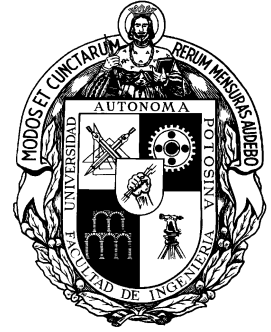




UNIVERSIDAD AUTÓNOMA DE SAN LUÍS POTOSÍ



FACULTAD DE INGENIERÍA ELÉCTRICA

“USO DE LAS ESPECTROSCOPIAS ÓPTICAS Y MÉTODOS MULTIVARIANTES APLICADOS AL ANÁLISIS DE MUESTRAS BIOLÓGICAS”

Tesis que presenta:

M.I. Raquel Avila Rodríguez

Para obtener el grado de
Doctor en Ingeniería Eléctrica
En la especialidad de Materiales y
Dispositivos Optoelectrónicos.

Asesores de Tesis:

Dr. Facundo Ruíz

Dr. Cuauhtémoc Araujo Andrade

San Luis Potosí, S.L.P., Mayo de 2010

DEDICATORIA

A Dios por ser la fuente de mi vida, la luz que me ilumina , la alegría de mi existencia, mi fortaleza, mi todo.....

Por regalarme la más sublime y excelsa de las bendiciones, al darme la vida a través de mis amados Padres.

A mi Preciosa Mamá: Ma. Teresa Rodríguez González, porque su gran Amor está vivo en cada uno de sus hijos.

A mi Papá: José Guadalupe Avila Ruíz, Amoroso Padre que con su excelente ejemplo me enseñó a luchar por mis sueños.

A mis hermanos y hermanas: Miguel Avila Rodríguez, Javier Avila Rodríguez, José Guadalupe Avila Rodríguez, Elizabeth Avila Rodríguez, Rosa Velia Avila Rodríguez, Joel T. Pérez palomares y Yolanda Avila Rodríguez, son una bendición para mi vida.

A Andrea Teresa Pérez Avila.

Y a todos mis sobrinos: Joel David, Eli, Hana, Pepe, Uli, Vero, Miguel, Juan de Dios, German, Caro, Tere.

AGRADECIMIENTOS

A mis asesores:

Dr. Facundo Ruíz, Dr. Cuauhtémoc Araujo Andrade.

Al Dr. José Refugio Martínez Mendoza.

Al Dr. Palomares, al Dr. Gerardo, y al Dr. Gabriel.

A todos los integrantes del Cuerpo Académico de Materiales de la Facultad de Ciencias de la UASLP.

A todas mis amistades que me apoyaron en este proceso.

Contenido General

PARTE PRELIMINAR

DEDICATORIAS Y AGRADECIMIENTOS

ÍNDICE

INTRODUCCIÓN	Página
Estructura de la tesis.....	8
Antecedentes.....	9
Justificación	12
Hipótesis	12
Objetivo General.....	12
Objetivos Específicos.....	12
Importancia del estudio	13
Límites del estudio	13
Definición de términos.....	13
1. ESPECTROSCOPIAS Y MÉTODOS MULTIVARIANTES: FUNDAMENTOS TEÓRICOS	
1.1 Espectroscopias ópticas.....	17
1.2 Espectroscopia Raman	24
1.3 Espectroscopia infrarroja	29
1.4 Análisis multivariante	41
1.4.1 Análisis de Componentes Principales (PCA).....	42
1.4.2 Mínimos Cuadrados Parciales (PLS).....	45
1.4.3 Preprocesamiento de datos	53

2. METODOLOGÍA DE PREPARACIÓN DE MUESTRAS Y MEDICIONES ESPECTROSCÓPICAS.

2.1 Instrumentación NIR.....	62
2.2 Instrumentación Raman.....	63
2.3 Mediciones con NIR para la determinación del error.....	65
2.4 Mediciones Raman para la detección de anticuerpos.....	66
2.5 Mediciones Raman para el análisis del Biofilm.....	67

3. ANÁLISIS MULTIVARIANTE Y DISCUSIÓN DE RESULTADOS

3.1 Determinación del error “ <i>in vitro</i> ” con espectroscopia NIR.....	70
3.2 Utilización de la espectroscopia Raman y análisis multivariante para la detección de anticuerpos.....	81
3.3 Espectroscopía Raman para el análisis del Biofilm.....	93

CONCLUSIONES

PERSPECTIVAS DEL TRABAJO

REFERENCIAS BIBLIOGRÁFICAS

ANEXOS

INTRODUCCIÓN

I INTRODUCCIÓN

En este trabajo de tesis se muestran los resultados del estudio de las espectroscopías ópticas, tales como la espectroscopia de absorción en el cercano a Infrarrojo (NIR, Near Infrared) y la espectroscopia de dispersión Raman, así como la aplicación de los métodos multivariantes, para la caracterización, análisis e interpretación de los datos espectroscópicos obtenidos de muestras biológicas.

La identificación del origen de alguna enfermedad causada por microorganismos, es de gran relevancia para el diagnóstico de diversas enfermedades, los métodos actuales necesitan de agentes reactivos, de alguna preparación especial de la muestra, y de tiempo para analizar dependiendo del tipo de estudio, por ejemplo, las técnicas clásicas de análisis microbiológico para la identificación de bacterias causantes de gastroenteritis puede llevar 4 días, lo cual podría tener serias consecuencias en niños, ancianos y pacientes con una gastroenteritis severa, es por ello que la investigación de diversos métodos y técnicas para identificar microorganismos causantes de una enfermedad específica está cobrando importancia [1].

Actualmente los niveles de glucosa se obtienen mediante métodos electroquímicos o enzimáticos, en los cuales es necesario el uso de agentes reactivos y la extracción de la muestra a analizar puede ser: sangre, orina, saliva, sudor, entre otros [2-3]. Sin embargo, en los últimos años, se han propuesto diversas técnicas menos invasivas como por ejemplo, la implantación subcutánea de biosensores [4], iontoforesis y el monitoreo no invasivo de glucosa en sangre y tejido, utilizando técnicas espectroscópicas tales como absorción en el cercano infrarrojo, Raman, foto acústica y fluorescencia [5-8], también se ha demostrado que las técnicas espectroscópicas, aunadas a las técnicas de análisis multivariante, en particular mínimos cuadrados parciales (PLS), son técnicas confiables para identificar y cuantificar analitos, en medios turbios como la sangre.

Las técnicas espectroscópicas han sido promovidas por diversos laboratorios y compañías, sin embargo, presentan algunos errores de precisión y exactitud en los resultados. Uno de los objetivos de la primera sección esta de tesis va enfocado a encontrar el error inducido en la predicción de glucosa debido a variaciones en la temperatura.

Otra aplicación de esta tesis está relacionada con la detección de anticuerpos en fluidos biológicos, específicamente en esta sección de la tesis se analizó los anticuerpos para el *Toxoplasma Gondii*, utilizando espectroscopia de dispersión Raman y análisis multivariante. Nuestro organismo cuenta con anticuerpos o inmunoglobulinas (ACs), que son moléculas de proteína, cuya función es identificar a las sustancias extrañas y microorganismos que invaden nuestro cuerpo y eliminarlas a través de la acción de otras moléculas. En el humano, existen 5 clases de Anticuerpos (ACs) o inmunoglobulinas (Igs) que son: IgG, IgD, IgG, IgA, e IgE. Actualmente la detección se realiza con pruebas serológicas, como la prueba de fijación de complemento, la prueba de ELISA (Enzyme-Linked ImmunoSorbent Assay, Ensayo por inmunoabsorción ligado a enzimas) [9], la prueba de Sabin-Feldman [10], la prueba de inmunofluorescencia [11], la prueba de aglutinamiento diferencial [12].

En esta tesis, utilizamos la espectroscopía Raman como una herramienta que puede brindarnos información de la composición bioquímica de muestras biológicas de tal forma que es posible obtener información de los anticuerpos presentes en las muestras, utilizando técnicas de análisis multivariante.

Otra de las aplicaciones de las espectroscopias ópticas plasmadas en esta tesis es la caracterización del biofilm producido por la bacteria *E. faecalis*, utilizando espectroscopía Raman. El biofilm es una población de células que crecen unidas a una superficie envuelta en una matriz de exopolisacáridos [13-15] que la protege del ataque de antibióticos [16-18]. Se ha reportado que más del 60% de las infecciones microbianas son causadas por el biofilm.

En Endodoncia, la bacteria *E. faecali* ha demostrado tener la habilidad de invadir los túbulos dentinarios [19], mientras otras bacterias no poseen esa característica, además, la bacteria *E. faecalis* tiene la capacidad de formar biofilm como factor de supervivencia y de esta forma infectar crónicamente los conductos radiculares [20-23], por ello la importancia de estudiar este microorganismo cuando se presenta en una comunidad tan especializada como el biofilm, hasta el momento la espectroscopía Raman en el área de odontología solo se ha utilizado para el análisis de tejido óseo y componentes de calcio como la hidroxiapatita [24], para analizar resinas adhesivas dentales [25], y la composición de hidróxido de calcio posterior a una mediación intraconducto. En esta sección de la tesis, utilizamos la espectroscopia de dispersión Raman para la caracterización del biofilm producido por la *E. faecalis*, lo cual no se ha hecho antes, y poder conocer su composición bioquímica, para posteriormente por medios enzimáticos deshacer ese biomaterial y dejar actuar a los antibióticos.

