral shifts, could also be used for this purpose. As the abscissa error of the calibrating standard increases, the SGSG network will show a
corresponding increase in the RMS error. Adjusting the instrumentation until the RMS error is minimized will result in abscissa error correction.

Finally, these findings suggest that using a system of cascaded ANNs with different transfer functions could produce a superior identification system. Such a system would be a refinement of the method of Tanabe et al. (1992). The first network in the cascade should be a SGSG network to effect a spectral classification. Once a spectrum has been classified (e.g., belonging to the class of alcohols), it is then given as an input to a SNSG network to uniquely identify that spectrum. If high quality Raman spectra could be obtained, the SGSG-SNSG cascade could be expected to perform very well, even on very similar spectra. If, on the other hand, spectra suffer from random distortions, a SGSG-SGSG cascade could be expected to perform better.

13.4. REFERENCES.


14.1. INTRODUCTION
14.2. MIXTURE RESOLUTION
14.2.1. Artificial neural networks
   14.2.1.1. Training
   14.2.1.2. Layer number and size
   14.2.1.3. Transfer functions
   14.2.1.4. Split layers
   14.2.1.5. Local connections
14.2.2. Classical least squares
14.3. DISCUSSION
14.4. REFERENCES

14.1. INTRODUCTION

Two widely used techniques for the \emph{in vivo} measurements of neurotransmitters are \emph{in vivo} voltammetry and \emph{in vivo} dialysis. With both dialysis and voltammetry, it is important to be able to decompose signals into their constituent components since, in a given sample volume, several neurotransmitters may coexist as a mixture, and the detection of any given neurotransmitter is often confounded by the presence of other compounds. For instance, the voltammetric measurements of \emph{in vivo} monoamine signals are prone to interference by other substances with similar oxidation potentials (e.g. Gonzalez-Mora et al., 1991). Although the voltammetric techniques are naturally selective for neurotransmitters that are electrochemically active at useful potentials, some compounds such as ascorbic acid and uric acid nevertheless interfere with the measurements of the catecholamines (e.g. Stamford, 1989). Furthermore, the signals of the catecholamines overlap, making accurate measurements of individual components in mixtures of catecholamines difficult.

Dialysis, in contrast to voltammetry, can be used with a variety of detection methods to measure almost anything. Many of the commonly used detection methods are susceptible to interferents (e.g. Roettger and Goldfinger, 1991). Overlapping dialysis
signals have to be resolved with the standard methods of chromatography: different stationary phases, different temperatures, different mobile phase flow rates, and so on.

In general, the identification and quantification of individual components in a sample mixture can be a serious methodological problem. Consequently, several methods have been advanced to deal with this problem (e.g. Lo and Brown, 1992; Gonzalez-Mora et al., 1991). Some of these methods employ artificial neural networks (ANNs) and are applied to various spectra, e.g. colorimetric (Lin et al., 1992), and near infrared spectra (Liu et al., 1993). Another method, the classical least squares method (CLS), employs a least squares analysis based on linear combinations of known spectra as described in Chapter 11. Neural networks and the classical least squares method were employed to identify the components in neurotransmitter mixtures.

14.2. MIXTURE RESOLUTION

In the previous chapter it was shown that neural networks were capable of identifying the 10 small-molecule neurotransmitters based on their Raman spectra. Therefore, it remained to be demonstrated that quantitative information could also be obtained with neural networks. The purpose of the research reported in this chapter was to investigate the effectiveness of artificial neural networks and the classical least squares method for quantitative mixture analysis of Raman spectra of the 10 small-molecule neurotransmitters as a further step in the development of the neuroprobe.

14.2.1. Artificial neural networks

Although neural networks are often used as pattern classifiers (e.g. Tanabe et al., 1992), they have not traditionally been used for quantitative estimation. Recently, however, some investigators have successfully demonstrated the utility of ANNs for quantitative estimation. Borggaard and Thodberg (1992) combined principal component analysis with neural networks for qualitative and quantitative spectral processing. Lin et
al. (1992) used a calibration curve to interpret the outputs of a self-organizing network when given unknown spectra from their samples, while Liu et al. (1993) used some data preprocessing and multiple networks, each with a single output node, to estimate the concentrations of different chemical components in a mixture.

Several parameters affecting network architecture and function were investigated in order to optimize the network for mixture resolution. These were layer sizes, transfer functions, layer splitting, and local connections.

14.2.1.1. Training

Backpropagation artificial neural networks were trained, as in Chapter 13, using the Raman spectra of the small-molecule neurotransmitters, covering the spectral region between 2000 cm\(^{-1}\) to 500 cm\(^{-1}\) from the laser excitation frequency. Training to the criteria described in Chapter 13 does not guarantee optimal training, and it should be noted that there are different ways to estimate the point of optimal training for a network, see for example the discussion by Hush and Horne (1993). Furthermore, the quantitative interpretation of ANN outputs are not theoretically well-developed, hence mixture quantification could be difficult. This was addressed by including scaled spectra in the training set and, in one case, using a linear transfer function.

The training set consisted of 36 spectra. These were normalized spectra, spectra with values of 50%, and 20% of the normalized spectra, as well as some linear combinations of those spectra. The 24 testing spectra consisted of different linear combinations of the normalized spectra (BASIC), the BASIC testing data set to which Gaussian noise with a standard deviation of 0.20 was added (NOISE), and the BASIC testing data set with the spectra shifted toward the excitation frequency by 5 cm\(^{-1}\) (SH5). The networks were also tested with the two acetylcholine difference spectra (CSFDIF, DMDIF) described in a previous section, and finally with three neurotransmitter spectra and one neurotransmitter mixture spectrum (INDEP) obtained under similar conditions,
but independently from the BASIC data. The makeup of the training and testing data sets are given in Tables 14.1 and 14.2, respectively. For networks with 150 input nodes, these spectra were sampled at uniform intervals of 10 cm\(^{-1}\) and for networks with 250 input nodes, at uniform intervals of 6 cm\(^{-1}\). Linear transfer functions were used for all input layers except where otherwise noted.

Table 14.1.
This table shows the desired output node values for the neurotransmitter Raman spectra used in the training data set. Abbreviations: ACH, A = acetylcholine; DOP = dopamine; EPI, EP = epinephrine; NOR = norepinephrine; SER = serotonin; HIS = histamine; ASP = aspartate; GAB = γ-amino butyric acid (GABA); GLU = glutamate; GLY = glycine; and combinations of these.

<table>
<thead>
<tr>
<th>Spectrum</th>
<th>ACH</th>
<th>DOP</th>
<th>EPI</th>
<th>NOR</th>
<th>SER</th>
<th>HIS</th>
<th>ASP</th>
<th>GAB</th>
<th>GLU</th>
<th>GLY</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACH</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>DOP</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>EPI</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NOR</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SER</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HIS</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ASP</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GAB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GLU</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>GLY</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>ACH5</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GLY5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>.5</td>
</tr>
<tr>
<td>ACH2</td>
<td>0.2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GLY2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>.2</td>
</tr>
<tr>
<td>ACHDOP</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ACHSER</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ACHGAB</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>EPGLY</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>ASERGAB</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>EPNORGLY</td>
<td>0</td>
<td>0</td>
<td>0.3</td>
<td>0.3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 14.2.
This table shows the desired output node values for the neurotransmitter Raman spectra in the BASIC, NOISE, SH5, difference spectra, and INDEP testing data sets. See text for further explanation. Abbreviations as in the text and in Table 1; MIX = equimolar mixture of ACH, DOP, and EPI (the spectrum was not normalized and the table values reflect the relative Raman signal intensities for these three neurotransmitters).

<table>
<thead>
<tr>
<th>Spectrum</th>
<th>Neurotransmitter</th>
<th>ACH</th>
<th>DOP</th>
<th>EPI</th>
<th>NOR</th>
<th>SER</th>
<th>HIS</th>
<th>ASP</th>
<th>GAB</th>
<th>GLU</th>
<th>GLY</th>
</tr>
</thead>
<tbody>
<tr>
<td>BASIC/NOISE/SH5</td>
<td>EPIDOP</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>HISER</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>ACH5DOP</td>
<td>0.5</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>DOPHIS5</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>AEPILGLY</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>ADOPEPI</td>
<td>1</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Difference spectra</td>
<td>CSFDFIF</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>DMDIF</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>INDEP</td>
<td>ACH</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>DOP</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>EPI</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>MIX</td>
<td>0.62</td>
<td>1</td>
<td>0.58</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

14.2.1.2. Layer number and size

It is known that a three-layer network can solve virtually any nonlinear decision problem, but in some cases an intractably large number of hidden nodes are required - this can be avoided by using a small four-layer network instead (see discussion by Hush and Horne, 1993). It is also known that the size of the hidden layer can influence network performance (e.g. Hush and Horne, 1993). Consequently, the number of hidden layers and later the size of one of the hidden layers were increased to investigate the effects of these changes on network performance in mixture analysis. All the networks used for this part of the study employed the hyperbolic tangent function. The networks were trained with 10,000 random presentations of the training data and these results are shown in Table 14.3.
Table 14.3.
The test results of three artificial neural networks when given the BASIC testing set as input. The networks had different configurations but used only hyperbolic tangent transfer functions. Consult the text for further details. Tfunction = transfer function; RMS = root mean square error value.

<table>
<thead>
<tr>
<th>Network</th>
<th>Tfunction</th>
<th>RMS</th>
<th>Trials</th>
</tr>
</thead>
<tbody>
<tr>
<td>150-24-10</td>
<td>TNTN</td>
<td>1.408</td>
<td>10000</td>
</tr>
<tr>
<td>150-24-10-10</td>
<td>TNTNTN</td>
<td>1.568</td>
<td>10000</td>
</tr>
<tr>
<td>150-50-10-10</td>
<td>TNTNTN</td>
<td>1.180</td>
<td>10000</td>
</tr>
</tbody>
</table>

It can be seen from Table 14.3 that enlarging the network from a 150-24-10 network to a 150-24-10-10 network by adding another hidden layer did not improve the network performance. However, increasing the size of the first hidden layer to give a 150-50-10-10 network, did produce better results.

It can be argued that the performance of an ANN ought to improve with an increase in input information. In a previous investigation (Schulze et al., 1994), it was found that increasing the number of input nodes resulted in an improvement of network operation. Hence, the number of input nodes were increased from 150 to 250, while keeping the ratio of hidden to input nodes constant at 0.16. The output of this 250-40-10 network failed to converge, even after 20,000 random presentations of the training data. However, when the size of the hidden layer was increased to produce a 250-200-10 network, the size of the RMS error was much reduced. This result confirmed earlier observations that an increase in the number of hidden layer nodes produced an improvement in performance. In spite of this noted improvement, though, the larger network could not reduce its RMS error any further after about 4,000 training presentations. When the same network was retrained, but this time using sigmoid transfer functions, another improvement was observed.
14.2.1.3. Transfer functions

The abilities of networks with different transfer functions to recognize the individual components of neurotransmitter mixtures from the Raman spectra of the mixtures were tested. The abbreviations SG, SN, and TN for sigmoid, sine and hyperbolic tangent functions, respectively, were used. The order in which they are used reflects the order in which information is propagated through the network starting with the input layer and ending with the output layer. Since input layers frequently use linear transfer functions, these were omitted except where otherwise noted. Previous work indicated that different transfer functions can have very different effects on the performance characteristics of networks (Schulze et al., 1994). In particular, networks using a sine and sigmoid transfer function for the hidden and output layers (SNSG), respectively, learn much faster and are more sensitive to finer spectral detail than similar networks using only sigmoid transfer functions (SGSG). It was thus hypothesized that a network with a sine transfer function would learn the training data set fast, but that it would generalize poorly to mixtures not contained in the training set. In contrast, the SGSG networks were better at generalization and were expected to perform better on mixtures not encountered before. In addition to the sine and sigmoid transfer functions, a hyperbolic tangent function (TNSG) was also investigated for comparison. The latter function is rather similar to the sigmoid function and it was expected to perform at a comparative level. Networks used were 150-24-10 (input-hidden-output) networks trained to the same RMS error (0.1) criterion. The results, shown in Table 14.4, confirmed the hypothesis that the SNSG network would train more rapidly but perform more poorly on novel data. The TNSG network performed better than the SGSG network; it reached the training criterion faster and it was better at mixture identification.
Table 14.4.
The test results of three artificial neural networks when given the BASIC testing set as input. The networks used different hidden layer transfer functions: sigmoidal (SGSG); hyperbolic tangent (TNSG); and sine (SNSG) transfer functions. Consult the text for further details. Tfunction = transfer function; RMS = root mean square error value.

<table>
<thead>
<tr>
<th>Network</th>
<th>Tfunction</th>
<th>RMS</th>
<th>Criterion</th>
<th>Trials</th>
</tr>
</thead>
<tbody>
<tr>
<td>150-24-10</td>
<td>SGSG</td>
<td>1.988</td>
<td>0.1</td>
<td>3856</td>
</tr>
<tr>
<td>150-24-10</td>
<td>SNSG</td>
<td>2.675</td>
<td>0.1</td>
<td>1776</td>
</tr>
<tr>
<td>150-24-10</td>
<td>TNSG</td>
<td>1.544</td>
<td>0.1</td>
<td>1457</td>
</tr>
</tbody>
</table>

14.2.1.4. Split layers

Further scrutiny of the network output produced from the testing data revealed that the use of different transfer functions in the larger networks also resulted in differential performances. The network using two sigmoid transfer functions (SGSGL) produced better results on the binary mixtures in the testing data set, whereas the network with hyperbolic tangent functions (TNTNL) was better at identifying the ternary mixtures.

This suggested that using a combination of these two transfer functions would enhance the network performance. Given the large difference in the number of intermediate and output layer nodes, it was speculated that combining these two transfer functions by employing a different function in each layer would merely skew the results in favor of the transfer function used by the layer with the most nodes. It was therefore decided to split the hidden layer into two equal sections and to employ a hyperbolic tangent function in the one part of the layer and a sigmoid function in the other part. The outputs from both parts of the split hidden layer were combined at the output layer where a sigmoid function was used. As expected, this network (SG/TN) yielded better results than the TNTNL network on the binary mixtures and the SGSGL network on the tertiary mixtures. This could be due to differential processing of spectral regions by different transfer functions in a split hidden layer (Gemperline et al., 1991). A further improvement was obtained when the hidden layer was split into three sections with a
different number of nodes per section and using sigmoid (90), hyperbolic tangent (70), and sine transfer functions (40), respectively. The latter network (SG/TN/SN) even outperformed the SGSGGL network (good with binary mixtures) and the TNTNL network (good with ternary mixtures) on average. Table 14.5 contains a summary of the results obtained in this part of the investigation.