Estructura de la Tesis

La introducción de la sección de esta tesis describe de manera general los objetivos del trabajo, los fundamentos teóricos, el problema específico y la estrategia de investigación llevada a cabo.

Después de la introducción, la primera sección corresponde a la descripción y análisis de las investigaciones realizadas con las espectroscopías NIR y Raman en análisis de fluidos biológicos, incluye también el marco teórico de referencia que orientó al diseño de la investigación y a la interpretación de los resultados.

La segunda sección describe el proceso de investigación, el material e instrumentación utilizados, el procedimiento de análisis de datos, las técnicas utilizadas para el procesamiento de datos, pruebas de hipótesis paramétricas o no paramétricas.

En la tercera sección se presentan los resultados obtenidos con el uso de las espectroscopias ópticas y el análisis multivariante para cada una de las aplicaciones con muestras biológicas.

En la última sección se presentan las conclusiones de esta tesis.

Antecedentes

El mundo es energía en transformación que existe desde el origen del Universo. Las distintas formas de energía interactúan entre sí y con la materia, produciendo diversas alteraciones gracias a las cuales obtenemos información sobre la materia. La espectroscopia trata de las distintas interacciones de la radiación electromagnética con la materia. La investigación de la estructura molecular ha sido siempre un problema complicado, ya que antes de que se desarrollaran los métodos espectroscópicos era necesario un largo trabajo que comenzaba desde la determinación del peso molecular y de la fórmula molecular, para luego, romper la molécula en compuestos más sencillos conocidos, identificarlos por sus propiedades físicas y químicas, realizar síntesis parciales informativas y construir la estructura de los datos obtenidos.

La confirmación definitiva se lograba cuando se conseguía la síntesis, de acuerdo con la estructura propuesta, y las propiedades del compuesto sintetizado coincidían con las del problema. Los métodos espectroscópicos han simplificado notablemente este trabajo, de modo que, en muchos casos, la estructura química puede determinarse en pocas horas, estos métodos se basan en que las moléculas orgánicas interactúan con las radiaciones electromagnéticas, de distintas longitudes de onda debido a las vibraciones de sus electrones, la longitud de onda absorbida depende de la estructura molecular, y esto es lo que nos proporciona información sobre la muestra que se quiere analizar [26].

En el desarrollo de esta tesis utilizamos espectroscopía de absorción en el cercano infrarrojo y la espectroscopia de dispersión Raman. La espectroscopía infrarroja tiene casi 125 años de existencia. El primer espectro de vibraciones moleculares fue observado en 1881 por Abney y Festing, quienes prepararon emulsiones fotográficas sensibles al cercano infrarrojo y fotografiaron el espectro de absorción de 48 líquidos orgánicos. Encontraron bandas características en estos espectros, las cuales asociaron con la presencia de hidrógeno en las moléculas estudiadas. En 1892 Julius observó el espectro infrarrojo de 20 compuestos orgánicos, encontrando que todos los compuestos que contienen metilo (CH_3) exhiben una banda de absorción de $3.45\mu\text{m}$ y llegó a la conclusión de que la absorción de “ondas caloríficas” se debe a movimientos intramoleculares [27], en otras palabras, la estructura interna de la molécula determina el tipo de absorción, es decir los espectros infrarrojos de una sustancia determinada pueden ser considerados como la huella digital de dicha sustancia.

La espectroscopía en el cercano infrarrojo, es una técnica que proporciona información química y estructural de los fluidos biológicos. Las intensidades espectrales son proporcionales a la concentración de la mezcla, por lo cual es posible cuantificar la proporción de ellos; es una herramienta valiosa que actualmente se utiliza para analizar mezclas de aminas aromáticas, para la determinación de proteínas, grasas, humedad, azúcares y aceites [28]. Los primeros equipos comerciales de espectroscopia infrarroja aparecieron a mediados del siglo XX, habiéndose impulsado su desarrollo durante la segunda guerra mundial, cuando se utilizaron para síntesis de caucho sintético, utilizado extensamente en neumáticos para automóviles y camiones.

La espectroscopia infrarroja se basa en el hecho de que las moléculas vibran a frecuencias específicas en respuesta a la interacción de la luz con la materia, es decir, los movimientos de rotación y vibración tienen niveles de energía discretos (modos normales vibracionales).

Las frecuencias resonantes o frecuencias vibracionales son determinadas por las formas de las superficies de energía potencial molecular, las masas de los átomos y, eventualmente por el acoplamiento vibrónico asociado.

Cabe mencionar que como antecedente a esta tesis, se tomó como punto de partida un trabajo previo de tesis doctoral de Espectroscopías ópticas para análisis químico de la sangre, que se realizó en colaboración con el hospital central de San Luís Potosí, utilizando el estándar de oro para el análisis químico de sangre y con espectroscopia NIR para detectar cambios de concentración de glucosa en sangre [29].

El fenómeno conocido como efecto Raman fue descrito por primera vez por el físico hindú Chandrasekhara Venkata Raman en el año de 1928, lo cual le significó la obtención del premio Nobel de física en 1930. La espectroscopía Raman es una técnica fotónica de alta resolución, que proporciona en pocos segundos; información química y estructural de casi cualquier material o compuesto orgánico y/o inorgánico permitiendo así su identificación.

El análisis mediante espectroscopía Raman se basa en analizar la luz dispersada por un material al hacer incidir sobre él un haz de luz monocromático. Al interactuar con la materia, una pequeña porción de la luz es dispersada inelásticamente experimentando ligeros cambios de frecuencia que son característicos del material que se está analizado, y es independientes de la frecuencia de la luz incidente.

Una de las ventajas de las espectroscopias ópticas es que se realiza directamente sobre la muestra, sin necesitar ningún tipo de preparación especial, y sin alterar la superficie sobre la que se realiza el análisis, es decir, es no-destructiva [30], en la parte 2 de la tesis, se describirán a detalle las técnicas de espectroscopia NIR y Raman.

Actualmente el análisis de fluidos biológicos se lleva a cabo mediante diversas técnicas como los son las pruebas serológicas, las pruebas de

inmunofluorescencia de anticuerpos, y la prueba de ELISA [31-32]; sin embargo, se han venido desarrollando otras técnicas ópticas, donde se muestra que es posible utilizar la espectroscopia Raman para el estudio de tejidos biológicos [33]. También el monitoreo no invasivo de glucosa en sangre por medio de técnicas espectroscópicas tales como la absorción en el cercano a infrarrojo, y Raman que con el apoyo del análisis multivariante [34-35] son técnicas confiables para identificar analitos en fluidos biológicos.

JUSTIFICACIÓN

El éxito en el análisis de diversos fluidos biológicos depende en gran parte de la efectividad de los resultados realizados en el análisis químico. En esta tesis se propone el uso de la espectroscopia de absorción en el cercano a infrarrojo y de la espectroscopia de dispersión Raman como técnicas para el análisis de fluidos biológicos, con fines de diagnóstico clínico, representando ventajas sobre los métodos tradicionales y en base a los resultados reportados en la literatura, así como el uso del análisis multivariante, que nos permite tener una nueva forma de identificar y cuantificar componentes químicos en fluidos biológicos.

HIPOTESIS

La espectroscopia de absorción en el cercano infrarrojo y la espectroscopia de dispersión Raman son técnicas alternativas y complementarias que pueden ser utilizadas con fines de diagnóstico clínico en el análisis de muestras biológicas, así como el uso de los métodos multivariantes para el análisis de los datos espectroscópicos.

OBJETIVO GENERAL

Evaluar el uso de las espectroscopías ópticas NIR y Raman, así como el uso de los métodos multivariantes para la caracterización y análisis de muestras biológicas.

OBJETIVOS ESPECÍFICOS

1. Evaluar la influencia de factores externos que influyen en la calibración y robustez de modelos de predicción usando datos espectroscópicos.
2. Factibilidad de las técnicas espectroscópicas para brindar información química y bioquímica específica en fluidos biológicos y la capacidad de los métodos multivariantes para el análisis e interpretación de datos espectroscópicos con fines de discriminación y/o clasificación.
3. Caracterización química y bioquímica del Biofilm, producido por la bacteria *E. faecalis*, mediante técnicas espectroscópicas.

Importancia del estudio

La importancia de los resultados experimentales abordados en esta tesis muestran las capacidades de las técnicas espectroscópicas para el estudio de muestras biológicas, con fines de diagnóstico clínico, y la capacidad de los métodos multivariantes para el análisis los datos espectroscópicos con fines de discriminación, clasificación y/o predicción. Se debe destacar que estos procedimientos ópticos en comparación con los métodos de referencia tradicionales tienen la virtud de su rapidez, su bajo costo y el empleo de pequeñas cantidades de muestras para el análisis (1 a 2mL), no obstante, en el estado actual de su desarrollo no pretenden desplazar a los métodos tradicionales.

Limitaciones del estudio

Algunas de las mediciones realizadas con las espectroscopias ópticas NIR (Cercano a infrarrojo) y Raman (Espectroscopia de dispersión), se obtuvieron con un número limitado de muestras.

Definición de términos

Antígeno: Del Griego ANTI que significa opuesto y GENO que significa producir, generar, (que genera oposición). Es una sustancia que desencadena la formación de anticuerpos y puede causar una respuesta inmune. Los antígenos son usualmente proteínas o polisacáridos. Esto incluye partes de bacterias (cápsula, pared celular, flagelos y toxinas) de virus y otros microorganismos.

Anticuerpo: Se conocen también como inmunoglobulinas, son glucoproteínas del tipo gamma globulina. Pueden encontrarse de forma soluble en la sangre u otros fluidos corporales de los vertebrados, disponiendo de una forma idéntica que actúa como receptor de los linfocitos B y son empleados por el sistema inmunitario para identificar y neutralizar elementos extraños tales como bacterias, virus o parásitos.

Phantom: Sustancia química que puede utilizarse como precursor químico.

Anaerobiosis: Proceso metabólico que tiene lugar en ausencia de oxígeno.

ELISA: Es un examen de laboratorio, para confirmar la presencia de anticuerpos, que se realiza en las enzimas y su nombre viene del término inglés “ Enzyme-linked immunosorbent assay” que quiere decir “ Ensayo Inmonoenzimático ligado a enzimas”

1. ESPECTROSCOPIAS Y MÉTODOS MULTIVARIANTES: FUNDAMENTOS TEÓRICOS

1.1 ESPECTROSCOPIAS ÓPTICAS:

Prácticamente todo lo que sabemos sobre la materia proviene de la interacción de la materia con la radiación electromagnética, dichas interacciones definen a la espectroscopia. La espectroscopia investiga los niveles de energía de átomos y moléculas, en consecuencia es una aplicación de la mecánica cuántica.

El análisis espectral en el cual se basa, permite detectar la absorción, o emisión de radiación electromagnética a ciertas longitudes de onda, y relacionar éstas con los niveles de energía implicados en una transición cuántica. Actualmente, las espectroscopías ópticas han despertado gran interés como una herramienta para análisis cuantitativo de muestras biológicas. Esta absorción o emisión se encuentra asociada a los cambios de estado de energía de las especies químicas interactuantes y, puesto que cada especie tiene estados energéticos que la caracterizan, la espectroscopía puede utilizarse con fines de identificación y también información cuantitativa [36].

Muchas propiedades de la radiación electromagnética se explican adecuadamente con un modelo clásico de onda sinusoidal, que utiliza parámetros como la longitud de onda, la frecuencia, la velocidad y la amplitud. Sin embargo, para explicar fenómenos asociados con la absorción o la emisión de energía radiante, hay que acudir a un modelo corpuscular en el que la radiación electromagnética se contempla como un flujo de partículas discretas, o paquetes ondulatorios de energía denominados fotones, en los que la energía de un fotón es proporcional a la frecuencia de la radiación. Este doble punto de vista de la radiación como partícula y como onda no es mutuamente excluyente, sino complementario. Los espectros se pueden observar mediante espectroscopios, que permiten realizar mediciones sobre éste, como la longitud de onda, la frecuencia y la intensidad de la radiación.

Atendiendo a su longitud de onda, la radiación electromagnética recibe diferentes nombres, y varía desde los energéticos rayos Gamma, hasta las ondas de radio, pasando por el espectro visible, el rango completo de longitudes de onda es lo que se denomina el espectro electromagnético.

Cuando una molécula absorbe energía de radiación electromagnética, puede sufrir varios tipos de excitación:

- Excitación electrónica
- Excitación rotacional
- Excitación que induce cambios del spin nuclear
- Excitación de deformaciones de enlace
- Ionización

La tabla número 1 muestra los intervalos de longitud de onda de las regiones del espectro que interesan para fines analíticos, así como los nombres de los diversos métodos espectroscópicos asociados a cada uno. Los métodos espectroscópicos se basan en la interacción radiación electromagnética-materia, y la materia puede estar en forma de átomos libres o moléculas, por lo que pueden distinguirse la espectroscopía atómica y la molecular.

Tabla 1. Región del espectro electromagnético y métodos espectroscópicos basados en la radiación electromagnética [26].

TIPO DE ESPECTROSCOPIA	INTERVALO DE LONGITUD DE ONDA
Emisión de rayos gamma	0.005-1 Angstrom
Absorción, emisión, fluorescencia y difracción de rayos X	0.1-100Angstrom
Absorción Ultravioleta de vacío	10-180nm
Absorción, emisión y fluorescencia ultravioleta visible	180-780nm
Absorción Infrarroja y dispersión Raman	0.78-300 μ m
Absorción de microondas	0.75-3.75mm
Resonancia de espín electrónico	3 cm
Resonancia magnética nuclear	0.6-10m

Niveles de energía atómica

De acuerdo con la teoría cuántica, los átomos sólo pueden existir en niveles discretos de energía potencial. La energía potencial de un átomo depende de su configuración electrónica. La transición entre los electrones externos y los niveles de energía producen la emisión o absorción de radiación en valores discretos de

energía. La frecuencia de radiación es proporcional al cambio en la energía potencial involucrada y está dada por la ecuación $E = h\nu$, donde cada transición da lugar a una frecuencia específica de radiación y, por consiguiente a la aparición de una línea en los espectros de absorción o de emisión.

Niveles de energía electrónica molecular

Los electrones de una molécula pueden excitarse a estados de mayor energía y se puede estudiar la radiación absorbida, o la energía emitida al regresar a su estado fundamental o basal. Las transiciones entre los diferentes estados electrónicos están acompañados de transiciones entre los niveles de energía rotacional y vibracional, lo cual imparte estructura fina a bandas de absorción electrónica; como resultado, los espectros de las moléculas son mucho más complicados que los correspondientes a los átomos. Las vibraciones y rotaciones de las moléculas revelan gran cantidad de información acerca de la estructura molecular. En la espectroscopía molecular se tienen las siguientes clasificaciones:

Espectroscopía de microondas

Espectroscopía de infrarrojos

Espectroscopía visible-ultravioleta

Espectroscopía Raman

Espectroscopía de RMN (Resonancia Magnética Nuclear)

En esta tesis se utilizaron la espectroscopia en cercano infrarrojo y la espectroscopia Raman.

Las medidas de absorbancia son utilizadas en química analítica, ya que la absorbancia es proporcional a la concentración de la sustancia.

Métodos basados en la absorción

Los métodos cuantitativos basados en la absorción requieren dos medidas de potencia: una, antes de que el haz haya pasado a través del medio que contiene el analito (P_0), y la otra después (P). La transmitancia y la absorbancia son los dos términos que se utilizan ampliamente en la espectrometría de absorción y se relacionan por la razón de P_0 y P .

Transmitancia

En la figura 1 se muestra un haz de luz de radiación paralelo, antes y después de atravesar un medio que tiene un espesor de b cm y una concentración c de una especie absorbente. Como consecuencia de las interacciones entre los fotones y los átomos o moléculas absorbentes, la potencia del haz disminuye de P_0 a P . La transmitancia T del medio es la fracción de radiación incidente transmitida por el medio:

$$T = \left(\frac{P}{P_0} \right) \quad (1)$$

La transmitancia se expresa con frecuencia como porcentaje o

$$\%T = \left(\frac{P}{P_0} \right) \times 100\% \quad (2)$$

$$A = \log \left(\frac{P_0}{P} \right) \quad (3)$$

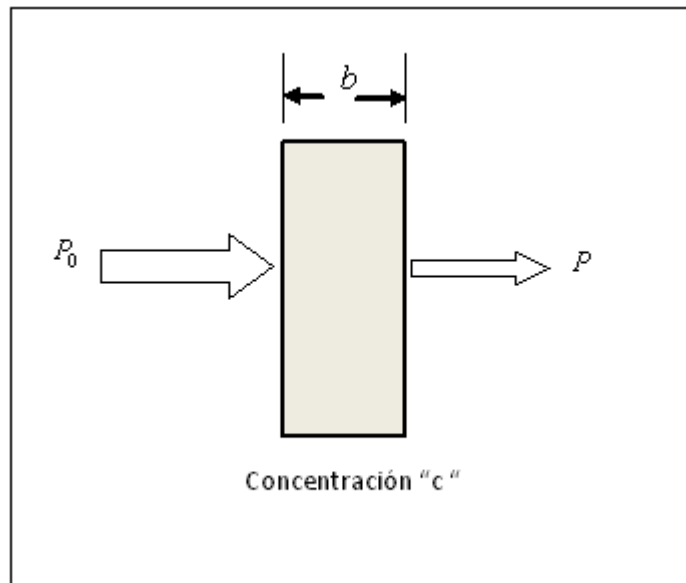


Figura 1. Atenuación de un haz de radiación por una disolución absorbente.

Absorbancia

La espectroscopia por absorción molecular se basa en la medición de la transmitancia T o en la absorbancia A de las soluciones que están en celdas transparentes que tienen una longitud de trayectoria de b cm. Normalmente, la concentración de un análito absorbente se relaciona en forma lineal con la absorbancia, según la Ley de Beer.