Table 14.5.
The test results of artificial neural networks when given the BASIC testing set as input. (a) The networks had configurations and used transfer functions as indicated in the table. (b) The table shows a detailed breakdown of the network performance (RMS error values) on the binary and tertiary mixtures in the BASIC set plus two additional binary mixtures. Also shown are RMS error values based only on the average values of the training set (AVG). Consult the text for further details. Tfunction = transfer function (or network name as it appears in the text); RMS = root mean square error value.

<table>
<thead>
<tr>
<th>(a) Network</th>
<th>Tfunction</th>
<th>RMS</th>
<th>Trials</th>
</tr>
</thead>
<tbody>
<tr>
<td>150-24-10</td>
<td>TNTN</td>
<td>1.408</td>
<td>10000</td>
</tr>
<tr>
<td>250-40-10</td>
<td>TNTN</td>
<td>7.403</td>
<td>20000</td>
</tr>
<tr>
<td>250-200-10</td>
<td>(TNTNL)</td>
<td>1.341</td>
<td>20000</td>
</tr>
<tr>
<td>250-200-10</td>
<td>(SGSGGL)</td>
<td>0.790</td>
<td>20000</td>
</tr>
<tr>
<td>250-200-10</td>
<td>(SG/TN)</td>
<td>0.632</td>
<td>20000</td>
</tr>
<tr>
<td>250-200-10</td>
<td>(SG/TN/SN)</td>
<td>0.535</td>
<td>20000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(b) Network</th>
<th>(TNTNL)</th>
<th>(SGSGGL)</th>
<th>(SG/TN)</th>
<th>(SG/TN/SN)</th>
<th>AVG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Binary</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EPIDOP</td>
<td>.125</td>
<td>.022</td>
<td>.038</td>
<td>.036</td>
<td>0.417</td>
</tr>
<tr>
<td>HISSE</td>
<td>.251</td>
<td>.090</td>
<td>.069</td>
<td>.049</td>
<td>0.417</td>
</tr>
<tr>
<td>ACH5DOP</td>
<td>.077</td>
<td>.022</td>
<td>.037</td>
<td>.034</td>
<td>0.327</td>
</tr>
<tr>
<td>DOPHIS5</td>
<td>.163</td>
<td>.059</td>
<td>.119</td>
<td>.112</td>
<td>0.327</td>
</tr>
<tr>
<td>GLUGAB</td>
<td>.316</td>
<td>.098</td>
<td>.039</td>
<td>.047</td>
<td>0.417</td>
</tr>
<tr>
<td>ACHHIS5</td>
<td>.109</td>
<td>.040</td>
<td>.035</td>
<td>.020</td>
<td>0.327</td>
</tr>
<tr>
<td>Σ</td>
<td>1.041</td>
<td>.331</td>
<td>.337</td>
<td>.298</td>
<td>2.232</td>
</tr>
<tr>
<td>Ternary</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AEPIGLY</td>
<td>.215</td>
<td>.319</td>
<td>.190</td>
<td>.166</td>
<td>0.507</td>
</tr>
<tr>
<td>ADOPEPI</td>
<td>.085</td>
<td>.140</td>
<td>.105</td>
<td>.071</td>
<td>0.352</td>
</tr>
<tr>
<td>Σ</td>
<td>.300</td>
<td>.459</td>
<td>.295</td>
<td>.237</td>
<td>0.859</td>
</tr>
</tbody>
</table>

14.2.1.5. Local connections

Both of the acetylcholine difference spectra were incorrectly identified as mixtures although both the biomatrices were spiked only with acetylcholine. The SG/TN/SN network identified the CSF difference spectrum as a mixture of dopamine (0.829) and acetylcholine (0.610) and the DMEM difference spectrum also as a mixture.
of dopamine (0.977) and acetylcholine (0.687). These results were not satisfactory given the fact that ANNs which had not been trained with mixtures readily identified these difference spectra as acetylcholine spectra (Schulze et al., 1994).

It is known that the nodes of a network with local connections process only a local region of the input (see discussion by Hush and Horne, 1993). Furthermore, it has been found that networks with restricted receptive fields produce efficient non-redundant data representations (Idiart et al., 1995). Hence a network (LC1) with local connections to enhance the network's sensitivity to local spectral features was designed. In a network with local connections every hidden node is connected to only a small number of immediate neighbor input nodes instead of to all input nodes. The input layer of this network used a sigmoidal transfer function, the hidden layer used a sine transfer function, and the output layer used a sigmoidal transfer function. Although this network performed less well than the SG/TN/SN network on the mixture testing data, it performed better on the difference spectra. The CSF spectrum was identified as a mixture of acetylcholine (0.575) and dopamine (0.334) and the DMEM spectrum as a mixture of acetylcholine (0.736) and dopamine (0.585).

Although it was earlier found that increasing the size of the hidden layer improved network performance, it is believed that networks with fewer hidden layer nodes than training samples (i.e. fewer than 36 nodes in the present case) would yield the best results (Hush and Horne, 1993). Consequently, a 250-20-10 network (LC2) with sigmoid, sine and linear transfer functions respectively, on the same spectra used before, but now sampled at 4 cm\(^{-1}\) intervals in the 1500 to 500 cm\(^{-1}\) region was trained and tested. This network showed an improved performance over the LC1 network.

14.2.2. Classical least squares

The normalized spectra from the BASIC testing data set were used to create a spectral library for the CLS method. The CLS method was then tested with the NOISE,
SH5, CSFDIF, DMDIF, and INDEP testing data. The ANN and CLS methods were compared on the basis of their root mean square (RMS) testing results.

Since the artificial mixture spectra are simply linear combinations of the neurotransmitter spectra in the spectral library, the CLS method was expected to identify the mixture components perfectly. These results are shown in Table 14.6 and, for comparison, also the results of the LC1 network on the same data. Clearly, the CLS method is superior on this type of data.

<table>
<thead>
<tr>
<th>Spectrum</th>
<th>Neurotransmitter</th>
<th>CLS</th>
<th>LC1 network</th>
<th>Neurotransmitter</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPI DOP</td>
<td>ACH</td>
<td>DOP</td>
<td>EPI</td>
<td>NOR</td>
</tr>
<tr>
<td>HIS SER</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>ACH5 DOP</td>
<td>0.5</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>DOP HIS5</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AEPI GLY</td>
<td>1</td>
<td>0</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>ADO PEPI</td>
<td>1</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>ACH DOP</td>
<td>.05</td>
<td>.83</td>
<td>.96</td>
<td>.32</td>
</tr>
<tr>
<td>HIS SER</td>
<td>.08</td>
<td>.02</td>
<td>.00</td>
<td>.01</td>
</tr>
<tr>
<td>ACH5 DOP</td>
<td>.51</td>
<td>.76</td>
<td>.08</td>
<td>.03</td>
</tr>
<tr>
<td>DOP HIS5</td>
<td>.10</td>
<td>.63</td>
<td>.04</td>
<td>.04</td>
</tr>
<tr>
<td>AEPI GLY</td>
<td>.39</td>
<td>.04</td>
<td>.91</td>
<td>.02</td>
</tr>
<tr>
<td>ADO PEPI</td>
<td>.83</td>
<td>.51</td>
<td>.32</td>
<td>.04</td>
</tr>
</tbody>
</table>

In order to determine the abilities of the CLS method to identify distorted spectra, we processed the acetylcholine difference spectra with this method. In addition, we processed the SH5, NOISE, and INDEP testing data sets with the CLS method. These spectra were also processed with the LC2 network and the results are given in Table 14.7 for comparison. Table 14.7 shows that the CLS method identified the shifted spectra better than the LC2 network with average RMS values of 0.134 and 0.188, respectively,
and on the NOISE data set with average RMS values of 0.114 and 0.166, respectively. We have previously shown that ANNs can be very sensitive to spectral shifts, especially those that use hidden layer sigmoid transfer functions, and that data preprocessing could reduce this sensitivity (Schulze et al., 1994). These results seem to indicate that the networks investigated here were more sensitive to spectral shifts than the CLS method. The spectra not in the training set or library were identified as well as or better than the CLS method by the LC2 network with average RMS values of 0.181 and 0.161 ± 0.004 (S.E.M.) respectively. The LC2 network was trained and evaluated 5 times with different sets of starting weights.

Table 14.7.
The RMS values and composition estimates of the acetylcholine difference spectra and data in the INDEP testing set obtained with the classical least squares method (CLS) and a neural network (LC2) using local processing. Also shown are the RMS values for the NOISE (5 runs) and SH5 testing sets.

<table>
<thead>
<tr>
<th>Spectrum</th>
<th>CLS</th>
<th>Neurotransmitter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ACH</td>
<td>DOP</td>
</tr>
<tr>
<td>CSFDF</td>
<td>.66</td>
<td>.00</td>
</tr>
<tr>
<td>DMDF</td>
<td>.70</td>
<td>.10</td>
</tr>
<tr>
<td>INDEP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACH</td>
<td>.74</td>
<td>.17</td>
</tr>
<tr>
<td>DOP</td>
<td>-.01</td>
<td>.76</td>
</tr>
<tr>
<td>EPI</td>
<td>.12</td>
<td>.16</td>
</tr>
<tr>
<td>MIX</td>
<td>.42</td>
<td>.68</td>
</tr>
<tr>
<td>NOISE</td>
<td>RUN1</td>
<td>RUN2</td>
</tr>
<tr>
<td>EPIDOP</td>
<td>0.133</td>
<td>0.162</td>
</tr>
<tr>
<td>HISER</td>
<td>0.056</td>
<td>0.062</td>
</tr>
<tr>
<td>ACH5DOP</td>
<td>0.114</td>
<td>0.135</td>
</tr>
<tr>
<td>ACHHIS5</td>
<td>0.073</td>
<td>0.127</td>
</tr>
<tr>
<td>AEPIGLY</td>
<td>0.141</td>
<td>0.176</td>
</tr>
<tr>
<td>ADOPEPI</td>
<td>0.129</td>
<td>0.168</td>
</tr>
<tr>
<td>SH5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EPIDOP</td>
<td>0.135</td>
<td></td>
</tr>
<tr>
<td>HISER</td>
<td>0.117</td>
<td></td>
</tr>
<tr>
<td>ACH5DOP</td>
<td>0.176</td>
<td></td>
</tr>
<tr>
<td>ACHHIS5</td>
<td>0.150</td>
<td></td>
</tr>
<tr>
<td>AEPIGLY</td>
<td>0.106</td>
<td></td>
</tr>
<tr>
<td>ADOPEPI</td>
<td>0.119</td>
<td></td>
</tr>
</tbody>
</table>
14.4. Discussion

This investigation revealed for the first time that the abilities of artificial neural networks to identify the components of neurotransmitter mixtures from their Raman spectra depended on several factors. Networks with hyperbolic tangent transfer functions learned relatively fast and generalized well to the test mixtures, those with sine functions learned fast but generalized poorly, while networks with sigmoidal functions learned more slowly but generalized well. Neither adding a second hidden layer nor increasing the size of the input layer markedly improved network performance, but when coupled with an increase in the number of nodes in the first hidden layer, both manipulations
produced improved mixture identification. In the enlarged networks, the choice of transfer function still produced noticeable effects; networks with sigmoidal functions yielded good results with binary mixtures and networks with hyperbolic tangent functions gave better results with ternary mixtures. A network with a hidden layer sectioned into three parts, with each part employing a different transfer function (hyperbolic tangent, sigmoid, and sine, respectively), produced better results than either of the preceding two. The best overall results on all the testing data sets, however, were obtained with a network using a form of local input processing, a combination of sigmoid, sine, and linear transfer functions, and a reduced sampling interval over a smaller spectral region. A further improvement is likely if all the data points in critical spectral regions common to all the neurotransmitters are used as input to networks with local processing architectures. In addition, the preprocessing of the spectral input data such as using interval means instead of uniform interval sampling (e.g. Schulze et al., 1994), or the expansion of the training data set to include more examples of single neurotransmitter and mixture spectra, could also enhance the network performance.

The classical least squares method proved excellent at mixture resolution with artificially produced spectra consisting of linear combinations of spectra contained in the spectral library used by this method, and performed well on shifted and noisy spectra. However, with spectra of the same neurotransmitters that were measured under similar conditions, but independently of the library spectra, the neural networks performed better. It should be noted that the CLS method is relatively easy to implement and to modify. For instance, expanding the number of compounds to be analyzed involves merely the addition of new spectra to the spectral library. Furthermore, the CLS method can be fine-tuned by appropriate weighting. The ANNs in contrast, often require extensive investigation and training, and adding new compounds to the set of analytes often involves complete retraining of the networks.
The ability of a network to generalize from the training to the testing data is considered to be mainly influenced by the number of data samples, the complexity of the underlying problem, and the network size (Hush and Horne, 1993). We have recently shown that the use of different transfer functions also influences generalization (Schulze et al., 1994). The effects of different transfer functions on the training speed of a network have been established and are understood (e.g. Zhang and Wang, 1992), but their precise effects on network generalization have not yet been mathematically elucidated (Lee, 1992). However, it is believed that different transfer functions, as well as other factors such as the number of hidden layers, provide networks with different capabilities to form complex, nonlinear mappings of the input (Lee, 1992). These differences in network mapping abilities based on different transfer functions and network architectures (e.g. number of hidden layers, number of nodes per layer, structure of connections between layers, etc.) could provide some explanation for the findings presented here that network data processing differed due to differences in such factors as transfer functions and network architecture. Since the details of these relationships are not yet fully understood, it remains necessary to systematically investigate these factors in the context of specific applications in order to optimize performance. However, the results reported here demonstrate the feasibility of using Raman spectroscopy, artificial neural networks, and the classical least squares method for the analysis of mixtures of small-molecule neurotransmitters.

14.5. REFERENCES


15.1. INTRODUCTION.

Artificial neural networks employing non-linear transfer functions are excellent tools for pattern classification of non-linearly separable data. For this task, the role of the hidden layer(s) is to project the patterns of the input space into an intermediate space in which they can be easily separated by the output layer. In so doing, however, the data structure and interpattern distances are severely distorted (Mao and Jain, 1995). Principal component analysis networks, in contrast, are less effective in pattern classification, but do preserve data structure, cluster shape, interpattern distances, and nearest-neighbor category information (Mao and Jain, 1995).

Therefore, it was an interesting question to consider whether the same ANN could be employed for both pattern classification and quantification. The purpose of this chapter was to continue to develop neural network-based methods for the quantification and identification of neurotransmitters in mixtures based on the Raman (and by implication resonance Raman) spectra of the neurotransmitters and of mixtures containing them.
15.2. QUANTIFICATION WITH NEURAL NETWORKS.