La absorbancia A de un medio se define por la ecuación:

$$A = -\log_{10} T = \log\left(\frac{P_0}{P}\right) \quad (4)$$

Cabe hacer notar que, al contrario de la transmitancia, la absorbancia de un medio aumenta cuando la atenuación del haz se hace mayor.

La Ley de Beer

Para una radiación monocromática, la absorbancia es directamente proporcional al camino óptico b a través del medio y la concentración c de la especie absorbente. Estas relaciones vienen dadas por:

$$A = abc \quad (5)$$

Donde a es una constante de proporcionalidad denominada absorptividad. Las unidades de a dependerán de las unidades utilizadas para b y c . Con frecuencia para disoluciones de una especie absorbente, b se da en centímetros y c en gramos por litro. En caso de expresar la concentración en términos de moles por litro y la longitud de la cubeta en centímetros, la absorptividad se denomina absorptividad molar y se representa por el símbolo ε . Así, cuando b está en centímetros y c en moles por litro.

$$A = \varepsilon bc \quad (6)$$

El valor del coeficiente de absorción a varía según los materiales absorbentes y con la longitud de onda para cada material en particular, cabe mencionar que la ley tiende a no ser válida para concentraciones muy elevadas, especialmente si el material dispersa mucho la luz. La relación de la ley entre la concentración y la absorción de luz está basada en el uso de espectroscopía para identificar sustancias. Cuando una molécula absorbe radiación electromagnética y pasa de un estado de baja energía a otro de energía mayor, la frecuencia de la radiación absorbida viene dada por la ecuación:

$$E = h\nu \quad (7)$$

Donde E es la energía absorbida, ν es la frecuencia de la radiación electromagnética y h es la constante de Planck = 6.624×10^{-27} erg/s, la energía también se puede expresar en función de la longitud de onda de la radiación con la siguiente ecuación:

$$E = h\nu = h(c / \lambda) \quad (8)$$

Siendo λ la longitud de onda y c la velocidad de la luz. En el espectro se utiliza el inverso de la longitud de onda en cm^{-1} para señalar la posición de las bandas de absorción.

1.2 ESPECTROSCOPIA RAMAN

En 1928 el físico Hindú Chandrasekhara Venkata Raman encontró que la longitud de onda de una pequeña fracción de la radiación dispersada por ciertas moléculas, difiere del haz incidente y, que los desplazamientos de la longitud de onda dependen de la estructura química de las moléculas responsables de la dispersión. Debido a ello, recibió el premio Nobel de Física en 1931.

El análisis mediante espectroscopía Raman está relacionado con el mismo tipo de cambios vibracionales cuantizados que se producen en la absorción infrarroja. Una gran ventaja de los espectros Raman respecto de los de infrarrojo, se debe al hecho de que el agua no produce interferencias y, por tanto, es posible obtener los espectros Raman de disoluciones acuosas.

La espectroscopia Raman se basa en irradiar una muestra con una fuente láser monocromática visible o infrarroja.

En la figura 2 se muestra un diagrama energético en el que cada estado de energía se representa por una línea horizontal, y proporciona una visión cualitativa del origen de la dispersión Raman.

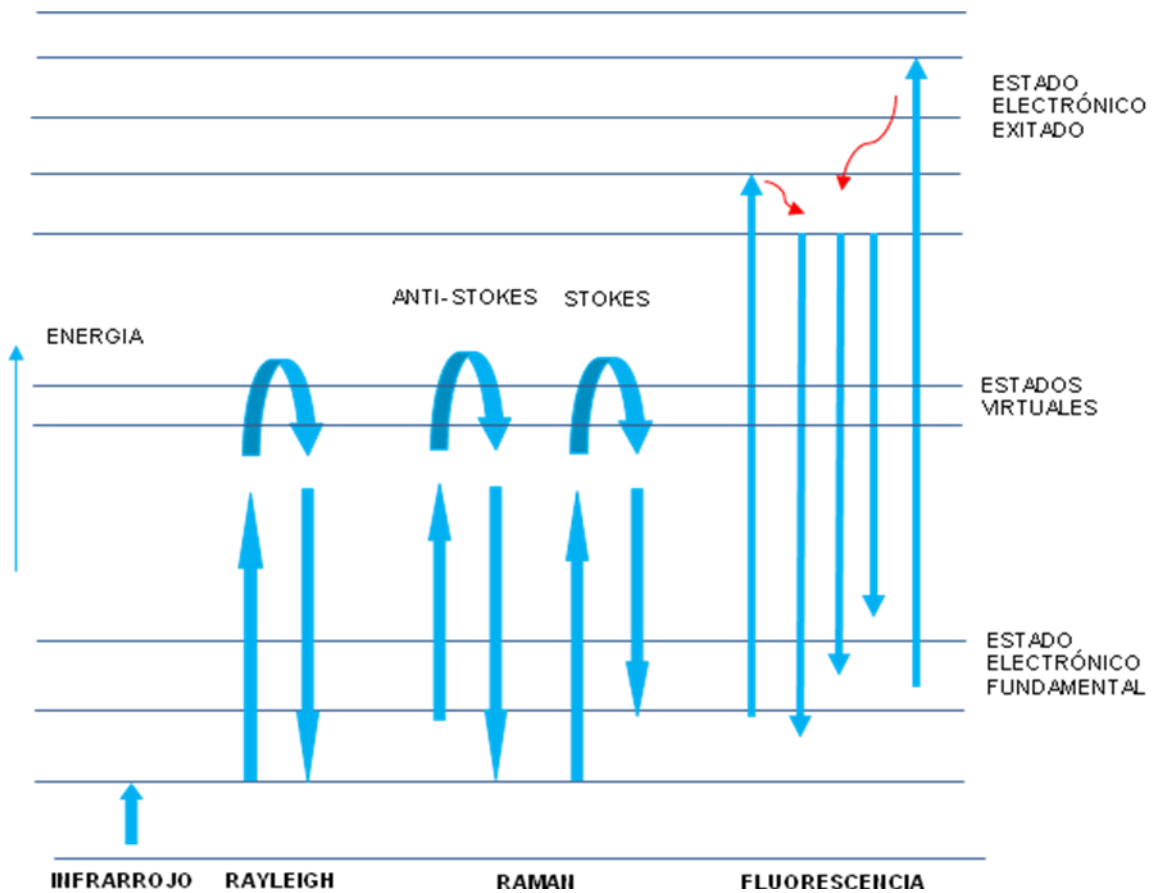


Figura 2. Diagrama energético en el que las líneas horizontales representan distintos estados vibracionales y en el que se muestran las transiciones entre estados energéticos para diferentes interacciones luz-materia.

La interacción de los fotones incidentes con la muestra produce que una fracción sea dispersada, esta dispersión significa que el fotón incidente lleva a la molécula transitoriamente a un nivel de energía vibracional llamado (nivel virtual excitado), el cual abandonará rápidamente para pasar a uno de los niveles de energía permitidos emitiendo un fotón, la frecuencia a la cual es liberado este fotón dependerá del salto energético realizado por la molécula.

Resultando los siguientes casos:

Si un fotón es dispersado a la misma frecuencia que el fotón incidente, se dice que el choque es elástico, no hay cambios en su estado energético, y con la misma frecuencia ν_0 que la frecuencia incidente, a esta dispersión se le llama dispersión Rayleigh.

Se puede presentar un choque inelástico, si como resultado de la interacción fotón-molécula se presenta una frecuencia distinta a la incidente, a esta se le conoce como dispersión Raman.

En caso de que la frecuencia sea mayor a la incidente, se produce una transferencia de energía de la molécula al fotón y el fotón es dispersado a una frecuencia $\nu_0 + \nu_r$, llamada dispersión Raman anti-Stokes.

En caso de que la frecuencia resultante sea menor a la incidente, el fotón es dispersado a una frecuencia $\nu_0 - \nu_r$, llamada dispersión Raman Stokes

En la figura 3 se observa un espectro Raman el cual está formado por una banda principal o Rayleigh y dos series de bandas secundarias correspondientes a las bandas Raman stokes y Raman anti-stokes.