Artificial neural networks (ANNs) can be trained to partition a pattern space with arbitrary precision and are consequently often employed as pattern classifiers for tasks such as spectral identification (e.g. Tanabe et al., 1992; Schulze et al., 1994). Some researchers have extended this application of ANNs to determine the composition of mixtures from spectral data (Lerner and Lu, 1993; Liu et al., 1993; Lin et al., 1992). In order to quantitatively estimate all the chemical components of a mixture simultaneously, some researchers have used single (e.g. Schulze et al., 1995) and others multiple (e.g. Liu et al., 1993) networks. However, due to the non-linear nature of ANNs, it is not easy to interpret the network outputs quantitatively.

In order to obtain quantitative information from ANN outputs, an iterative approach applied to network outputs, was developed. This consisted of the identification of the major constituent of a mixture and then subtracting that spectrum from the mixture spectrum, rescaling the residual mixture spectrum, and repeating the process until some predetermined termination threshold value was reached.

15.2.1. Training.

The Raman spectra of the small-molecule neurotransmitters, as described above, were used for ANN training and testing. A backpropagation artificial neural network with local connections consisting of 250 input nodes, 30 hidden nodes, and 10 output nodes, one output node corresponding to each of the 10 small-molecule neurotransmitters, was used. The input layer of this network used a sigmoidal transfer function, the hidden layer used a sine transfer function, and the output layer used a linear transfer function. In a previous study it was found that networks with hidden layer sine transfer functions were more accurate in spectral identification, less sensitive to spectral shifts, but generalized less well than sigmoidal functions (Schulze et al., 1994). It was
envisaged that a network with a hidden layer sine transfer function and local connections would perform well on mixture recognition tasks, while retaining generalizability.

The network was trained with a data set that consisted of spectra that were low pass filtered by smoothing with a symmetrical 5-point moving average to remove high frequency random noise. The spectra used were sampled at uniform intervals of 4 cm\(^{-1}\) (to generate 250 network inputs from the 1001 points in each spectrum). The training set consisted of 36 spectra. These were normalized spectra, spectra with values of 50%, and 20% of the normalized spectra, as well as some linear combinations of those spectra. The 6 testing spectra, used to test the ability of the network to generalize to new data, consisted of different linear combinations of the normalized spectra. The analyzed HISSER (a binary compound mixture consisting of histamine and serotonin) and AEPIGLY (a tertiary compound mixture consisting of acetylcholine, epinephrine, and glycine) spectra belonged to this group. Although the number of training spectra could be considered small for the adequate training of the network, the primary aim here was neither to validate a neural network architecture nor to optimally train one, but to demonstrate a procedure for obtaining quantitative information from a neural network output. The network was trained for 20,000 trials consisting of the random presentation of the training spectra.

15.2.2. Spectral removal procedure.

Quantitative mixture analysis with ANNs using the spectral removal technique consisted of the following steps:

(i) A neural network was trained to identify all of the individual components likely to occur in the mixtures under investigation. These normalized spectra of individual components were referred to as the library spectra. The training set also included some mixture spectra.
(ii) Once trained, the components of a mixture were identified from a mixture spectrum normalized with respect to peak height. A fraction of the library spectrum of the component $p$ for which the network produced the largest output value was then removed from the mixture spectrum using that output value as the removal fraction (RF). Note that, due to the non-linear nature of ANNs, networks sometimes overestimated the amount of a component present in the mixture. Such overestimation was countered by using a smaller percentage ($\alpha$) of the RFs at the cost of increasing the number of iterations required for mixture resolution. Assessing the proper value for $\alpha$ is discussed in a later section. The removal fraction was defined by:

$$RF_i = \max_i [\text{ANN}(S_i)]$$

(15.1)

where $S_i$ is the $i$-th residual spectrum (for the original mixture spectrum $i = 0$), and $[\text{ANN}(S_i)]$ are the ANN outputs for $S_i$.

(iii) The remaining mixture spectrum ($S_{i+1}$) was then renormalized with regard to peak height and given to the network as input.

$$S_{i+1} = \frac{S_i - (\alpha/100)RF_i(L_p)}{SF_i}$$

(15.2)

where $L_p$ is the library spectrum of component $p$, $\alpha$ is the percentage by which $RF_i$ can be adjusted to improve spectral decomposition, and $SF_i$ is a scaling fraction given by

$$SF_i = \max_i [S_i - (\alpha/100)RF_i(L_p)]$$

(15.3).

(iv) The process was repeated until the recovered component fractions (CFs) fell below a certain preselected threshold value (e.g. 0.01). The recovered fraction of component $p$ consisted of the current RF multiplied by all the preceding SFs:

$$CF_{i,p} = (\prod(SF_i))(\alpha/100)RF_i$$

(15.4)

where $\prod(SF_i)$ is the product of all prior SFs.

(v) The summation of all the CFs for a particular component and their normalization with regard to the largest sum produced the relative amounts of the components present in the mixture. The final network estimates, only of those components recovered previously, were included in this sum.
\[(\text{Final Estimate})_p = \Sigma CF_{i,p}\] (15.5)

where \((\text{Final Estimate})_p\) was summed over all \(i\) to give the total contribution of component \(p\) to the mixture.

15.2.3. Binary mixtures.

The library spectra of histamine and serotonin were added in the ratio 1.000:1.000, respectively, and normalized with regard to peak height to produce the binary compound mixture spectrum HISSER. This spectrum was then given as input to the trained ANN which identified the mixture as consisting of histamine, serotonin, and acetylcholine in the ratio of 0.824:0.712:0.242, respectively. Subtracting 0.824 of the histamine library spectrum from the mixture spectrum yielded a residual spectrum with a maximum peak height of 0.789. The latter value was then used as a scaling fraction to renormalize the residual mixture spectrum. The ANN subsequently identified the renormalized residual spectrum as consisting of histamine, serotonin, and epinephrine in the ratio of 0.064:1.019:0.116, respectively. Subtracting 1.019 times the serotonin library spectrum from the residual mixture spectrum produced another residual mixture spectrum with a maximum peak height of 0.000. The latter value precluded renormalization of the remaining mixture spectrum and the process was terminated. The component fractions were determined by multiplying the RF with all the preceding SFs. For histamine: \(CF = 0.824 \times 1\) (\(RF \times SF\)) = 0.824. For serotonin: \(CF = 1.019 \times 1 \times 0.789\) (\(RF \times SF \times SF\)) = 0.804. The CFs of histamine and serotonin give the mixture composition then as 1.000:0.976.

Although the network sometimes identified mixture components spuriously (e.g. acetylcholine and epinephrine in the present example), the outputs associated with these components were smaller than those of the correct mixture components, often fluctuated and diminished with increasing iterations, and rarely contributed to the final composition estimates (see the discussion about spectral shifts below).
15.2.4. Tertiary mixtures.

The library spectra of acetylcholine, epinephrine, and glycine were added in the ratio 1.000:1.000:1.000, respectively, and normalized with regard to peak height to generate the tertiary compound mixture spectrum AEPIGLY. Using the spectral removal technique with an $\alpha$ value of 80%, the mixture composition was analyzed. The process was terminated when the CF fell below 0.01. The mixture proved to consist of acetylcholine, epinephrine, and glycine in the ratio of 0.994:1.000:0.991, respectively. The results obtained are given in Table 15.1.

<table>
<thead>
<tr>
<th>Spectrum</th>
<th>Ach</th>
<th>Epi</th>
<th>Gly</th>
<th>Other</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original</td>
<td>0.978</td>
<td>1.027</td>
<td>0.622</td>
<td>0.140 (ser)</td>
<td>0.150</td>
</tr>
<tr>
<td>Residual1</td>
<td>0.985</td>
<td>0.237</td>
<td>0.703</td>
<td>0.107 (ser)</td>
<td>0.076</td>
</tr>
<tr>
<td>Residual2</td>
<td>0.213</td>
<td>0.193</td>
<td>0.810</td>
<td>0.085 (nor)</td>
<td>0.061</td>
</tr>
<tr>
<td>Residual3</td>
<td>0.350</td>
<td>0.260</td>
<td>0.781</td>
<td>0.089 (ser)</td>
<td>0.067</td>
</tr>
<tr>
<td>Residual4</td>
<td>0.722</td>
<td>0.406</td>
<td>0.704</td>
<td>0.108 (ser)</td>
<td>0.082</td>
</tr>
<tr>
<td>Residual5</td>
<td>0.139</td>
<td>0.400</td>
<td>0.772</td>
<td>0.125 (nor)</td>
<td>0.072</td>
</tr>
<tr>
<td>Residual6</td>
<td>0.124</td>
<td>0.731</td>
<td>0.714</td>
<td>0.157 (nor)</td>
<td>0.096</td>
</tr>
<tr>
<td>Residual7</td>
<td>0.100</td>
<td>-0.049</td>
<td>0.871</td>
<td>0.068 (ser)</td>
<td>0.043</td>
</tr>
<tr>
<td>Residual8</td>
<td>0.018</td>
<td>-0.412</td>
<td>0.899</td>
<td>0.062 (ser)</td>
<td>0.009</td>
</tr>
</tbody>
</table>

When the same spectrum was analyzed with $\alpha = 100\%$, large negative values were generated in the residual spectra. Since the library spectra contained no negative values, the appearance of large negative values, resulting in negative average spectral values (the average of all the data points per spectrum), indicated incorrect spectral removal. In addition, the ANN identified 'negative' amounts of epinephrine in all the residual mixture spectra. This implied that the ANN initially overestimated the amount of epinephrine in the mixture. Consequently, too much of the epinephrine spectrum was
removed from the original mixture spectrum, resulting in normalized residual spectra that were difficult to interpret. These results are shown in Table 15.2.

Table 15.2
Consecutive network estimates of the components present in the mixture and the highest estimate for components not present in the mixture (Other) for the AEPIGLY spectrum and residual spectra after component removal. Also shown are spectral average values. $\alpha$ used: 100%. Abbreviations: ach = acetylcholine; epi = epinephrine; gly = glycine; nor = norepinephrine; ser = serotonin.

<table>
<thead>
<tr>
<th>Spectrum</th>
<th>Ach</th>
<th>Epi</th>
<th>Gly</th>
<th>Other</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original</td>
<td>0.978</td>
<td>1.027</td>
<td>0.622</td>
<td>0.140(ser)</td>
<td>0.150</td>
</tr>
<tr>
<td>Residual1</td>
<td>0.966</td>
<td>-0.045</td>
<td>0.764</td>
<td>0.080(ser)</td>
<td>0.057</td>
</tr>
<tr>
<td>Residual2</td>
<td>0.034</td>
<td>-0.151</td>
<td>0.893</td>
<td>0.065(ser)</td>
<td>0.034</td>
</tr>
<tr>
<td>Residual3</td>
<td>0.117</td>
<td>-0.775</td>
<td>0.493</td>
<td>0.197(nor)</td>
<td>-0.169</td>
</tr>
<tr>
<td>Residual4</td>
<td>0.634</td>
<td>-0.817</td>
<td>0.145</td>
<td>0.337(nor)</td>
<td>-0.394</td>
</tr>
<tr>
<td>Residual5</td>
<td>0.780</td>
<td>-0.794</td>
<td>0.152</td>
<td>0.347(nor)</td>
<td>-0.414</td>
</tr>
<tr>
<td>Residual6</td>
<td>0.848</td>
<td>-0.728</td>
<td>0.150</td>
<td>0.336(nor)</td>
<td>---------</td>
</tr>
</tbody>
</table>

15.2.5. Measured mixture spectra.

The Raman spectra of two mixtures of neurotransmitters were also analyzed with this technique. Both mixtures contained extensive overlap between major peaks of the neurotransmitters. Figure 15.1 shows these experimentally measured mixture spectra as well as an acetylcholine spectrum, a component in both mixtures, for comparison.

The first mixture (MIX1) contained acetylcholine and serotonin with maximum peak heights in the ratio 0.95:1.00, respectively. Using an $\alpha$ of 80%, the ratio of acetylcholine to serotonin was established as 0.89:1.00. The second mixture (MIX2) contained acetylcholine, dopamine, and epinephrine with maximum peak heights in the ratio 0.62:1.00:0.58, respectively. The spectral removal technique was applied to MIX2 using $\alpha$ values of 100%, 80%, 50%, and 30%. The results of using $\alpha = 50\%$ are shown in Table 15.3 and indicate the ratio of acetylcholine, dopamine, and epinephrine in the mixture to be 0.59:1.00:0.58, respectively.
Figure 15.1. *The Raman spectra of acetylcholine (a) and two experimentally measured mixtures, MIX1 (b) and MIX2 (c), both containing acetylcholine. In the mixture spectra, considerable overlap of the major component peaks occur.*

Table 15.3

Consecutive network estimates of largest component in mixture (Estimate), the fraction of that component removed from the mixture spectrum $RF (RF = \alpha \times \text{Estimate})$, the scaling factor $SF$ used to renormalize the remaining mixture spectrum, and CFs obtained for every component recovered from the MIX2 spectrum. Value of $\alpha$ used: 50%.

Abbreviations: ach = acetylcholine; dop = dopamine; epi = epinephrine; Resid = Residual.

<table>
<thead>
<tr>
<th>Spectrum</th>
<th>Estimate</th>
<th>RF</th>
<th>SF</th>
<th>CF</th>
<th>Component</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original</td>
<td>0.619</td>
<td>0.310</td>
<td>1.000</td>
<td>0.310</td>
<td>epi</td>
</tr>
<tr>
<td>Residual1</td>
<td>0.697</td>
<td>0.349</td>
<td>0.729</td>
<td>0.254</td>
<td>ach</td>
</tr>
<tr>
<td>Residual2</td>
<td>0.674</td>
<td>0.337</td>
<td>0.958</td>
<td>0.235</td>
<td>dop</td>
</tr>
<tr>
<td>Residual3</td>
<td>0.666</td>
<td>0.333</td>
<td>0.686</td>
<td>0.160</td>
<td>ach</td>
</tr>
<tr>
<td>Residual4</td>
<td>0.630</td>
<td>-0.315</td>
<td>0.966</td>
<td>0.146</td>
<td>dop</td>
</tr>
<tr>
<td>Residual5</td>
<td>0.594</td>
<td>0.297</td>
<td>0.734</td>
<td>0.101</td>
<td>epi</td>
</tr>
<tr>
<td>Residual6</td>
<td>0.828</td>
<td>0.414</td>
<td>0.738</td>
<td>0.104</td>
<td>dop</td>
</tr>
<tr>
<td>Residual7</td>
<td>0.877</td>
<td>0.439</td>
<td>0.662</td>
<td>0.073</td>
<td>dop</td>
</tr>
<tr>
<td>Residual8</td>
<td>0.881</td>
<td>0.441</td>
<td>0.727</td>
<td>0.053</td>
<td>dop</td>
</tr>
<tr>
<td>Residual9</td>
<td>0.762</td>
<td>0.381</td>
<td>0.837</td>
<td>0.038</td>
<td>dop</td>
</tr>
<tr>
<td>Resid.10</td>
<td>0.621</td>
<td>0.311</td>
<td>0.860</td>
<td>0.027</td>
<td>dop</td>
</tr>
<tr>
<td>Resid.11</td>
<td>0.498</td>
<td>0.249</td>
<td>0.886</td>
<td>0.019</td>
<td>dop</td>
</tr>
<tr>
<td>Resid.12</td>
<td>0.385</td>
<td>0.193</td>
<td>0.908</td>
<td>0.013</td>
<td>dop</td>
</tr>
<tr>
<td>Resid.13</td>
<td>0.295</td>
<td>0.148</td>
<td>0.971</td>
<td>0.010</td>
<td>ach</td>
</tr>
<tr>
<td>Resid.14</td>
<td>0.307</td>
<td>0.154</td>
<td>0.969</td>
<td>0.010</td>
<td>dop</td>
</tr>
<tr>
<td>Resid.15</td>
<td>0.240</td>
<td>0.240</td>
<td>0.993</td>
<td>0.008</td>
<td>dop</td>
</tr>
<tr>
<td>Resid.16</td>
<td>0.237</td>
<td>0.237</td>
<td>0.993</td>
<td>0.008</td>
<td>epi</td>
</tr>
<tr>
<td>Resid.17</td>
<td>0.179</td>
<td>0.179</td>
<td>0.993</td>
<td>0.006</td>
<td>ach</td>
</tr>
</tbody>
</table>
15.2.6. Removal fraction coefficient.

Given that a network could overestimate a mixture component and that a smaller RF coefficient (i.e. percentage of the RF) should then be used during the iteration process, it remains to be determined how this coefficient should be chosen. There appear to be three indicators of incorrect spectral decomposition. First, the average of the residual mixture spectral values generally becomes negative before iterations are terminated (i.e. before the CF falls below the selected small threshold value, as indicated in Table 15.2). Second, the network produces 'negative' estimates for a component(s) subtracted from the mixture spectrum before iterations terminate. As mentioned before, this is indicative of a prior overestimation of the mixture component(s). Third, the sum of the CFs of individual components converge to the correct mixture composition when the \( \alpha \) value used is such that the overall sum of CFs is a minimum. This is demonstrated in Figure 15.2 where the sum of the CFs for the AEPIGLY spectrum are plotted against the \( \alpha \) value used during iterations. The same figure also contains the analogous values for the MIX2 spectrum. In both cases the correct mixture composition corresponded to an \( \alpha \) of 50\%. Although it is unlikely that an \( \alpha \) of 100\% will produce the correct outcome, the a priori selection of an \( \alpha \) value that will result in convergence to the correct estimates must be determined on a case-by-case basis, according to an iteration strategy based on the indicators mentioned above.
Figure 15.2. The sum of all the recovered component fractions are plotted against the coefficient of the removal fraction used during spectral removal iterations. This graph illustrates that component fraction sums tend to converge to a minimum which is indicative of correct spectral recovery. The results shown are for the AEPIGLY (broken line, left scale) and MIX2 (solid line, right scale) spectra.

It can be seen from Figure 15.2 that, in both cases, the sum of CFs decline and level off with decreasing coefficient. The following iteration strategy is therefore proposed: if during an iteration the mean of the remaining spectrum becomes negative or the network output for a specific component becomes negative after its subtraction from the mixture spectrum, indicating that that component may have been overestimated, the value of $\alpha$ should be reduced. However, if the sum of CFs differ little from that obtained with the previous (i.e. higher) $\alpha$ value, there may be no need for a further reduction in the value of $\alpha$ used. For spectra with low signal-to-noise ratios, spectral subtraction will produce negative values, hence using mean spectral values to judge whether $\alpha$ should be adjusted will be difficult and the other two indicators should be relied upon instead. It becomes obvious, however, when an inappropriate choice for $\alpha$ is used, and 2 or 3 trials
will generally lead to convergence to the correct mixture composition. Note that the number of iterations increase with decrease in \( \alpha \). Figure 15.3 shows the number of iterations required to resolve a mixture against the \( \alpha \) used.

Figure 15.3. The number of iterations required for mixture decomposition as a function of the \( \alpha \) value used. The results shown are for the AEPILGY (broken line, right scale) and MIX2 (solid line, left scale) spectra. This figure indicates that the number of iterations required to resolve a mixture into its individual components is inversely proportional to the value of \( \alpha \) used.

15.2.7. Multi-component mixtures.

It was not within the intended scope of this investigation to carry out an exhaustive empirical characterization of the algorithm. The maximum number of components in a mixture that can be successfully recovered with this method likely depends on the signal-to-noise ratio of the spectra and on how distinct the component spectra are. We have analyzed mixtures containing as many as 10 components. For
example, a 5-component mixture spectrum was generated by adding the normalized spectra of dopamine, serotonin, histamine, γ-amino butyric acid, and glycine in equal ratios. These respective components were recovered in the following ratio 0.995:0.997:0.992:1.000:0.993. Spuriously identified components did not exceed 0.005. The value of $\alpha$ used was 80%. A 10-component mixture, containing randomly chosen amounts of all the small-molecule neurotransmitters, was resolved to within 1-6% of the component values, except for one component that was recovered to within 10% of its real value. The $\alpha$ used was 30% and it seems possible that a further improvement could be obtained with a lower $\alpha$. These successful recoveries demonstrate the usefulness of this method in analyzing multicomponent spectra.

15.2.8. Spectral shifts.

ANNs are sensitive to spectral phase shifts which suggests that ANNs could be vulnerable to errors of calibration. Networks with sine transfer functions, however, have been found to be less sensitive to such spectral shifts than those employing sigmoidal transfer functions (Schulze et al., 1994). The effects of spectral shifts of mixture spectra on ANN output was examined. The MIX2 spectrum was shifted by 1, 2, 3, and 4 cm$^{-1}$ toward the excitation frequency before being processed by the network. The resulting network estimates for the three mixture components for each shifted spectrum are presented in Table 15.4.

Table 15.4
The table shows the network estimates for the three components present in the mixture of the MIX2 spectrum when the spectrum was shifted towards the excitation frequency by 0, 1, 2, 3, and 4 cm$^{-1}$.

<table>
<thead>
<tr>
<th>Shift</th>
<th>Ach</th>
<th>Dop</th>
<th>Epi</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.562</td>
<td>0.529</td>
<td>0.619</td>
</tr>
<tr>
<td>1</td>
<td>0.457</td>
<td>0.480</td>
<td>0.632</td>
</tr>
<tr>
<td>2</td>
<td>0.376</td>
<td>0.492</td>
<td>0.613</td>
</tr>
<tr>
<td>3</td>
<td>0.282</td>
<td>0.446</td>
<td>0.652</td>
</tr>
<tr>
<td>4</td>
<td>0.185</td>
<td>0.444</td>
<td>0.656</td>
</tr>
</tbody>
</table>
It is clear from Table 15.4 that spectral shifts cause the network estimates to vary considerably. Furthermore, as discussed previously, the use of smaller α values are likely to reduce the effects of any relative overestimates because smaller fractions of the library spectra are subtracted from the mixture spectrum. An additional complication produced by spectral shifts is the corruption of the residual mixture spectra. Because the mixture spectrum to be analyzed is frequency shifted with regard to the library spectra, subtracting the library spectra from the mixture spectrum will result in 'derivative-like' phenomena in the residual mixture spectra. These errors become amplified during subsequent iterations and hence lead to the progressive deterioration of the mixture spectrum. Therefore, they will become increasingly difficult to interpret for the ANNs. When high values of α are used, the residual mixture spectra should also show a rapid increase in negative values. Conversely, when lower values of α are used, the residual spectra are not likely to show negative values, but they will become altered to the extent of being artifactual, thus leading to spurious component identification. For instance, when the AEP1GLY spectrum was shifted toward the excitation frequency by 4 cm\(^{-1}\) and α = 50% was used for component recovery, the network incorrectly identified an increasing amount of serotonin in the residual spectra. After 5 iterations, serotonin was in fact identified as the major component of the residual spectrum.

A remedy for problems associated with spectral shifts consists of using the component with the largest ANN output to estimate the spectral shift and adjusting the spectrum to be analyzed accordingly. In the case referred to above, the network originally identified epinephrine as the major component of the shifted AEP1GLY spectrum. The frequency of the highest peak in the mixture spectrum is compared to the frequency of the highest epinephrine peak in the library spectrum and the mixture spectrum is adjusted until the peaks have the same frequency. Given the non-linearity of ANNs, the highest peak in the mixture spectrum may on occasion not belong to that component associated with the highest network output value. Therefore, it may be
necessary to adjust the mixture spectrum based on components associated with other high output values.

Alternatively, a network could be specifically trained to recognize spectra despite the presence of small spectral shifts. The use of such a network would allow more accurate spectral corrections to be made as described above. Two networks with the same architecture, training files, and training criteria as described earlier, except that spectra were sampled at 6 cm\(^{-1}\) and that one network's training data set was augmented by 48 spectra consisting of each of the 10 neurotransmitter spectra as well as one binary and one tertiary mixture spectrum shifted away from the excitation frequency by 1, 2, 3, and 4 cm\(^{-1}\) were trained. When tested with a 4 cm\(^{-1}\) shifted mixture spectrum of acetylcholine, epinephrine, and glycine in a 1:1:1 ratio (not in the training set), the ANN not trained with shifted spectra identified the components in the ratio 0.690:1.043:1.022; the ANN trained with the shifted spectra in the ratio 0.897:0.972:1.044; and when employing the classical least squares method for comparison, the ratio obtained was 0.807:0.982:0.915.

15.3. DISCUSSION.

The work described in this chapter clearly demonstrates that ANNs can be used for the quantitative mixture resolution of small-molecule neurotransmitters. A novel iterative spectral removal technique was employed to remove estimated amounts, or fractions of the estimates of individual component spectra, from the mixture spectrum until the recovered increments fell below a predetermined threshold value. Reducing these fractions improved the mixture resolution (at the expense of increased computation time). A negative residual spectrum mean and/or a negative estimate of a component occurring after its removal were indicative of the need to reduce the fraction of the estimate used for recovery. Apparent frequency shifts in mixture spectra due to calibration errors caused corrupted residual mixture spectra and consequently incorrect
component identification. This was corrected by estimating the spectral shift from the component with the highest estimate and correcting the mixture spectrum by that amount. A more accurate estimate can be obtained with an ANN specifically trained to identify shifted spectra.

This spectral removal approach yielded satisfactory recovery of the individual components of artificially generated and experimentally measured Raman spectra of neurotransmitter mixtures (generally within 1-6%). Its accuracy can be adjusted depending on the quality of the spectra to be analyzed. Excessive noise levels and/or spectral shifts could be expected to make mixture resolution more difficult. Although normalized composite mixture spectra were not used in the training of the network, mixture spectra analyzed here were normalized before every analysis. Normalization often lead to less accurate initial network estimates. In spite of this, however, the method converged to the correct mixture compositions. Furthermore, as the successive residual mixture spectra were pruned of their components, they were simplified and were more readily identified by the network.

ANNs have the ability to perform a rich variety of nonlinear parallel computations (e.g. Lippmann, 1987) that enable researchers to model complex data sets without the need to understand the underlying physical principles giving rise to those data sets (e.g. Bakshi and Stephanopoulos, 1993). ANNs are therefore mainly applied to non-linear problems. However, due to the non-linearity of ANNs, it can be difficult to interpret the network outputs quantitatively. In the method presented here, an iterative subtraction technique was applied to data that was generated/or assumed to show a linear relationship between signal intensity and concentration. This method was purposely developed to allow quantitative information to be obtained from single ANNs.

It should be kept in mind that other (linear) methods, such as classical least squares decomposition and principal component analysis, exist that may be more suitable for application to linear regression problems. In addition, ANNs can be combined with
such methods (e.g. Gemperline et al., 1991; Borggaard and Thodberg, 1992; Blank and Brown, 1993). These hybrid regression-network methods lead to a reduction in the number of network inputs required, increases in training speed, and improved quantitative performance. However, the present method makes it possible to obtain quantitative information from single ANNs without the need for preprocessing and/or the use of hybrid methods.

15.4. REFERENCES.


PART VI
Evaluation, Summary and Future Directions

CHAPTER 16

16.1. INTRODUCTION

The 4 areas of investigation covered in this thesis were applied to the problem of neurotransmitter measurement and identification in order to integrate them into a single unit. To this end, cell secretions from PC12 cells were used as samples (Part IV, Chapter 9), a fiber-optic probe was employed (Part III, Chapter 8) to measure the UV resonance Raman spectra of neurotransmitters (Part II, Chapter 5) in these secretions and the signals were identified with an artificial neural network (Part V, Chapter 14). The application discussed in this chapter is a major milestone for gauging the degree of development of the neuroprobe and its readiness for in vivo use.

16.2. AN APPLICATION: THE MEASUREMENT OF NEUROTRANSMITTER SECRETIONS

16.2.1. The samples

Cell secretions from PC12 cells were obtained as discussed in detail in paragraph 9.3.3. However, in order to minimize the possible contamination of the sample with
components from the cell growth medium, a medium without the phenol red pH indicator was used for culturing the cells, trypsin solution without indicator was used to remove adhering cells from the flask (cell dissociation buffer 3153SA, GibcoBRL, Burlington, Ont.), and the cells were washed consecutively with 40 and 30 ml of physiological saline solution, respectively (as opposed to two consecutive washes of 5 ml each). In a further attempt to identify components secreted by the cells from those originating elsewhere, the cells were split after counting into 2 equal samples containing approximately 4.27x10^6 cells each and depolarized for different periods of time. Each sample received 0.5 ml depolarizing agent. Immediately after, the cells in sample F were centrifuged, the supernatant removed, ultrafiltered, and stored on ice as described before, while the cells in sample G were returned to the incubator for another 40 minutes, then centrifuged, the supernatant removed, filtered, and stored on ice. This procedure resulted in an affective depolarization period of 5 minutes for sample F (the shortest possible) and 45 minutes for sample G. Sample F was intended to serve as a baseline to determine whether the peaks in an observed spectrum were due to components released by the cells during the depolarization period. An aliquot of 0.1 ml each was obtained from both samples and stored at -70 °C for 5 days until HPLC analysis could be performed. The samples F and G were then subjected to resonance Raman measurement.

16.2.2. The fiber-optic probe

A dual optical fiber probe was fabricated (see Chapter 8) and used for resonance Raman measurements. The probe consisted of a 400 µm excitation fiber suitable for pulsed UV applications (Chapter 7) and an angled 600 µm collection fiber (Chapter 8). The angled face of the collection fiber was given 3 facets to improve its collection efficiency. The length of the excitation fiber was approximately 20 cm and that of the collection fiber
approximately 50 cm. The angled face was coated with an aluminum mirror as described in Chapter 8 and protected with a thin layer of epoxy resin.