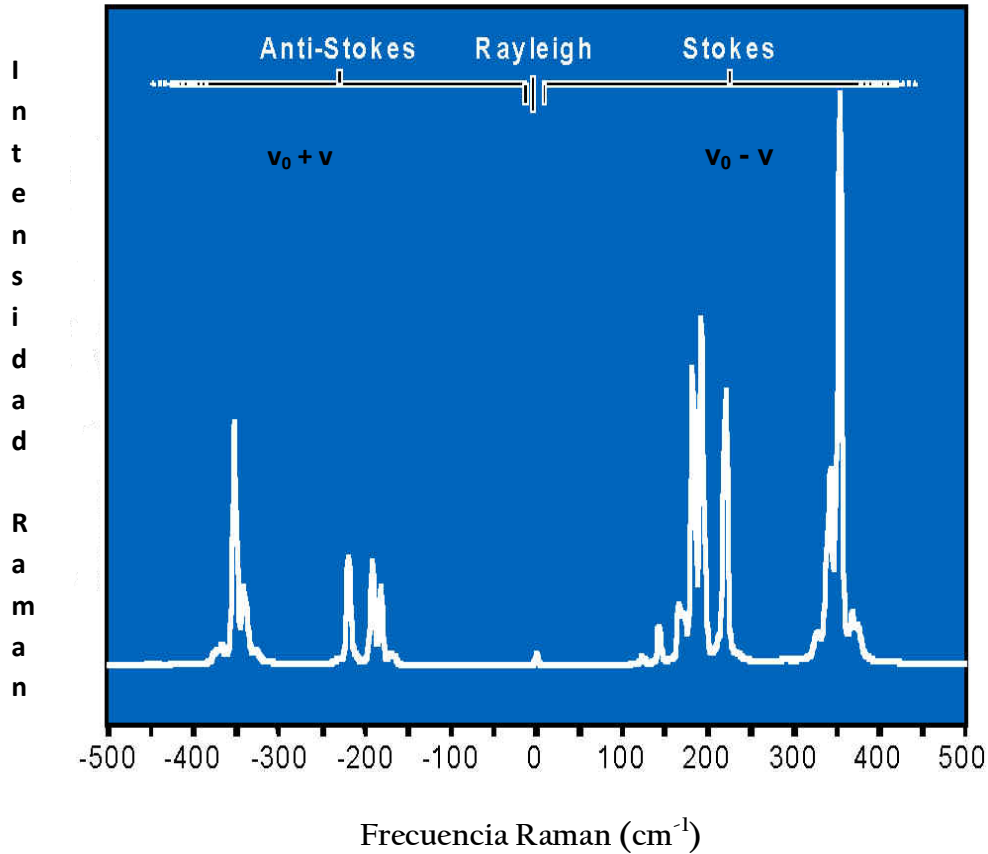


Figura 3 Frecuencias Raman Stokes, Raman Anti-stokes y Rayleigh

El espectro Raman recoge estos fenómenos representando la intensidad óptica dispersada en función del número de onda normalizado $\tilde{\nu}$ al que se produce. El número de onda es una magnitud proporcional a la frecuencia e inversamente proporcional a la longitud de onda, que se expresa en cm^{-1}

$$\tilde{\nu} = \frac{\nu}{\left(\frac{c}{n}\right)} = \frac{1}{\lambda} [\text{cm}^{-1}] . \quad (9)$$

Donde: c es la velocidad de la Luz (2.9979×10^{10} cm/s)

(c/n) es la velocidad de la luz en un medio cuyo índice refractivo es n

λ es la longitud de onda en nm.

Aplicaciones de la Espectroscopia Raman

La espectroscopia Raman se aplica en el análisis cualitativo y cuantitativo de sistemas inorgánicos, orgánicos y biológicos. Como se mencionó anteriormente, la espectroscopia Raman tiene la ventaja de se puede utilizar en disoluciones acuosas, se realiza directamente sobre el material a analizar, tiene las ventajas sobre otras técnicas de caracterización de compuestos que no requieren ninguna preparación y que no conlleva ninguna alteración de la superficie sobre la que se realiza el análisis.

Algunas aplicaciones son:

Aplicaciones Biológicas: Se ha utilizado también en la identificación de microorganismos como la *Candida* spp. [37]. De igual manera, ha servido como auxiliar en el diagnóstico de malignidad en cavidad oral [38].

Análisis de polímeros: La aplicación de la espectroscopia Raman en el análisis de polímeros, permite investigar cambios estructurales en pequeñas regiones de la muestra, estos cambios pueden ser debido a la aplicación de esfuerzos externos, microdefectos, impurezas, recubrimientos, etc. [36]

Agricultura: La espectroscopia Raman ha sido usada para analizar compuestos orgánicos o inorgánicos en suelos y aguas y para la caracterización de pesticidas orgánicos [39]

Nanotecnología: Se ha desarrollado una técnica más sensible y segura para la detección de células cancerígenas, esta técnica combina los avances más recientes de la nanotecnología con los principios del efecto Raman y permite detectar masas microscópicas de solo unas cuantas células a diferencia de los

métodos actuales como la tomografía de emisión de positrones que solo permite detectar masas tumorales a partir de 5 mm, esta nueva tecnología ha sido probada con éxito hasta el momento solo en ratones.

Espectroscopia Raman y Odontología:

Como se describió anteriormente, la espectroscopia Raman es útil para la detección de grupos funcionales que permiten la identificación de compuestos orgánicos e inorgánicos [40].

Estudios previos de espectroscopia Raman en odontología incluyen la estructura de materiales sintéticos, análisis de tejido óseo y componentes de fosfato de calcio incluyendo hidroxiapatita [41]. Recientemente también ha sido usado para analizar resinas adhesivas dentales [42] y composición de hidróxido de calcio posterior a una medicación intraconducto [43-44], pero la mayoría de los estudios han sido limitados a esmalte dental, el cual contiene solo un porcentaje de materiales orgánicos.

Como parte de este trabajo de tesis, se realizaron estudios sobre la estructura química de exopolosacáridos producido por la bacteria *E. faecalis*, formado en una cámara de anaerobiosis, utilizando la espectroscopia Raman.

1.3 ESPECTROSCOPIA INFRARROJA:

Desde el punto de vista de las aplicaciones como de la instrumentación, es conveniente dividir el espectro infrarrojo en tres regiones denominadas infrarrojo cercano (NIR), medio (MIR) y lejano (FIR). La espectroscopía de absorción en infrarrojo, es una espectroscopia molecular que se basa en la interacción de la radiación electromagnética con las moléculas. Debido a que la energía de la radiación depende de su longitud de onda, la interacción molecular comprende la excitación de los electrones, vibraciones moleculares y rotaciones moleculares, de tal forma que cuando la molécula absorbe radiación, su estado de energía vibracional y rotacional cambia.

Las bandas de absorción en infrarrojo se forman debido a transiciones entre niveles de energía vibracionales, de modo que cuando el número cuántico cambia en una unidad, se obtiene lo que conocemos como banda fundamental (mediano IR) y si cambia en más de una unidad, se le denominan sobretonos (cercano infrarrojo).

Los modos de vibración de una molécula de n -átomos es igual a $(3n-6)$ modos y de $(3n-5)$ modos, si es lineal y producen cambios en la distancia de enlace o elongaciones, las cuales pueden ser simétricas o asimétricas. También se producen cambios de los ángulos de enlace, los cuales pueden ser simétricos o asimétricos. Los espectros de absorción, emisión y reflexión en el infrarrojo, de especies moleculares, se pueden explicar asumiendo que todos son el resultado de los distintos cambios energéticos producidos en las transiciones de las moléculas de unos estados de energía vibracionales y rotacionales a otros. La absorción de radiación en el infrarrojo se limita en gran parte, a especies moleculares para las cuales existen pequeñas diferencias de energía entre los distintos estados vibracionales y rotacionales. Para absorber radiación en el infrarrojo, una molécula debe sufrir un cambio neto en el momento dipolar como consecuencia de su movimiento de vibración o de rotación.

Sólo en estas circunstancias, el campo eléctrico alterno de la radiación puede interactuar con la molécula, y provocar cambios en la amplitud de alguno de sus movimientos. El momento dipolar está determinado por la magnitud de la diferencia de carga y por la distancia entre los dos centros de carga. Dado que la molécula de ácido clorhídrico vibra, se produce una constante de variación del momento dipolar, lo que origina un campo que puede interactuar con el campo eléctrico asociado a la radiación.

Si la frecuencia de radiación coincide exactamente con la frecuencia de vibración natural de la molécula, tiene lugar una transferencia neta de energía que origina un cambio en la amplitud de la vibración molecular; la consecuencia es la absorción de radiación. De manera análoga, la rotación de las moléculas

asimétricas, alrededor de sus centros de masa, produce una variación periódica en el dipolo que puede interactuar con la radiación. Cuando se trata de especies homonucleares como O_2 , N_2 o Cl_2 , el momento dipolar no sufre un cambio neto durante la vibración o la rotación y, como consecuencia, este tipo de compuestos no absorben en el infrarrojo. Con excepción de algunos compuestos de este tipo, todas las demás especies moleculares absorben radiación en el infrarrojo.

Transiciones rotacionales

La energía necesaria para provocar un cambio en los niveles rotacionales es muy pequeña y corresponde a radiaciones de 100 cm^{-1} o menores ($>100\mu\text{m}$). Dado que los niveles rotacionales están cuantizados, la absorción por los gases en la región del infrarrojo lejano se caracteriza por líneas discretas bien definidas. En los líquidos o los sólidos los choques e interacciones intramoleculares causan el ensanchamiento de las líneas originando un espectro continuo [45].

Tipos de vibraciones moleculares

Las posiciones relativas de los átomos en una molécula no son exactamente fijas, sino que fluctúan continuamente como consecuencia de una multitud de tipos de vibraciones y rotaciones diferentes alrededor de los enlaces en la molécula. Para una molécula sencilla diatómica o triatómica es fácil definir el número y la naturaleza de tales vibraciones, y relacionarlas con las energías de absorción [45].

Sin embargo, en moléculas poliatómicas, se hace difícil, por las interacciones que aparecen entre algunos de sus centros de vibración. Pueden distinguirse dos categorías básicas de vibraciones: de tensión y de flexión.

Una vibración de tensión supone un cambio continuo en la distancia interatómica a lo largo del eje del enlace entre dos átomos. Las vibraciones de flexión se caracterizan por un cambio en el ángulo entre dos enlaces y son de cuatro tipos: de tijereteo, de balanceo, de aleteo y de torsión. Los distintos tipos de vibraciones se representan esquemáticamente en la figura 4. En una molécula que contiene más de dos átomos, son posibles todos los tipos de vibraciones mostrados en la

figura 3. Además puede producirse una interacción o acoplamiento de las vibraciones si éstas involucran enlaces con un mismo átomo central. El resultado del acoplamiento es un cambio en las características de las vibraciones implicadas.

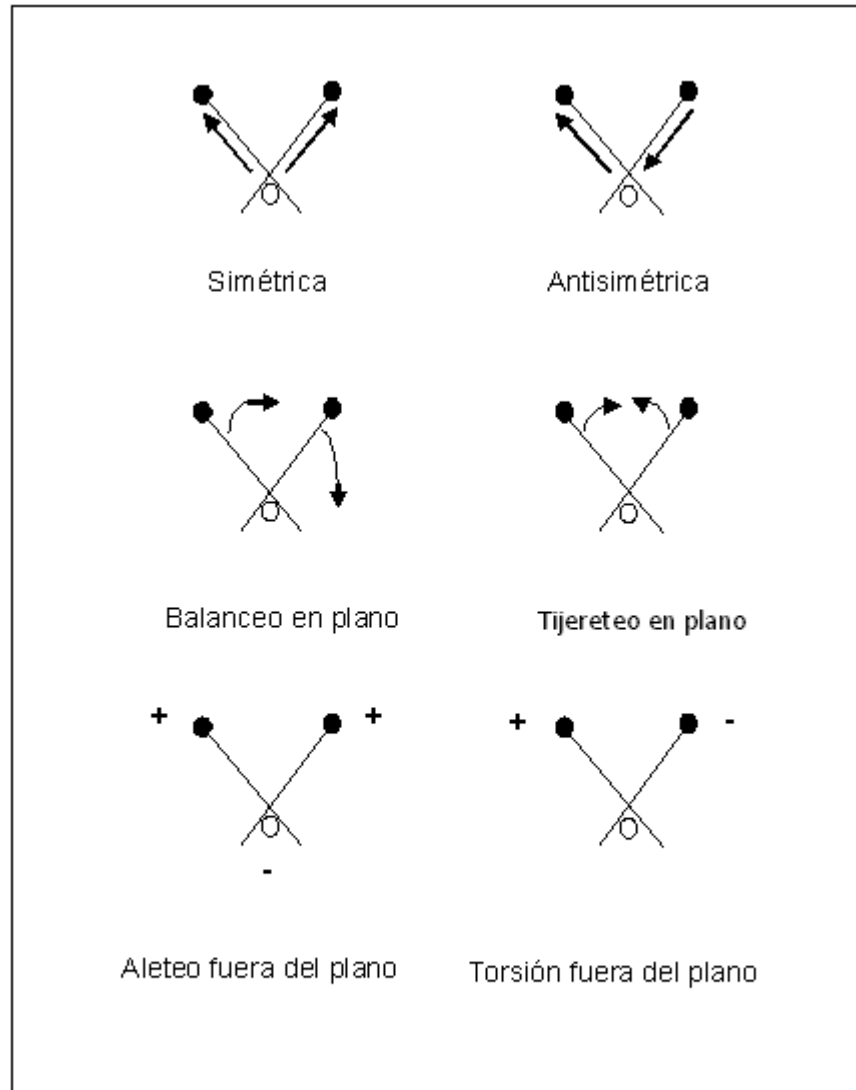


Figura 4. Tipos de vibraciones moleculares. El signo + indica un movimiento desde el plano de la página hacia el lector y – indica un movimiento desde el plano de la página alejándose del lector [45].

Modos de vibración

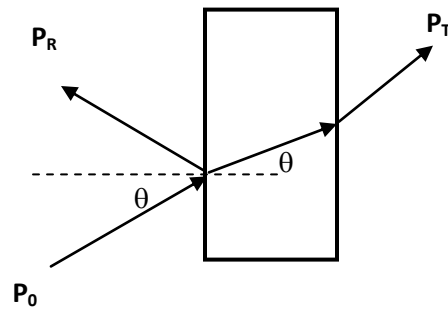
La cantidad de vibraciones posibles en una molécula poliatómica se puede calcular como sigue. Como se necesitan tres coordenadas para localizar un punto en el espacio, para fijar N puntos se requieren tres coordenadas para cada uno, es decir, un total de $3N$. Cada coordenada corresponde a un grado de libertad para uno de los átomos en una molécula poliatómica. Por este motivo, si una molécula contiene N átomos, se dice que tiene $3N$ grados de libertad [44].

Para definir el movimiento de una molécula, se tiene que considerar:

- 1) El movimiento de toda la molécula en el espacio, es decir, el movimiento de traslación de su centro de gravedad.
- 2) El movimiento de rotación de la molécula completa alrededor de su centro de gravedad.
- 3) El movimiento de cada uno de sus átomos respecto a los otros átomos, o en otras palabras, sus vibraciones individuales.

Tipos de medidas en infrarrojo

Cuando la radiación incide en la muestra, figura 5, ésta puede sufrir diferentes fenómenos: absorción, transmisión y reflexión. La intensidad de la luz transmitida a través de la muestra (P_T) es menos que la intensidad incidente (P_0). Una parte de esta intensidad incidente se ha reflejado (P_R), mientras que otra parte ha sido absorbida por la sustancia (P_A).



$$P_0 = P_A + P_R + P_T$$

Figura 5. Fenómenos de absorción, transmisión y reflexión de la radiación electromagnética al interactuar con la materia.

La medida más común en el infrarrojo es la que se basa en la absorción, aunque también se ha desarrollado espectroscopias basadas en el fenómeno de la reflexión como son la reflectancia total atenuada y la reflectancia difusa, en el desarrollo de esta tesis hablaremos solo de la reflectancia difusa.

Reflectancia difusa

Cuando la luz incide sobre una superficie y no absorbe, tiene lugar el fenómeno de reflexión especular regido por las ecuaciones de Fresnel. Como se muestra en la Figura 6.

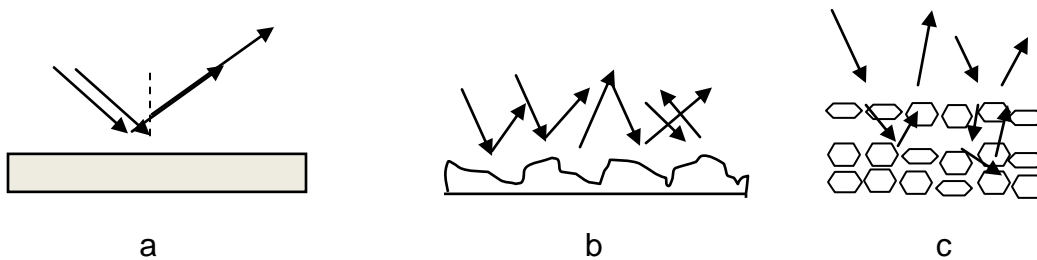


Figura 6. Procesos de reflexión en un material especular(a) o irregular (b). Y el fenómeno de reflectancia difusa (c).

La intensidad reflejada sobre el total incidente depende de los índices de refracción del aire y la muestra (n_1, n_2). Para el caso de un ángulo de incidencia igual a cero la expresión es [44]:

$$\frac{P_R}{P_o} = \frac{(n_2 - n_1)^2}{(n_2 + n_1)^2} \quad (10)$$

Cuando la luz incide sobre una superficie irregular, se puede considerar que la frontera entre el medio y la muestra está formado por una serie de pequeñas interfaces orientadas en todos los ángulos posibles. De esta forma, aunque cada una de estas pequeñas interfaces reflejan la luz siguiendo la ley de Fresnel, el efecto global es una reflexión de la luz a cualquier ángulo (reflectancia difusa).

La radiación que se transmite a través de la primera interface, puede sufrir absorción por parte de la muestra, por lo que la intensidad de la luz se verá atenuada según la ley de Beer. Esta radiación que ha atravesado la primera capa de partículas se difunde a las siguientes capas a través de reflexiones aleatorias, refracciones y dispersión y puede sufrir nuevos fenómenos de atenuación. La longitud de camino óptico seguido por la luz es muy difícil de describir matemáticamente, sobre todo si las partículas de la muestra tienen tamaños heterogéneos, por lo que no se ha desarrollado una teoría rigurosa de la reflectancia difusa.

Sin embargo existen teorías basadas en la práctica, la más conocida la de Kubelka- Munk.

$$\frac{(1 - R_\infty)}{2R_\infty} = \frac{k}{s} = \frac{\epsilon c}{s} \quad (11)$$

Donde R_{∞} es la reflectancia de la capa infinitamente gruesa, k y s son las constantes de scattering y absorción, respectivamente. El coeficiente de absorción (k), es igual a la concentración multiplicada por la absortividad definida en la ley de Beer (ϵ, c).