16.2.3. The neural network

The baselines of some measured UV resonance Raman spectra of neurotransmitters and precursors/metabolites (e.g. Chapter 5) were removed by polynomial fit, and the spectra smoothed by autoaccumulation (30 cm⁻¹ bandwidth, normally distributed) according to the procedures described in Chapter 12. The spectra were then normalized and sampled at 10 cm⁻¹ from 1800 cm⁻¹ to 910 cm⁻¹ giving 90 points per spectrum in total. In the case of the monoamines, dopamine, epinephrine, and norepinephrine, a Lorentzian distribution was fitted to the resonance peak in order to more clearly discriminate between these compounds (see Chapter 5). The Lorentzian peak was then spliced into the original spectrum replacing the original peak. These spliced spectra were then sampled as described above. Spectra of dopamine, epinephrine, serotonin, histamine, melatonin, tryptophan and 2 linear combinations of these (0.5 dopamine + 0.5 melatonin; 0.5 epinephrine + 0.5 serotonin) were used to create a training file. Another 2 linear combinations (0.5 histamine + 0.5 serotonin; 0.5 histamine + 0.5 tryptophan), the spliced epinephrine spectrum as well as a different non-spliced epinephrine spectrum were used to create a test file. Because of the great similarity between epinephrine and norepinephrine spectra, norepinephrine was not included in the training or testing sets.

A 3-layer 90-9-6 neural network with sigmoid, sine, and linear transfer functions per layer, respectively, was created for signal identification (see Chapter 14). This local connection network had connections between every 10 input nodes and a hidden layer
node. The hidden and output layers were fully connected. The network was instantiated 5 times with different starting conditions and each instance was trained for 20,000 trials using the training file and then tested with the testing file. This number of trials was sufficient to produce output convergence and the correct identification of the compounds in the training and testing sets with a RMS mean and standard error of the mean of \(0.007 \pm 0.001\), and \(0.043 \pm 0.003\), respectively.

16.2.4. The resonance Raman measurements

The resonance Raman spectra obtained from both samples F and G are shown in Figure 16.1 and the difference spectrum in Figure 16.2.

![Raman spectra graph](image)

Figure 16.1. The UV resonance Raman spectra of cell secretion samples F (5 minute depolarization, top trace) and G (45 minute depolarization, bottom trace) obtained at 227 nm, 20 Hz, and approximately 400 μW employing a fiber-optic probe.
These spectra were measured with the probe using excitation at 227 nm (optimal for dopamine) and average power at the sample of approximately 400 μW. Monochromator settings were 233.8 nm, the entrance slit open to 300 μm, and a total of 180 s integration time per spectrum was used. Spectra of pure ethanol (for calibration), depolarizing solution (for background correction), and 0.5 mM epinephrine (for peak identification) were also collected.

16.2.5. The results

16.2.5.1. HPLC results

Both samples were analyzed in duplicate by HPLC as described in Chapter 9. While the results for dopamine, epinephrine, and norepinephrine were similar to those
observed before (Table 9.2), those for serotonin were about tenfold lower. It is possible that the change in culture medium was responsible for these variations. There was also a noticeable difference between the samples with regard to depolarization time. The sample depolarized for the longest period yielded the highest amounts of neurotransmitter. Furthermore, except for epinephrine, the concentrations differed by a similar factor between samples F and G indicating that proportionately more epinephrine than other neurotransmitters was released during the depolarization period. The results are shown in Table 16.1.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Norepinephrine</th>
<th>Epinephrine</th>
<th>Dopamine</th>
<th>Serotonin</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>8.15x10^-8</td>
<td>2.28x10^-8</td>
<td>5.49x10^-6</td>
<td>8.51x10^-8</td>
</tr>
<tr>
<td></td>
<td>9.46x10^-8</td>
<td>2.79x10^-8</td>
<td>5.88x10^-6</td>
<td>8.90x10^-8</td>
</tr>
<tr>
<td>G</td>
<td>3.07x10^-7</td>
<td>5.63x10^-7</td>
<td>2.30x10^-5</td>
<td>2.78x10^-7</td>
</tr>
<tr>
<td></td>
<td>3.41x10^-7</td>
<td>6.26x10^-7</td>
<td>2.47x10^-5</td>
<td>2.87x10^-7</td>
</tr>
<tr>
<td>Average factor of increase</td>
<td>3.70</td>
<td>23.45</td>
<td>4.20</td>
<td>3.25</td>
</tr>
</tbody>
</table>

The chromatograms of the two samples are shown in Figure 16.3 and on an expanded vertical scale in Figure 16.4. From these figures, it is clear that, in addition to the neurotransmitters reported in Table 16.1, several unknown compounds were also released by the cells during the extended depolarization period. As can be seen from Figure 16.3, one of these unknown compounds was secreted in copious quantities and increasing with a factor of 14 over the depolarization period.
Figure 16.3. The full-scale chromatograms of cell secretion samples F (5 minute depolarization, dotted line) and G (45 minute depolarization, solid line).

Figure 16.4. The scale-expanded chromatograms of cell secretion samples F (5 minute depolarization, dotted line) and G (45 minute depolarization, solid line).
16.2.5.1. Neuroprobe results

The difference spectrum of samples F and G were obtained by subtracting the spectrum of sample F from 6 times that of sample G. The baseline of this spectrum was corrected, sampled every 10 cm\(^{-1}\) from 1800 cm\(^{-1}\) to 910 cm\(^{-1}\), and given without any smoothing as input to the 5 versions of the neural network described above for identification. Consistent with the HPLC results, the neuroprobe results indicated a strong presence of dopamine. However, in contrast with the HPLC results, epinephrine, norepinephrine, and serotonin, were not observed since these components were present in concentrations below those detectable with the neuroprobe. Interestingly, signals indicative of histamine and melatonin were also identified in the cell secretion difference spectrum. The results of the neuroprobe analysis of the cell secretion samples are shown in Figure 16.5.
Dopamine could be clearly detected with the neuroprobe in secretions from rat pheochromocytoma cells. The presence of dopamine was confirmed by HPLC analysis of the same samples and indicated an approximately four-fold increase in concentration over the 40 minute additional depolarization period. This was in general agreement with the fact that an approximately six-fold attenuation occurred in the Raman scattering produced by water (see Figure 16.1) due to increased absorption (mostly by dopamine, but also by other transmitters).

In addition to dopamine, histamine appeared to be present in these samples. Histamine is present in all mammalian tissues, mostly in mast cells, but also in adrenal
chromaffin cells (see Carmichael, 1986), the normal counterparts of pheochromocytoma
cells (Greene and Rein, 1977). Rat enterochromaffin-like cells secrete histamine in
response to secretagogues such as gastrin and acetylcholine by calcium-dependent
exocytosis (Loo et al., 1996). Several lines of human glioma cells have been shown to
secrete histamine into their culture media (Hirohata et al., 1995). In addition, histamine
preferentially stimulates the release of epinephrine from bovine adrenergic chromaffin
cells (Nunez et al., 1995). Taken together, the evidence strongly suggests that histamine
could indeed be present in secretions from PC12 cells. The sharp increase observed in
epinephrine secretion (see Table 16.1) is certainly consistent with this possibility.
However, histamine is also released by cells in response to trauma (see Kadar, 1985),
hence the histamine could have originated from manipulation of the cells (e.g. washing,
centrifuging, trypsinizing, etc.). This could account for the observed release of
neurotransmitters after the wash but before depolarization as shown in Figure 9.1. A useful
modification to the investigation described here would be to add two additional controls
consisting of cells ‘depolarized’ respectively for 5 and 45 minutes with the vehicle only in
order to tease apart sample components originating in the cell culture medium, those
released due to trauma but not depolarization, and those released due to depolarization
only.

Although histamine was a likely constituent of the cell secretion samples and there
was excellent agreement between the spectrum of histamine and the peaks in the cell
secretion difference spectrum attributable to histamine, there was also a noticeable shift
(about 40 cm\(^{-1}\)) between the histamine peaks in the two spectra. It is possible that the
histamine precursor, histidine, produces a spectrum similar to that of histamine, but
displaced by the amount noticed here. This displacement precluded the use of the spectral removal method to identify other neurotransmitters possibly present in the cell secretions (e.g. melatonin) because such peak shifts induce artifacts into the residual spectra leading to spurious component identification (see Chapter 15). An attempt to positively identify histamine with the HPLC failed because the detector on the instrument was not suitable for histamine detection. This experiment also highlighted the need for continuous and careful calibration of the instrument and the utilization of water as an internal standard to verify consistency of the calibration between spectra. The presence of several compounds with highly overlapping and similar spectra (the monoamines and their precursors and metabolites - see also Chapter 5) furthermore emphasized the need to attain and maintain the highest resolution possible with the instrument in order to obtain positive identification of these constituents.

Although much work remains to be done to optimize the probe for in vivo applications, it is practically ready for in vivo testing given a tunable deep ultraviolet continuous wave laser, a high quality detector (e.g. charge coupled device), and a high throughput spectrograph with high dispersion (see also Chapter 17). In addition to these improvements, a more sensitive probe would be very useful. Some effort should be made to realize the benefit of multiple collection fiber probes. To gain greater confidence in the performance of the neuroprobe, the positive identification of histamine as well as the unknown compounds shown in Figure 16.2 are required. Regarding signal identification and quantification, artificial neural networks are well suited to the considerable problem of dealing with the nonlinear variation in signal intensity of multiple components with concentration and excitation wavelength. Wavelength tuning can provide a significant
means of selectivity, but requires the measured spectra of all the possible components of a sample at the different wavelengths to be employed. However, the moderate success of the neuroprobe in analyzing cell secretions, especially in the light of the great progress made since its inception, provides reason for optimism. The results reported in this chapter clearly demonstrate the ability of the neuroprobe to detect neurotransmitters in complex samples. With state-of-the-art equipment (detailed in Chapter 17), physiologically relevant detection limits for many neurotransmitters will be attainable and the prospect of practical in vivo ultraviolet resonance Raman spectroscopy is now a real possibility.

16.4. References


17.1. INTRODUCTION

Several factors related to the development of a resonance Raman fiber-optic-based neurotransmitter probe for \textit{in vivo} use have been discussed in the preceding chapters. In this final chapter, the current state of the neuroprobe will be compared to some ideal biosensor characteristics (mostly discussed by Buerk, 1993). This will provide both a sense of the considerable progress made so far and an idea of the work still required to complete a fully-functional resonance Raman neuroprobe.
17.2. IDEAL BIOSENSOR CHARACTERISTICS

Buerk (1993) discusses several pertinent properties and characteristic behaviors of biosensors. Most of these are addressed in the following sections. However, the discussion of dynamic range has been added, and some were extended (e.g. temperature dependence augmented to general dependencies) to reflect those issues most relevant to the neuroprobe.

17.2.1. Dynamic range

The dynamic range of an instrument is the range from the detection limit to response saturation. In the case of resonance Raman spectroscopy, the dynamic range depends on the analyte and solvent characteristics. For non-absorbing solutions, the dynamic range reaches a maximum where the solution becomes saturated with analyte because of the direct dependence of the Raman signal on concentration (see Chapter 2.2). For absorbing solutions, the dynamic range extends to the concentration where the probe signal peaks (e.g. $2 \times 10^{-5}$ M in Figure 8.10, 300x600 probe).

However, due to the use of the probe in aqueous media where a water Raman signal can be used as an internal standard, the dynamic range can be extended. For instance, measuring the water peak at 1630 cm$^{-1}$, which is in the fingerprint region, simultaneously with the analyte signal in this region, would allow one to construct a working curve similar to the methyl orange/potassium nitrate working curves shown in Figure 8.10. The analyte concentration, given a measured analyte signal intensity which could correspond to two possible analyte concentrations, can then be unambiguously determined based on the intensity of the internal standard. For example, a methyl orange signal of 500 measured with a 300x600 probe (Figure 8.10) could indicate a concentration of either $1 \times 10^{-3}$ M or $2 \times 10^{-4}$ M. A simultaneously measured KNO$_3$ signal of 2500 would indicate $1 \times 10^{-3}$ M methyl orange to be the correct concentration. For in vivo work, water would always be present and could fulfill the role of KNO$_3$. 
17.2.2. Linearity

A perfectly linear sensor shows a directly proportional relationship between analyte concentration and signal intensity over the entire possible concentration range of the analyte in a particular solvent. A subset of the data shown in Figure 8.12 are replotted in Figure 17.1 on a linear scale to show the linear range of the front- and side-casting probes in absorbing solutions.

![Graph showing linear range of probes](image)

Figure 17.1. The working curves for front-(dotted lines) and side-casting probes (solid lines) plotted on a linear scale to show the linear ranges of the two probes for Raman resonant (positive slope) and non-resonant (negative slope) analytes in an absorbing medium.

The linear ranges of the probes in non-absorbing solutions have not been determined, but should extend from the detection limit to the onset of non-linear
molecular effects at high concentrations (i.e. should well exceed the linear range in absorbing solutions). Note that the linear range for the non-absorbing analyte in an absorbing solution differs from that of the absorbing analyte for the same probe, due to the non-linear nature of the absorption curve.

17.2.3. Sensitivity

The sensitivity of a sensor is defined as the slope of the sensor signal versus analyte concentration in the sensor's linear region. The sensitivity depends on the efficiency with which a signal can be generated given a specific analyte concentration, i.e., how efficient the analyte can be excited, how efficient scattered or fluorescent light can be collected, etc. Efficiencies have been investigated in Chapter 8. It can be seen from Figure 17.1 that the front- and side-casting probes exhibited different sensitivities, that of the side-casting probe being nearly 3 times higher (calculated for the methyl orange signal).

17.2.4. Dynamic response

The dynamic response of a sensor indicates how quickly it responds to changes in the concentration of the target analyte. Given that resonance Raman scattering occurs on the $10^{-12}$ s time scale, the response to concentration changes should be immediate. Rapid response times is one of the advantages of this detection method. The effective dynamic response of the neuroprobe however, also depends on the time taken to tune to the required wavelengths and to acquire and process spectra. Throughout the development process, the need for the rapid collection of spectra has been emphasized, therefore, spectra were rarely collected using prolonged collection times. The identification and quantification of neurotransmitters with artificial neural networks based on their Raman and resonance Raman spectra were undertaken in order to provide rapid (real time) signal processing.
17.2.5. Hysteresis

Although the hysteresis of the different probes were not investigated, there is little reason to believe that hysteresis would be observed with respect to changes in analyte concentration, except perhaps if some degree of analyte adsorption to the probe occurs. On the other hand, it is likely that hysteresis will be observed with respect to changes in input pulse energies given the sensitivity of the optical fibers to this variable.

17.2.6. Stability

The possibility of signal drift during \textit{in vivo} use was investigated and indicated that some degree of protein fouling did occur (Chapter 10). Quantification with a linear fouling rate gave an \textit{in vivo} lifetime of 140 minutes. Although the stability of the optical system has also been investigated (Chapter 3), continual changes to this system as part of the development process have occurred, hence influencing its stability. Furthermore, laser dyes have finite lifetimes depending on the dye, the peak pulse energies, and the laser repetition rate. These factors often put practical limits on the length of time for which a calibration may be valid or for which spectra could be collected.

17.2.7. Calibration

Two types of calibration are important to ensure obtaining useful measurements with this probe. Wavelength calibration is necessary to enable peak identification and hence analyte identification. Given that various factors could produce slight variations of the optical path inside the monochromator resulting in peak position shifts, the use of a calibrant was necessary prior to commencing measurements to alleviate wavelength calibration problems. For instance, optical path variations could be produced by fiber endfaces not perpendicular to the fiber axis, insertion errors of the fiber into the fiber chuck, or the chuck into the fiber-optic adapter. Pure ethanol was routinely used as a
calibrant, providing 5 well-defined peaks spread across the region of interest. Another potential standard that could be used is tryptophan.

Calibration of signal strength with regard to analyte concentration is necessary for quantification purposes. The working curves shown in Chapter 8 represent such calibration. Ideally, calibration should only be required once and be valid for the entire lifetime of the sensor. If calibrations are required more frequently, they should preferably require measurement of only a few points of the calibration curve. The frequency of calibration and the minimum number of calibration points required have to be established for these probes in the future.

17.2.8. Selectivity

The ideal sensor is only selectively sensitive to the analyte to be measured. As pointed out before, this ideal also constitutes a limitation on the number of analytes that could be measured. The resonance Raman probe effectively combines selectivity with versatility via wavelength tuning (see Chapters 1, 2, and 5) and is one of the strengths of this method.

17.2.9. Background signal

Due to the weakness of Raman relative to Rayleigh scattering and/or fluorescence, a strong background signal that needs to be removed is often present (Davies, 1991, Durman, 1988). Rayleigh scattering can be reduced with the use of optical filters (e.g. Chase, 1994). For use with tunable UV radiation, a tunable UV filter is required. Background reduction with the present optical filter was unsatisfactory (see Chapter 3) and a prism-based system may perform better. For some excitation frequencies, liquid filters may prove useful (e.g. about 30x10^-4 M acenaphthene in methanol for excitation near 225 nm). Digital methods of background removal have been discussed in Chapters 11 and 12.
17.2.10. Noise characteristics

Extrinsic noise is noise arising from non-specific background signals even when no analyte is present while intrinsic noise is noise associated with the detection of the analyte itself (Winefordner et al., 1994). For Raman spectroscopy, the types of extrinsic noise are:

(i) Rayleigh scattering shot noise;
(ii) Rayleigh scattering flicker noise;
(iii) detector dark current;
(iv) detector shot noise (from dark current and Rayleigh scattering);
(v) detector readout noise.

Rayleigh scattering can be reduced with the use of optical filters, and dark current with detector cooling (see Chapter 3). Digital noise reduction methods have been discussed in Chapters 11 and 12.

Sources of intrinsic noise are the following:

(i) Raman scattering shot noise;
(ii) Raman scattering flicker noise;
(iii) detector shot noise (from Raman scattering);
(v) detector readout noise.

For all practical purposes, Raman measurements of analytes in a solvent are extrinsic noise limited, however, both extrinsic and intrinsic noise contribute to the overall noise. The complete characterization of all the sources of noise relevant to this probe would be another worthwhile future project since a knowledge of these would allow further improvements in sensitivity to be had.
17.2.11. Limit of detection

The limit of detection is generally defined as a SNR of 3 (see Winefordner et al., 1994). The current limit of detection for an aromatic molecule with the neuroprobe is about $10 \times 10^{-6}$ M (at 225 nm, 20 Hz, 20 μJ/pulse, 30 s exposure). However, the use of state-of-the-art equipment with the probes developed here could dramatically improve this detection limit to approximately $1 \times 10^{-9}$ M for aromatic neurotransmitters and $1 \times 10^{-6}$ M for aliphatic neurotransmitters using a continuous-wave deep ultraviolet laser (e.g. Russell et al., 1995). The sources of improvement are shown in Table 17.1.

<table>
<thead>
<tr>
<th>Improvement</th>
<th>Improvement factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increased fiber transmittance (cw laser)</td>
<td>4</td>
</tr>
<tr>
<td>Multiple-collection-fiber probe</td>
<td>5</td>
</tr>
<tr>
<td>Prism spectrometer</td>
<td>9</td>
</tr>
<tr>
<td>Increased average laser power (to 10 mW)</td>
<td>25</td>
</tr>
<tr>
<td>Total estimated improvement factor (cw)</td>
<td>4500</td>
</tr>
</tbody>
</table>

These do not include possible enhancements from high-performance detectors or as yet unforeseen improvements in the design of the optical system (e.g. better coupling efficiency) or better probes (e.g. using a higher numerical aperture collection fibers).

17.2.12. Dependencies

Biosensors may have various dependencies such as that on temperature, pressure, etc. The dependencies of the neuroprobe are mostly those associated with the transmission of pulsed UV radiation (e.g. pulse energy, pulse length, laser repetition rate, radiation frequency) and have been examined in some detail in Chapter 7. The temperature dependence of the probe was not investigated. For in vivo use, the temperature would be generally stable at 36.6 °C.
17.2.13. Biocompatibility

The biocompatibility of the probe glue was addressed in Chapter 8 and probe fouling in Chapter 10. The biocompatibility of the probe in toto was not examined.

17.2.14. Lifetime

The lifetime of current probes exceed well over 80 h of use. Factors affecting probe lifetime include solarization of the fiber, deterioration of the reflective coating of the fiber, deterioration of the fiber input and output endfaces due to the cumulative effect of laser microdamage, possible deterioration of the silastic glue holding the probe together. Other factors such as the lifetime of the laser dyes, lasing rods, laser flash lamps, and the detector also have a bearing on the total lifetime of the probe. These, however, are unlikely to be current limiting factors.

17.3. Future Work

It is abundantly clear that this method is nowhere near complete, however, it is likely to be viable, especially given the technical advances occurring apace in the field. These advances consist of improved lasers and detectors (Chase, 1994), and improved ultraviolet transmission light guides (e.g. Kubo, 1994). There are, however, several other areas related to the current set-up that could benefit from attention. These would constitute additional suggested future work and consist of improving the throughput of the system, filtering, and continued work on the development of digital signal processing methods.

17.3.1. Efficiency improvements

Two improvements could be made to increase the throughput of the system. These consist of modifying the optical filter and frequency doubling unit.
17.3.1.1. Double monochromator filter

It is clear from Chapter 3 that the throughput of the spectrometer and attached optical filter left much to be desired. The main losses in the system occurred at the holographic gratings and were measured to be around 75%. As a result, the optical filter could not be used since Raman signals were often too weak to be detected after filtering. When the single monochromator only was used, spectra had large backgrounds. Therefore, either a different filter has to be used or the throughput of the existing filter has to be improved.

UV frequency decoding could also be effected with a quartz prism and the manufacturer of the present optical filter supplies a prism predisperser that could be used for optical filtering (e.g. Russell et al., 1995). Because the throughput of the prism is about 95%, more if appropriately anti-reflection coated, the throughput of a prism optical filter could be about 90% compared to the current 7%. A logical extension of this argument leads to a prism spectrometer with prism filter and has been suggested. This could be combined with a tunable prism filter for laser plasma line removal (Sakuntala and Arora, 1993) or dichroic separation. The throughput of the spectrometer and filter system could be improved by perhaps as much as 2 orders of magnitude.

Alternatively, the use of a spatial-filter-equipped single monochromator for Raman spectroscopy has been reported (Trulson et al., 1994). Another way in which the throughput of a dispersing instrument could be improved, is by collecting light from unused orders and recombining them with the input via optical fibers, an arrangement that is likely to be less effective than a prism system.
17.3.1.2. Frequency doubler

The frequency doubler has a measured conversion efficiency of about 6% at 225 nm (Chapter 3), implying that about 90% of the input energy is wasted if a reflection loss of 0.5% per surface is taken into account. Clearly, recovering any amount of the wasted energy would improve the efficiency of the system. An attempt was made to recirculate the fundamental to the doubler input (slightly off-axis with regard to the input beam) with 1 front-surface mirror and 3 right-angle prisms. About 80% of the fundamental output beam was recirculated in this manner, but failed to produce a noticeable effect due to the rapid beam divergence. By incorporating a collimating lens and Galilean telescope, the beam could be collimated before redirecting it to the frequency doubler input. The use of high quality optical components (<0.5% reflectance per surface by 13 surfaces) could realize the recovery of about 94% of the fundamental output beam. It seems reasonable to expect an increase in efficiency of 2 to 3 times (or more) under these conditions.

17.3.2. Filtering improvements

Given the very low efficiency of Raman scattering, a high degree of rejection of elastically scattered light is essential (Chase, 1994). Improvements to the double monochromator optical filter have been discussed above. Two other practical, convenient, and economical methods of Rayleigh rejection and/or Raman amplification are discussed below.
17.3.2.1. Liquid filters

An ideal Raman filter should have a sharp cut-off just toward the red of the excitation frequency and negligible emission in the spectral region of interest after absorption of the excitation light (Chou et al., 1991). Some liquids possess this attribute and have been used as filters for Raman excitation at 355 and 337 (Chou et al., 1991).

Although, with some exceptions (e.g. naphthalene, anthracene), it is unlikely that such sharp-cut liquid filters could be found for Raman applications in the deep UV, a search of library absorption spectra of various compounds may nevertheless prove useful, even if filters suitable only for visible Raman excitation are uncovered.

17.3.2.2. Optical amplification

Above, it has been stated that an ideal Raman filter should have a sharp cut-off to the red of the excitation frequency and negligible emission in the spectral region of interest after absorption of the excitation light. One could consider non-negligible emission in the spectral region of interest after absorption of the excitation frequency to constitute interference. However, a second analysis reveals the possibility of harnessing such secondary emission, coupled with pumping, to amplify the Raman scattering. A two-fold advantage would then occur: reduction of the Rayleigh scattering and amplification of Raman scattering. A liquid used for this purpose should have a sharp absorption cut-off to the red of the excitation wavelength and a correspondingly sharp emission onset to the blue of the spectral region of interest. Some commercially available laser dyes fulfill this requirement for excitation of Raman scattering from NO$_3^-$ (1050 cm$^{-1}$) with different lines from an Ar$^+$ laser.

As with tunable lasers, amplified spontaneous emission (ASE) will be a problem with amplifying filters. A careful selection of the concentration of the amplifying filter dye may alleviate this problem. It is expected that several amplification stages should be
used, each with a different dye concentration. The initial stage should have a dye concentration that would give a number of dye molecules in the volume defined by the optical path comparable to the number of Raman photons in this volume at any one instant. These Raman photons would then be multiplied before entering the next stage of the amplifier. The same criterion should be applied to each successive amplification and filtering stage until the excitation photons are reduced to an acceptable level.

Incorporation of these dyes in plastics may stabilize the excited metastable state from which fluorescence occurs and hence minimize ASE. Such stabilization of organic dyes for flashlamp-excited stimulated emission has been demonstrated for polymethyl methacrylate (Peterson and Snavely, 1968) and may vary with the solidifying agent used. Although amplifying filters are not likely to be found for deep UV applications, the theoretical possibility for converting deep UV Raman scattering to visible wavelengths before amplification, exists.

17.3.3. Digital signal processing

Research to improve digital signal processing methods in order to extract the most information possible from a given spectrum, continues. This includes investigations of neural networks for signal extraction and mixture resolution. The ideal neural network to use with the Raman neuroprobe would generalize well from training data to noisy real data and discriminate well between closely related compounds. As mentioned in Chapters 11 and 12, the network's training and architecture influence generalization and discrimination. Future work could involve research into the use of new training methods, transfer functions, cost functions, and different connectivities.
17.3.3.1. Neural networks

The bearing of training methods and architectures on network function are some of the unresolved issues concerning neural networks (Baldi, 1995; Baldi and Hornik, 1995; Kappen, 1995). Some further investigation of these issues is therefore in order. In Chapter 11, the entire neural computational process has been shown to be equivalent to a series of matrix operations. Neural networks are normally trained by assigning random values to the network weights (weights 1 and 2 below). Because transfer 1 is known (often linear), [transformed values 1] are known by extension. All the known information before training commences is indicated in bold in the schematic representation below.

[Input layer values]

TRANSFER 1

[transformed values 1] x [weights 1]

= [hidden layer values]

TRANSFER 2

[transformed values 2] x [weights 2]

= [output layer values]

TRANSFER 3

[final network output (= desired output when trained)]

The unknown values are now simply calculated from the known ones by matrix multiplication. Because the weights were randomly assigned, the [final network output] will not equal the [desired network output] within the given tolerance. An inspection of the above schematic, however, reveals another way of training the network. This could be accomplished by assigning random values to [hidden layer values]. All the known information before training commences is again indicated in bold below (provided the inverse of the transfer functions is defined).
The unknown values are now the weight matrices and the problem reduces to the solution of the matrix equation $A.X = B$ for the unknown $X$. Methods such as singular value decomposition (SVD) and QR decomposition (QRD) (because the matrix $A$ can be decomposed into an orthogonal matrix $Q$ and an upper triangular matrix $R$) can be used for this purpose (Press et al., 1994). Existing methods of multidimensional optimization can be applied equally to optimizing the network weights (the conventional approach) or the hidden layer values (the approach suggested here).

Table 17.3 shows the results of two networks with sine transfer functions and no bias connections, trained by random assignment of the weights and by random assignment of the hidden layer values. Optimization with the simplex method (Nelder and Mead, 1965) was used in both cases. The training data consisted of the 10 neurotransmitter spectra and 3 of their linear combinations. The testing set consisted of 5 different linear combinations of the neurotransmitter spectra. The spectra were divided into 10 sections each and the average per section used as network input.
Table 17.2.
The results (after 2000 trials) of two networks with sine transfer functions trained by initial random assignment of either the weights (W) or the hidden layer values (H) followed by optimization with the simplex method.

<table>
<thead>
<tr>
<th>Architecture</th>
<th>RMS(W)</th>
<th>RMS(H)</th>
<th>time(W)</th>
<th>time(H)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10-1-10</td>
<td>0.314</td>
<td>0.326</td>
<td>156 s</td>
<td>227 s</td>
</tr>
<tr>
<td>10-2-10</td>
<td>0.288</td>
<td>0.320</td>
<td>150 s</td>
<td>378 s</td>
</tr>
<tr>
<td>10-3-10</td>
<td>0.298</td>
<td>0.236</td>
<td>123 s</td>
<td>569 s</td>
</tr>
<tr>
<td>10-4-10</td>
<td>0.287</td>
<td>0.221</td>
<td>140 s</td>
<td>774 s</td>
</tr>
<tr>
<td>10-5-10</td>
<td>0.307</td>
<td>0.210</td>
<td>139 s</td>
<td>1002 s</td>
</tr>
<tr>
<td>10-6-10</td>
<td>0.317</td>
<td>0.188</td>
<td>136 s</td>
<td>1261 s</td>
</tr>
</tbody>
</table>

Although SVD and QRD methods are computationally intensive and using them lengthened training as is evident from Table 17.3, the weights calculated were best fit weights given the input parameters and resulted in smaller RMS errors. Furthermore, some methods exist to update or speed up SVD and QRD calculations (Hansen, 1988; Hsieh et al., 1991; Moonen et al., 1991; Press et al., 1994).