En la práctica la reflectancia difusa se mide respecto a un estándar no absorbente y a continuación se calcula el logaritmo para llegar a una relación lineal con la concentración.

$$\log \frac{R_{s\text{tan}d}}{R} = \log \frac{1}{R} + \log R_{s\text{tan}d} \propto \frac{\epsilon c}{s} . \quad (12)$$

Donde R y $R_{s\text{tan}d}$ representan la reflectancia de la muestra y del estándar respectivamente (siendo mayor la reflectancia del estándar que la de la muestra), Para la luz monocromática $\log R_{s\text{tan}d}$ es constante y puede ser ignorado. Si se cumple la condición de aditividad del espectro, la expresión anterior puede ser escrita como:

$$c = k + \frac{s}{\epsilon} \log \frac{1}{R} . \quad (13)$$

Con lo que, cómo en el caso de la ley de Beer, existe una relación entre la concentración y la radiación medida en el espectro.

La presencia de agua en el espectro provoca la aparición de bandas características y, además afecta a la totalidad del espectro debido a la variación que se introduce en el índice de refracción del medio (n_o), provocando un aumento de $\log \left(\frac{1}{R} \right)$. Puesto que el grado de humedad puede variar entre muestras, en la práctica la constante de dispersión (scattering) (s) se convierte en una incógnita para cada una de las nuevas muestras, por lo que no se puede llevar a cabo el análisis cuantitativo.

Asignación de bandas: En el espectro infrarrojo cercano, predominan las bandas debidas a sobretonos y combinaciones de enlaces en los que participa el hidrógeno.

Región del cercano Infrarrojo:

El descubrimiento de la radiación invisible se debe a Herschel en el año de 1800, cuando observó que el efecto de calentamiento relativo generado por las diferentes porciones del espectro de la luz solar al atravesar un prisma, continuaba aumentando cuando colocaba el termómetro más allá del rojo del espectro visible.

En la región del infrarrojo cercano que se toca con la región visible aproximadamente a los 12500 cm^{-1} ($0.8\text{ }\mu\text{m}$) y se extiende a los 4000 cm^{-1} ($2.5\text{ }\mu\text{m}$), hay muchas bandas de absorción que provienen de sobretonos armónicos de la banda fundamental o de bandas de combinaciones asociadas con los átomos de hidrógeno.

Entre estas se encuentran los primeros sobretonos de las vibraciones de estiramiento de los enlaces O-H y N-H, cerca de los 7140 cm^{-1} ($1.4\text{ }\mu\text{m}$) y de los 6667 cm^{-1} ($1.5\text{ }\mu\text{m}$), respectivamente, también se encuentran bandas de combinación que resultan de las vibraciones de estiramiento y de deformación del enlace C-H de los grupos alquilo, a los 4548 cm^{-1} ($2.2\text{ }\mu\text{m}$) y a los 3850 cm^{-1} ($2.6\text{ }\mu\text{m}$). La absorptividad de las bandas del NIR de 10 a 1000 veces menor que las bandas del infrarrojo intermedio, pero el espesor de las capas de muestra (0.5 a 10mm) compensa este inconveniente. Ya que las absorptividades son tan pequeñas, el haz de NIR penetra más profundo en la muestra en las técnicas de reflectancia, además las impurezas menores dan menos problemas que el infrarrojo medio, tanto en los métodos de transmisión como en los de reflectancia.

La espectrometría en infrarrojo cercano también es una herramienta valiosa para analizar mezclas de aminas aromáticas. La espectroscopía en el cercano infrarrojo (NIR), ha sido aplicada para el análisis de algunos parámetros de la sangre, tales como la oxigenación en el tejido y la sangre [46]. También los espectros de reflectancia en el infrarrojo cercano, encuentran amplia aceptación en la industria de granos y alimentos para la determinación de proteínas, grasas, humedad, azúcares, aceites [47]. Estos espectros también se han utilizado para la determinación de sustancias en la madera, de los componentes de los polímeros y aún en la exploración geológica área.

La espectroscopia de infrarrojo cercano resulta una técnica atractiva porque:

1. Proporciona una información química muy versátil. La espectroscopia molecular, como consecuencia del tipo de transiciones propias de esta zona, aporta información de una gran cantidad de compuestos químicos. Esta propiedad, combinada a las técnica multivariantes, permite considerar la determinación simultánea de más de un analito o propiedad de interés a partir del espectro registrado de la muestra.
2. La obtención del espectro se hace de una forma rápida y no es necesario un pre tratamiento de la muestra, de la misma forma que la espectroscopia con Raman.

Los espectros NIR consisten generalmente en una superposición de bandas, las cuales en apariencia no aportan gran información, sin embargo, el uso de técnicas de análisis multivariante para el procesado de estos datos espectrales son utilizados para un análisis cuantitativo. Hay que recalcar el hecho de que un espectro NIR es la unión de funciones instrumentales de la medición y las características ópticas de la muestra analizada, es decir, la muestra es un elemento ópticamente activo del espectrómetro.

Un espectro NIR, puede o no contener información relacionada con la muestra analizada por un determinado método de referencia, de esta forma, la estructura espectral nos permite conocer la relación causa y efecto entre la respuesta del instrumento y los datos de referencia, lo cual sienta las bases científicas para la calibración multivariante utilizando espectroscopia NIR.

En esta tesis doctoral, se ha abordado la determinación de muestras biológicas utilizando espectroscopia NIR y Raman.

Comparación del espectro Raman e IR

La base de espectroscopia IR es la absorción de la energía de fotones ($h\nu$) para la excitación de vibraciones moleculares. Los espectros Raman e IR no coinciden en cuanto a las bandas que aparecen, si bien pueden tener algunas bandas en común.

Para que se pueda dar una absorción IR el momento dipolar de la molécula o del grupo de moléculas debe cambiar durante la vibración, es decir, la componente eléctrica podrá transmitir su energía, a condición de que exista una concordancia entre la frecuencia mecánica de vibración y la frecuencia de radiación. Sin embargo, para obtener una línea de dispersión Raman, la polarizabilidad de la molécula debe cambiar durante la vibración. Por consiguiente, especialmente para moléculas simétricas, hay bandas que aparecen en Raman y que están ausentes en IR, y viceversa.

Enlaces altamente polares como C-F pero poco polarizables aparecen en IR, pero no en Raman, sin embargo, enlaces como S-S muy polarizable pero poco polar no aparecen en IR pero si en Raman. Por lo que ambas técnicas se pueden considerar complementarias. En muchos casos hay bandas comunes en los dos espectros porque hay una contribución simultánea de la vibración a un cambio de momento dipolar y polarizabilidad del grupo o la molécula, pero aun así, las intensidades de las bandas difieren totalmente de uno a otro espectro debido a la diferencia de principios físicos de interacción de la luz con la molécula.

La espectroscopia Raman puede ser usada para identificar grupos de átomos no polares, para investigar grupos simétricos o moléculas inactivas al IR (vibraciones simétricas se detectan) Mientras que IR va mejor para investigar moléculas poco simétricas y con grupos muy polares.

Aplicaciones de la espectroscopia infrarroja:

Los espectros de absorción en el cercano a infrarrojo se utilizan para el análisis cuantitativo de compuestos que contengan grupos funcionales formados por la unión de hidrógeno con carbono, nitrógeno y oxígeno

Algunos ejemplos son:

En alimentos: En el estudio de la deshidratación de la uva, en estudios para la detección de plomo en alimentos de origen vegetal, para identificar la autenticación de aceites de oliva virgen [47].

En la industria: Para la determinación del índice de octano en gasolinas [48], para medir el grado de polimerización en la manufactura de polímeros.

En medicina: Para la determinación de oxigenación en el cerebro [49], y para la determinación de los niveles de alcohol [50], y en el monitoreo de glucosa en sangre.

El medio ambiente y las artes: en el análisis de contaminantes atmosféricos provenientes de los procesos industriales, uso para la caracterización de muestras del patrimonio cultural, para la clasificación de maderas en la industria forestal.

En bioquímica: Para diferenciar bacterias de importancia clínica [51].

1.4 Análisis multivariante

El análisis multivariante tiene una historia que data desde el uso de regresión lineal por parte de Gauss en 1809 y, posteriormente, por otros estadísticos como Markov en 1900. Las técnicas más recientes datan desde los años 1930. El análisis multivariante es el conjunto de métodos estadísticos cuya finalidad es analizar simultáneamente conjuntos de datos multivariantes, en el sentido de que hay varias variables para cada individuo u objeto de estudio. Las técnicas de análisis multivariante están siendo ampliamente aplicadas a la industria, administración y centros de investigación, nos permiten explicar algunas de las principales relaciones que se establecen entre pares o un conjunto de variables y, dentro de la variedad de las mismas, nuestro objeto de estudio recomienda la aplicación de técnicas de reducción de casos. En principio reducir la información de toda una matriz o base de datos a un conjunto de elementos que nos permitan explicar el máximo de nuestro objeto con una pérdida mínima de información, es decir, reducir la información a aquello que realmente es sustantivo para la investigación. Una vez realizado este proceso de “factores explicativos” pueden convertirse en variables y tratarse de manera relacional con las demás variables del estudio. En el campo de la biología se pueden formular muchas preguntas sobre determinados fenómenos que se producen en el contexto de complejidad, pues en general dispondremos de muchos individuos sobre los que se miden muchos valores. Pero estas preguntas tienen en común que su respuesta ha de basarse en el análisis conjunto de muchas variables, es decir en el contexto del análisis estadístico multivariante. El elemento esencial del análisis multivariante es el valor teórico, una combinación lineal de variables con ponderaciones determinadas empíricamente. El investigador especifica las variables, y las ponderaciones son objeto específico de determinación por parte de la técnica multivariante.