The question arises as to the optimum hidden layer values and nodes to use in the random assignment. Since the function of the hidden layer is to project patterns from the input space into an intermediate space whence they can be easily separated by the output layer (Mao and Jain, 1995), one would wish the number of dimensions of the intermediate space to be as few as possible while permitting signal projection without information loss. Principal component analysis (PCA), which uses SVD, is a well-known optimal linear technique for signal extraction and data compression (Karhunen and Joutsensalo, 1995; Scharf, 1991) and can be used to separate the input space into signal and noise subspaces. This suggests using the number of dimensions of the signal subspace for the hidden layer space, in fact, the eigenvalues defining the signal subspace could be used for the initial hidden layer values instead of assigning them randomly (or they could be components obtained from other methods such as Fourier transforms, discrete cosine transforms, etc.).
From Table 17.3 it is clear that training using the SVD method produced good results and future work may consist of employing some of the acceleration methods to reduce training time. In addition, encoding schemes for the hidden layer and optimization of the hidden layer size should be investigated.

17.4. CONCLUSIONS

To conclude this work, it is appropriate to briefly summarize the objectives and the progress made toward attaining them. The purpose of this work was to develop a neurotransmitter sensor, in principle capable of measuring an unlimited number of neurotransmitters concurrently, in real time, and in vivo. The basic approach therefore has been to develop a probe with general capabilities. This created a considerable challenge (despite limiting the scope to the 10 small-molecule neurotransmitters), but the benefit of this approach has already been demonstrated: the neuroprobe of which the development has been detailed in this work, is capable of measuring not only neurotransmitters, but any chemical species producing a (resonance) Raman spectrum (e.g. steroids, proteins, alcohols, etc.). Furthermore, it could also be used with other analytical methods such as fluorescence and, with modifications, infrared absorption and coherent Raman spectroscopy. As such this probe is likely to find applications in a great number of fields where remote chemical analysis is required.

In order to realize the research objective, optical fiber probes have been fabricated for use with resonance Raman spectroscopy as analytical method. These represent a major advance in the field since it has been considered impossible to perform pulsed deep ultraviolet resonance Raman spectroscopy with optical fibers. This problem has been overcome by utilizing a recently developed optical fiber with relatively high transmittance of pulsed deep ultraviolet light and by designing novel optical fiber probes suitable for use under a variety of absorbing conditions. With the current experimental system, detection limits for aromatic amino acids are near 10 μM and with technically
more sophisticated equipment their detection limits can likely be lowered to 1 nM. The detection limits are also dependent on probe sensitivity and these are continually being improved.

Data processing methods for signal enhancement and identification have been developed as part of this work. In particular, artificial neural networks have been trained to identify specific neurotransmitters from their Raman and resonance Raman spectra obtained via optical fiber probe which constitutes a novel application of these algorithms.

Finally, these probes have been shown capable of measuring neurotransmitters in complex biological samples, including those mimicking in vivo conditions. Dopamine secreted by rat adrenal cancer cells in vitro has been detected and confirmed with HPLC analysis. The many problems that remain have been detailed throughout, nevertheless, a firm basis has been established by the work reported here to enable the continued and detailed development of this neuroprobe. With sophisticated ancillary instruments, as described at various points in the thesis, this neuroprobe will prove a valuable addition to the number of neurotransmitter sensors used to investigate the functioning of the brain, and in addition, to the array of existing analytical instruments for other applications.

17.5. REFERENCES.


APPENDIX A

Probe simulation program in MicroSoft Quick Basic 4.0

'parameter list
'cll  cross line length, i.e. divergence due to expanding light cone
'clv  cross line value, value of distribution at distance from fiber tip
'psn  position of grid crossing taking expanding cone into account
'vel  value of light intensity at psn
'x    x position in grid where light intensity has a specific value
'y    y position
'z    symmetry x position in grid where light intensity has a specific value
'absorb  absorption factor of solution
'clcr cross line core right, divergence cone for collection fiber
'sep separation between fibers
'geo excitation tip geometry
'marg margin for graphics
'sp$  spacer for printing to screen

DIM psn(11, 401), vle(ll, 401), cll(401), clv(401), x(401), y(401), z(401)
DIM distr(ll)
pi = 3.142: sp$ = " "
CLS

PRINT "This program calculates the exit profile of light from a fiber"
PRINT "with different core diameters, different propagation profiles, and"
PRINT "in media with different absorptivities."
PRINT
PRINT "Some parameters can be set by the user."
PRINT
INPUT "Press ENTER to continue..."; l$
CLS

INPUT "Enter excitation core diameter"; corel
INPUT "Enter excitation tip geometry (1=flush/2=angled)"; geo
INPUT "Enter collection core diameter"; corer
INPUT "Enter separation distance between fibers"; sep
PRINT
INPUT "Enter fiber divergence in degrees (half angle)"; na
PRINT
PRINT "Enter number of intensity gradations to be plotted,"
INPUT "e.g. 10 will cause light intensities to be plotted at 10% intervals"; level
PRINT
INPUT "Enter % absorptivity of solution (0-100)"; absorb
PRINT
INPUT "Enter beam profile (Gaussian = 1/Other = 2)"; profile
PRINT
IF profile = 1 THEN GOSUB gauss
marg = INT(640 - (corel + sep + corer)) / 2
10
'calcs for excitation fiber
FOR n = 0 TO 400
    cll(n) = corel / 2 + n * TAN(pi * na / 180)
    clcr = corer / 2 + 400 * TAN(pi * na / 180) 'collection fiber cone
    'PRINT cll(n): INPUT ; t$
NEXT n

clv(0) = 1
FOR n = 1 TO 400
    clv(n) = ((cll(0))^2 / (cll(n))^2)
    'IF n = 1 THEN PRINT cll(n), clv(n)
    'IF n = 400 THEN PRINT cll(n), clv(n): INPUT ; t$
    'PRINT clv(n): INPUT ; t$
NEXT n

FOR n = 0 TO 400
    FOR m = 0 TO 10
        psn(m, n) = m * cll(n) / 10
        'PRINT m, n, psn(m, n): INPUT ; t$
        vle(m, n) = clv(n) * distr(m) * (100 - absorb) / 100
        'PRINT m, n, vle(m, n): INPUT ; t$
    NEXT m
NEXT n

'graphics
SCREEN 12

IF geo = 2 THEN GOTO 500
400 'tip geometry flush draw fiber sides
FOR n = 400 TO 459
    PSET (marg, n): PSET (marg + corel, n): PSET (marg + corel + sep, n): PSET (marg + corel + sep + corer, n)
NEXT n
GOTO 600

500 'tip angled draw fiber sides
LINE (marg, 400)-(marg, 459): LINE (marg + corel, 400 - corel)-(marg + corel, 400), 12:
LINE (marg + corel, 400)-(marg + corel, 459)
LINE (marg + corel + sep, 400)-(marg + corel + sep, 459): LINE (marg + corel + sep + corer, 400)-(marg + corel + sep + corer, 459)
600 iso = 0: 1 = 1
    IF geo = 1 THEN LINE (marg, 400)-(marg + corel, 400), 12: LINE (marg + corel + sep, 400)-(marg + corel + sep + corer, 400), 10
    IF geo = 2 THEN LINE (marg, 400)-(marg + corel, 400 - corel): LINE (marg + corel + sep, 400)-(marg + corel + sep + corer, 400), 10
FOR k = 1 TO level
    iso = 1 - k / level
    1 = 1 + level / 10
    IF geo = 1 THEN x(0) = marg + corel / 2: y(0) = 400: z(0) = marg + corel / 2
    IF geo = 2 THEN x(0) = 400 - corel / 2: y(0) = marg + corel: z(0) = 400 - corel / 2
    FOR n = 1 TO 400
        FOR m = 0 TO 9
IF iso < vle(m, n) AND iso > vle(m + 1, n) THEN 100 ELSE 200

100 IF geo = 1 THEN x(n) = marg + corel / 2 + psn(m, n) + (psn(m + 1, n) - psn(m, n)) * ((vle(m, n) - iso) / (vle(m, n) - vle(m + 1, n)))
    IF geo = 2 THEN x(n) = 400 - corel / 2 + psn(m, n) + (psn(m + 1, n) - psn(m, n)) * ((vle(m, n) - iso) / (vle(m, n) - vle(m + 1, n)))
    IF geo = 1 THEN z(n) = marg + corel / 2 - (psn(m, n) + (psn(m + 1, n) - psn(m, n)) * ((vle(m, n) - iso) / (vle(m, n) - vle(m + 1, n))))
    IF geo = 2 THEN z(n) = 400 - corel / 2 - (psn(m, n) + (psn(m + 1, n) - psn(m, n)) * ((vle(m, n) - iso) / (vle(m, n) - vle(m + 1, n))))
    IF geo = 1 THEN y(n) = 400 - n
    IF geo = 2 THEN y(n) = marg + corel + n
't locate: PRINT n, iso, x(n), y(n): INPUT ; t$

IF x(n) = 0 THEN 300
IF geo = 1 THEN PSET (x(n), y(n)), 12
IF geo = 2 THEN
    IF POINT(y(n), x(n)) = 10 THEN PSET (y(n), x(n)), 14 ELSE PSET (y(n), x(n)), 12

END IF

200 NEXT m

300 NEXT n

NEXT k

FOR n = 1 TO 400
    IF geo = 1 THEN PSET (marg + corel / 2 - cll(n), n), 12
    IF geo = 2 THEN PSET (n + marg + corel, 400 - corel / 2 - cll(n)), 12
    IF geo = 1 THEN PSET (marg + corel / 2 + cll(n), n), 12
    IF geo = 2 THEN PSET (n + marg + corel, 400 - corel / 2 + cll(n)), 12
NEXT n

'collection fiber calcs

FOR n = 1 TO 400
    IF POINT(marg + sep + corel + corer / 2 - (corer / 2 + n * TAN(pi * na / 180)), 400 - n) = 12 THEN
        PSET (marg + sep + corel + corer / 2 - (corer / 2 + n * TAN(pi * na / 180)), 400 - n), 14
    ELSE

    END IF

NEXT n
PSET (marg + sep + corel + corer / 2 - (corer / 2 + n * TAN(pi * na / 180)), 400 - n), 10
END IF
IF POINT(marg + sep + corel + corer / 2 + (corer / 2 + n * TAN(pi * na / 180)), 400 - n) = 12 THEN
    PSET (marg + sep + corel + corer / 2 + (corer / 2 + n * TAN(pi * na / 180)), 400 - n), 14
ELSE
    PSET (marg + sep + corel + corer / 2 + (corer / 2 + n * TAN(pi * na / 180)), 400 - n), 10
END IF
NEXT n

'collection efficiency calcs
IF geo = 2 THEN GOTO 2000
excit = 0: coll = 0: fibl = 0: fib2 = 0: max = 0
FOR n = 0 TO 400
    fibl = 0: fib2 = 0:
    FOR m = 0 TO 639
        IF POINT(m, n) = 12 OR POINT(m, n) = 14 THEN fibl = fibl + 1
        IF POINT(m, n) = 10 OR POINT(m, n) = 14 THEN fib2 = fib2 + 1
        LOCATE 26: PRINT fibl, fib2
        max = fibl
    NEXT m
    fibl = 0: fib2 = 0:
    FOR m = 0 TO 639
        IF POINT(m, n) = 12 OR POINT(m, n) = 14 THEN fibl = fibl + 1
        IF POINT(m, n) = 10 OR POINT(m, n) = 14 THEN fib2 = fib2 + 1
        IF fibl > 0 AND fibl <= max / 2 THEN excit = excit + fibl
        IF fibl > max / 2 AND fibl <= max THEN excit = excit + max - fibl
        IF fib2 = 1 AND fibl > 0 AND fibl <= max / 2 THEN coll = coll + fibl
        IF fib2 = 1 AND fibl > max / 2 AND fibl <= max THEN coll = coll + max - fibl
        LOCATE 27: PRINT "excit="; excit; "coll="; coll; fibl; fib2
    NEXT m
NEXT n

INPUT "Enter to exit!"; IS$ END

2000
excit = 0: coll = 0: fibl = 0: fib2 = 0: max = 0
FOR m = marg + corel + 2 TO 639
    fibl = 0: fib2 = 0:
    FOR n = 439 TO 0 STEP -1
        IF POINT(m, n) = 12 OR POINT(m, n) = 14 THEN fibl = fibl + 1
        IF POINT(m, n) = 10 THEN fib2 = fib2 + 1
        IF POINT(m, n) = 10 AND POINT(m, n - 1) = 10 THEN fib2 = fib2 - 1
        IF POINT(m, n) = 10 AND POINT(m, n - 1) = 14 THEN fib2 = fib2 - 1
        'LOCATE 26: PRINT fibl, fib2
        max = fibl
fib1 = 0; fib2 = 0:
FOR n = 439 TO 0 STEP -1
    IF POINT(m, n) = 12 OR POINT(m, n) = 14 THEN fib1 = fib1 + 1
    IF POINT(m, n) = 10 THEN fib2 = fib2 + 1
    IF POINT(m, n) = 10 AND POINT(m, n - 1) = 10 THEN fib2 = fib2 - 1
    IF POINT(m, n) = 10 AND POINT(m, n - 1) = 14 THEN fib2 = fib2 - 1
    IF fib1 > 0 AND fib1 <= max / 2 THEN excit = excit + fib1
    IF fib1 > max / 2 AND fib1 <= max THEN excit = excit + max - fib1
    IF fib2 = 1 AND fib1 > 0 AND fib1 <= max / 2 THEN coll = coll + fib1
    IF fib2 = 1 AND fib1 > max / 2 AND fib1 <= max THEN coll = coll + max -
    fib1
LOCATE 27: PRINT "excit=", excit;  "coll="; coll; fib1; fib2
    IF excit > 0 THEN PRINT coll / excit, sp$
NEXT n
    IF fib1 / 2 - INT(fib1 / 2) > 0 THEN INPUT ; t$
NEXT m
INPUT "Enter to exit!"; IS
END

gauss:
INPUT "Enter distribution mean (0)"; mean
INPUT "Enter distribution std. deviation"; std
FOR n = 0 TO 10
    distr(n) = EXP(-(n - mean) / std) ^ 2)
NEXT n
PRINT
RETURN
APPENDIX B

Program to perform randomly displaced autoaccumulations of a spectrum, written in Microsoft Quick Basic 4.0.

'******--variable list--*****

xs(xxx) intermediate array
x1(xxx) input data values (dependent variable) file #1
x2(xxx) input data values (dependent variable) file #2
x3(xxx) input data values (dependent variable) file #3
xo(xxx) output data value (dependent variable)
xb(xxx) averaged input data for baseline correction (dependent variable)
w(xxx) input data values (independent variable)
sm(xxx) calculated product (output values)
m1,m2 start/end values of data region to process
m3 number of points to interpolate between adjacent input data values
i number of input data points
m4 number of accumulations of interpolated data
m5 accumulation bandwidth in terms of original input
m6 random symmetrical shift used for accumulations
m6x m6 components for generating a normal distribution
b$ indicates whether normal or uniform distribution wanted
b1$ indicates whether baseline correction wanted
b1 moving average window size
b2 number of iterations for moving average (normally 1)
mavg moving average window sum
resp response function

'**************--screen introduction--**************

CLS
PRINT "This program accumulates a number of the same or independent series."
PRINT "The series (1500 points max.) are displaced by a random amount (within the bandwidth)"
PRINT "from the original before accumulation starts."
PRINT "The intention is to simulate the accumulation of a number of"
PRINT "independent spectra of the same sample, but instead using only"
PRINT "a few (1-3) spectra."
PRINT ""
PRINT "If desired, a baseline correction can be made by subtracting a"
PRINT "moving average based on the original values from the calculated output."
PRINT ""

'**************--initialization statements--**************

DIM x1(1600)
DIM x2(1600)
DIM x3(1600)
DIM xb(1600)
DIM xo(1600)
DIM xs(1600)
DIM sm(1600)
DIM w(1600)
DIM resp(100)
c = 0
n3 = 0

3 '***************--Read input data file--***************
INPUT "Enter the number of files to read (max=3):"; n3

FOR n = 1 TO n3
    GOSUB 1000
NEXT n

'***************--prompts for parameter inputs--***************
5 n1 = 0: n2 = 0
'INPUT "Enter data region to process as STARTVALUE,ENDVALUE:"; m1, m2
INPUT "Enter number of accumulations:"; m4
INPUT "Enter the bandwidth for accumulation:"; m5
INPUT "Enter bandwidth type as uniform (U) or normal (N):"; b$
INPUT "Baseline to be corrected (y/n):"; b1$
b1 = 0: b2 = 0
IF b1$ = "Y" OR b1$ = "y" THEN INPUT "Enter moving average window size:"; b1
IF b1$ = "Y" OR b1$ = "y" THEN INPUT "Enter moving average # of iterations:"; b2
CLS

'***************--main program--***************
max = 0: x1 = 0: x2 = 0: y1 = 0: y2 = 0: allsum = 0: allavg = 0
FOR n = 1 TO i
    IF x1(n) > max THEN max = x1(n)
    IF x2(n) > max THEN max = x2(n)
    IF x3(n) > max THEN max = x3(n)
    allsum = allsum + x1(n)
NEXT n
allavg = allsum / i

8 IF b$ = "U" OR b$ = "u" THEN GOSUB 10 ELSE GOSUB 20 'random summation
GOSUB 3000 'plot results
IF n1 < m4 THEN GOTO 8
21 GOSUB 30 'average accumulations
IF b$ = "Y" OR b$ = "y" THEN GOSUB 500 'remove baseline
GOSUB 2000 'output results to file
GOSUB 3000 'plot results
GOSUB 4000 'another run?
END

'***************--summation routine--***************

'random summation (uniform distribution)
10
FOR n = 1 TO 10  'plot results every 10 iterations

RANDOMIZE TIMER
  n4 = INT((n3) * RND + 1)
  m6 = (RND * m5)
  m6 = INT(m6 - (m5 / 2))
  FOR k = 0 TO m5  'response function
    IF INT(m6 + (m5 / 2)) = k THEN resp(k) = resp(k) + 1: LOCATE 2: PRINT k, resp(k)
  NEXT k

FOR m = 1 + INT(m5 / 2) TO i - INT(m5 / 2)
  RANDOMIZE TIMER
  IF n4 = 1 THEN sm(m) = sm(m) + xl(m + m6)' + RND * allavg
  IF n4 = 2 THEN sm(m) = sm(m) + x2(m + m6)' + RND * allavg
  IF n4 = 3 THEN sm(m) = sm(m) + x3(m + m6)' + RND * allavg
  xo(m) = sm(m): IF n1 / 10 = INT(n1 / 10) THEN xs(m) = xs(m) + xo(m): sm(m) = 0
  'PRINT m, m6, m + m6, xi(m + m6), sm(m)
  NEXT m

LOCATE 1: PRINT n1
  n1 = n1 + 1: IF n1 > m4 THEN RETURN 21
  NEXT n

RETURN

'random summation (normal distribution)
20:
FOR n = 1 TO 10  'plot results every 10 iterations
  RANDOMIZE TIMER
  n4 = INT((n3) * RND + 1)
  m61 = (RND * m5)
  m62 = (RND * m5)
  m63 = (RND * m5)
  m64 = (RND * m5)
  m65 = (RND * m5)
  m66 = (RND * m5)
  m6 = (m61 + m62 + m63 + m64 + m65 + m66) / 6
  m7 = INT(m6 - (m5 / 2))
  FOR k = 0 TO m5
    IF INT(m6) = k THEN resp(k) = resp(k) + 1: LOCATE 2: PRINT k, resp(k)
    NEXT k
  FOR m = 1 + INT(m5 / 2) TO i - INT(m5 / 2)
    RANDOMIZE TIMER
    m6 = (m6 - m66 + (RND * m5)) / m5
IF n4 = 1 THEN sm(m) = sm(m) + x1(m + m7)' + m6 * allavg
IF n4 = 2 THEN sm(m) = sm(m) + x2(m + m7)' + m6 * allavg
IF n4 = 3 THEN sm(m) = sm(m) + x3(m + m7)' + m6 * allavg
xo(m) = sm(m): IF n1 / 10 = INT(n1 / 10) THEN xs(m) = xs(m) + xo(m): sm(m) = 0
'PRINT m, m6, m + m6, xi(m + m6), sm(m)
NEXT m
LOCATE 1: PRINT n1
n1 = n1 + 1: IF n1 > m4 THEN RETURN 21
NEXT n
RETURN

'get average from accumulations
30:
FOR n = 1 TO 1600
   xs(n) = xs(n) / m4: 'PRINT xs(n)
exo(n) = xs(n)
   'resp(n) = resp(n) / m4
NEXT n
RETURN

**********-baseline correction subroutine--**********
500:
b3 = INT(b1 / 2): mavg = 0
FOR l = 1 TO b2
   FOR n = b3 TO i - b3
      FOR m = 1 TO b1
         mavg = mavg + xo(n - b3 + m)
      NEXT m
      xb(n) = mavg / b1: mavg = 0
   NEXT n
NEXT l
600
FOR n = 1 TO i
   xo(n) = xo(n) - xb(n)
NEXT n
RETURN

**********---Input subroutine--**********
1000:
PRINT " "
1002 PRINT "Enter path and name of file # (X AND Y values, 3 decimals)"; n
   INPUT "to process"; f$
1008 OPEN f$ FOR INPUT AS #1
1010 i = 1
1020 DO WHILE NOT EOF(1)
1030   INPUT #1, w(i), x3(i)
xo(i) = x3(i); xb(i) = x3(i)
IF n = 1 THEN x1(i) = x3(i)
IF n = 2 THEN x2(i) = x3(i)
x3(i) = 0
i = i + 1
1040 LOOP
1050 CLOSE #1
i = i - 1
1060 RETURN

'***************--Output subroutine--***************
2000:
PRINT ""
2001 INPUT "Do you want to save the results (y/n):"; a$
2003 IF a$ = "N" OR a$ = "n" THEN GOTO 2060
2007 INPUT "Enter name of file to store accumulation data"; f$
2008 OPEN f$ FOR OUTPUT AS #1
2010
2020 FOR n = 1 TO i
2030 PRINT #1, w(n), xo(n)
2040 NEXT n
2050 CLOSE #1

INPUT "Enter name of file to store response function"; f$
OPEN f$ FOR OUTPUT AS #1
FOR n = 0 TO m5
   IF n < INT((m5 / 2) + 1) THEN PRINT #1, n, resp(LNT(m5 / 2) + n) ELSE PRINT #1, n, resp(n - INT(m5 / 2))
   NEXT n
C1SE #1
2060 RETURN

'***************--plot results--***************
3000:
CLS
SCREEN 12

max = 0: x1 = 0: x2 = 0: y1 = 0: y2 = 0
FOR n = 1 TO 640
   PRINT w(n), x1(n): INPUT t$
   IF x1(n) > max THEN max = x1(n)
   IF xs(n) / n1 > max THEN max = xs(n) / n1
   NEXT n
scale = 460 / max

PALETTE
FOR n = 1 TO 640
   x1 = n

y1 = 480 - INT(scale * xs(n) / n1)
y3 = (480 - INT(scale * (x1(n)))) - (.3 * max))

PSET (n, 479)
LINE (x2, y2)-(x1, y1)
LINE (x2, y4)-(x1, y3), 2
x2 = x1: y2 = y1: y4 = y3
IF x2 < 0 THEN x2 = 0: IF y2 < 0 THEN y2 = 0
NEXT n

RETURN

4000:
INPUT "Another run"; i$
SCREEN 0
IF i$ = "Y" OR i$ = "y" THEN CLS ELSE END
INPUT "Use the same files:"; i$
IF i$ = "Y" OR i$ = "y" THEN RETURN 5 ELSE RETURN 3
APPENDIX C

Program to amplify signals and reduce noise in a spectrum using a moving product method. Written in MicroSoft Quick Basic 4.0.

CLS
PRINT "This program calculates a moving average and/or product"
PRINT "and flattens the baseline if required."
PRINT "Most of the parameters can be changed by the user."
PRINT "It is suggested that the user experiment with different"
PRINT "parameter settings for optimal results."
PRINT "Suggested starting values are given (in parentheses)."
PRINT "Maximum file size = 2000 points."

PRINT ""

DIM sm(2000) AS DOUBLE
DIM xr(2000) AS DOUBLE
DIM bc(2000) AS DOUBLE
DIM pr(2000) AS DOUBLE
DIM w(2000) AS DOUBLE

C = 0

INPUT "Enter moving window size count:"; m
IF m / 2 > INT(m / 2) THEN k = INT(m / 2) + 1 ELSE k = m / 2
PRINT ""
INPUT "Do you want a moving average (a), product (p) or both (b):"; t$

'READ FILE
GOSUB 1000

5

'moving product/average
max = -32000: min = 10000: CLS : PRINT "Calculating...": PRINT ""
20 FOR n = k TO count - k
30 sum = 0: prod = 1: LOCATE 3: PRINT "run =": n
    IF xr(n) > max THEN max = xr(n)
    IF xr(n) > 0 AND xr(n) < min THEN min = xr(n)
60 FOR 1 = 1 TO m
70 prod = prod * ((xr(n + 1 - k) / max) + 1)
PRINT prod: ' INPUT TT$
75 sum = sum + ((xr(n + 1 - k) / max) + 1)
80 NEXT 1
90 sm(n) = sum / m
92 pr(n) = prod ^ 1 / m
95
100 'PRINT sm(n), pr(n)
110 NEXT n
'GOSUB 6000  'smooth pr(n)

'baseline correction
PRINT " "
INPUT "Do you require baseline correction (y/n):"; t1$
IF t1$ = "n" OR t1$ = "N" THEN GOTO 2000
CLS : PRINT "Calculating..." : PRINT " "
max = 0: min = 10000
FOR n = k TO count - k
    LOCATE 3: PRINT "run = "; n
    IF sm(n) = 1 THEN GOTO 2000
    IF sm(n) > max THEN max = sm(n)
    IF sm(n) > 0 AND sm(n) < min THEN min = sm(n)
    bc(n) = sm(n) - pr(n)
    PRINT pr(n), sm(n), bc(n)
NEXT n
GOTO 2000

'Input subroutine
n = 0
1000 PRINT " "
1002 INPUT "Enter path and name of file to process"; f$
1004 CLS
1008 OPEN f$ FOR INPUT AS #1
1010 PRINT "Reading file..."
1020 DO WHILE NOT EOF(1)  
     n = n + 1
1030 INPUT #1, w(n), xr(n)
1040 LOOP
1050 CLOSE #1
    CLS
    count = n
1060 RETURN

2000 'Output subroutine
CLS : INPUT "Do you want to save the data (y/n):"; a$
IF a$ = "N" OR a$ = "n" THEN GOTO 3000
PRINT ""
2002 INPUT "Enter name of file to store data"; f$
2008 OPEN f$ FOR OUTPUT AS #1
2010 2020 FOR n = 1 TO 2000
2030    IF t$ = "a" OR t$ = "A" THEN WRITE #1, sm(n)
    IF t$ = "p" OR t$ = "P" THEN WRITE #1, pr(n)
    IF t$ = "b" OR t$ = "B" THEN WRITE #1, sm(n), pr(n)
2040 NEXT n
2050 CLOSE #1
2060
3000 'plot data
CLS
SCREEN 11
max = 0: x1 = 0: x2 = 0: y2 = 480: y1 = 0: x3 = 0: x4 = 0: y3 = 0: y4 = 0
min = 32000
FOR n = k TO 680 - k
    IF bc(n) > max THEN max = bc(n)
    IF pr(n) > max THEN max = pr(n)
    IF sm(n) > max THEN max = sm(n)
    IF bc(n) > 0 AND bc(n) < min THEN min = bc(n)
    IF pr(n) > 0 AND pr(n) < min THEN min = pr(n)
    IF sm(n) > 0 AND sm(n) < min THEN min = sm(n)
NEXT n
scale = 460 / (max - min)

FOR l = k TO 638
    FOR m = 1 TO 480
        y1 = 480 - (scale * (pr(l) - min)): x1 = 1: y3 = 480 - (scale * (pr(l) - min)): x3 = 1
        PSET (x1, y1): ' PSET (x3, y3), 3
        LINE (x2, y2)-(x1, y1): ' LINE (x4, y4)-(x3, y3), 3
        x2 = x1: y2 = y1: x4 = x3: y4 = y3
    NEXT m
NEXT l

INPUT "Another run"; i$
IF i$ = "Y" OR i$ = "y" THEN CLS : GOTO 3
END

6000 FOR n = m TO 2000 - m
    sum = 0: prod = 1: LOCATE 3: PRINT "run =": n
    FOR l = 1 TO 2 * m
        sum = sum + pr(n + 1 - m)
    NEXT l
    pr(n) = sum / (2 * m)
NEXT n
RETURN