Un valor teórico de n variables ponderadas (X_1 a X_n) puede expresarse matemáticamente así:

$$\text{Valor Teórico} = W_1X_1 + W_2X_2 + W_3X_3 + \dots W_nX_n . \quad (14)$$

Donde X_n es la variable observada y W_n es la ponderación determinada por la técnica multivariante.

El resultado es un valor único que representa la combinación de todo el conjunto de variables que mejor se adaptan al objeto de análisis multivariante específico. En regresiones múltiples, el valor teórico se determina de tal forma que guarde la mejor correlación con la variable que se está prediciendo. En el análisis discriminante, el valor teórico se forma de tal manera que produzca resultados por cada observación que diferencien de forma máxima entre grupos de observaciones. Y en el análisis factorial, los valores teóricos se forman para representar mejor las estructuras subyacentes o la dimensionalidad de las variables tal y como se representan en sus intercorrelaciones.

En cada caso, el valor teórico capta el carácter multivariante del análisis, por lo tanto, en este trabajo el valor teórico es el punto central del análisis por varias razones. Debemos entender no sólo su impacto conjunto para lograr cumplir el objetivo de cada técnica, sino también la contribución de cada variable separada al efecto teórico en su conjunto [52].

1.4.1 Análisis de componentes principales (Principal Component Analysis, PCA)

El análisis de componentes principales es una técnica estadística multivariante de simplificación o reducción que permite transformar un conjunto de variables no correlacionadas, denominadas Factores o Componentes Principales. El análisis de componentes principales es una aproximación estadística que puede usarse para analizar interrelaciones entre un gran número de variables y explicar estas variables en términos de sus dimensiones subyacentes comunes (factores).

El objetivo es encontrar un modo de condensar la información contenida en un número de variables originales en un conjunto más pequeño de variables (factores) con una pérdida mínima de información. Seleccionaremos un número de componentes de tal forma que en la transformación no se pierda mucha información. Esta reducción puede simplificar la aplicación sobre estos últimos de otras técnicas multivariantes (regresión, clusters, etc.) A continuación se muestra el algoritmo para condensar la información, sin que esta pierda la esencia de los datos.

Algoritmo:

1.- Partimos de los datos contenidos en la tabla

$$X_n = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{pmatrix}$$

$n \longrightarrow$ para individuos

$p \longrightarrow$ variables

2.- Se calcula la matriz de correlaciones $p \times p$

$$R = \begin{pmatrix} 1 & r_{12} & \dots & r_{1p} \\ r_{21} & 1 & \dots & r_{2p} \\ \dots & \dots & \dots & \dots \\ r_{p1} & r_{p2} & \dots & 1 \end{pmatrix}.$$

Se calculan los componentes principales C_i combinaciones lineales de las p variables.

$$C = a_{11}x_{11} + \dots + a_{1p}x_p,$$

$$\dots,$$

$$\dots,$$

$$C_p = a_{p1}x_1 + \dots + a_{pp}x_p.$$

Donde:

- a_{ij} representa los pesos de cada variable en cada componente
- cada componente explica una parte de la varianza total
- se encuentran pocos componentes k que expliquen casi toda la varianza total.

Si calculamos los valores originales de la matriz de correlación (λ), la varianza de la nube de puntos explicada por cada componente coincide con el valor propio asociado.

La porción de varianza total que recoge cada componente será el porcentaje que presenta su valor propio frente al total.

$$\frac{\lambda_k}{\lambda_1 + \dots + \lambda_p} \times 100. \tag{15}$$

Como la varianza de cada variable estandarizada es 1, la varianza total de la nube será igual a p (p veces 1), con lo cual la expresión anterior quedaría

$$\frac{\lambda_k}{p} \times 100. \tag{16}$$

El coeficiente de correlación (r_{ij}) entre un componente C_j y una variable x_i se calcula multiplicando el peso de la variable en ese componente por la raíz cuadrada de su valor propio.

$$r_{ij} = \sqrt{\lambda_j a_{ij}}. \tag{17}$$

$$i: 1 \dots p, j: 1 \dots k.$$

Estos coeficientes representan la parte de la varianza que cada factor retenido explica de cada variable, se denominan pesos o cargas factoriales.

Obteniendose la matriz de cargas factoriales ($p \times k$)

$$x_1 = r_{11}c_{11} + r_{1k}c_k,$$

.....,

.....,

$$x_p = r_{p1}c_1 + r_{pk}c_k.$$

La suma factorial de los cuadrados de las cargas factoriales se denomina comunalidad (parte de la dispersión total de la variable explicada por el conjunto k de componentes). La suma de todas las comunidades coincide con el porcentaje de la dispersión total explicado por los componentes retenidos y coincide con la suma de los valores propios. La suma vertical de los cuadrados de las cargas factoriales de todas las variables coincide con su valor propio. Se pueden calcular las nuevas coordenadas de los puntos sobre los nuevos componentes o factores.

1.4.2 Mínimos cuadrados Parciales (PLS)

La regresión múltiple es el método de análisis apropiado cuando el problema del investigador incluye una única variable métrica dependiente que se supone está relacionada con una o más variables métricas independientes. El objetivo del análisis de regresión múltiple es predecir los cambios en la variable dependiente en respuesta a cambios en varias de las variables independientes, esta técnica es útil siempre que se esté interesado en predecir la cantidad o magnitud de la variable dependiente [52].

Los modelos de regresión se usan con varios fines, que incluyen los siguientes:

- Descripción de datos
- Estimación de parámetros
- Predicción y estimación.

La regresión lineal se ha utilizado en la determinación de glucosa subcutánea utilizando espectroscopia NIR, en la predicción de concentración de alcohol utilizando espectroscopia en cercano a infrarrojo de reflectancia difusa.

También se han utilizado cuando se tiene interés en analizar la relación entre variables cuantitativas como por ejemplo: Presión Sanguínea y edad, estatura y peso. En medicina, las primeras evidencias relacionando la mortalidad con el fumar tabaco vinieron de estudios que utilizaban regresión lineal.

En esta tesis se utilizan los modelos de regresión para la predicción de variables respuesta. Una vez calibrado un modelo en base a cambios en concentración de glucosa, podemos predecir concentración de futuras muestras.

Modelo de regresión lineal simple

El modelo de regresión lineal simple es:

$$y = \beta_0 + \beta_1 x + \varepsilon . \quad (17)$$

Donde la ordenada al origen β_0 y la pendiente β_1 son constantes desconocidas, y ε es un componente aleatorio de error. Se supone que los errores tienen promedio cero y varianza σ^2 desconocida. Además se suele suponer que los errores no están correlacionados.

Esto quiere decir que el valor de un error no depende del valor de cualquier otro error. Conviene considerar que el regresor x está controlado por el analista de datos, y se puede medir con error despreciable, mientras que la respuesta y es

una variable aleatoria. Con lo que hay una distribución de probabilidades de y para cada valor posible de x . La media de esta distribución es

$$E(y | x) = \beta_0 + \beta_1 x, \quad (18)$$

y la varianza es

$$Var(y | x) = Var(\beta_0 + \beta_1 x + \varepsilon) = \sigma^2. \quad (19)$$

Así, la medida de y es una función lineal de x , aunque la varianza de y no depende del valor de x . Además, ya que los errores no están correlacionados, las respuestas tampoco lo están. A los parámetros β_1 y β_2 se les suele llamar coeficientes de regresión. Estos tienen una interpretación simple y, frecuentemente útil. La pendiente β_1 es el cambio de la media de la distribución de y producido por el cambio unitario en x . Si el intervalo de los datos incluye a $x = 0$, entonces la ordenada al origen β_0 , es la medida de la distribución de la respuesta y cuando $x = 0$. Si se incluye al cero, β_0 no tiene interpretación práctica.

Estimación de los parámetros por mínimos cuadrados

Los parámetros β_0 y β_1 son desconocidos, y se deben estimar con los datos de la muestra. Supongamos que hay n pares de datos $(y_1, x_1), (y_2, x_2), \dots, (y_n, x_n)$.

Estos datos se obtienen de un experimento controlado, diseñado en forma específica para recolectarlos, o en un estudio observacional, o a partir de registros históricos existentes, en nuestro caso, en esta tesis se recolectaron espectros de cambios en concentración de glucosa y en bacterias.

Para estimar β_0 y β_1 se usa el método de mínimos cuadrados. Esto es, se estiman β_0 y β_1 tales que la suma de los cuadrados de las diferencias entre las observaciones y_i y la línea recta sea mínima, de forma que se puede escribir:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (20)$$

La ecuación (20) es un modelo de regresión, escritos en términos de los n pares de datos $y_i, x_i (i = 1, 2, \dots, n)$. Así el criterio de mínimos cuadrados es

$$S((b_0, b_1)) = \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2. \quad (21)$$

Los estimadores, por mínimos cuadrados, de β_0 y β_1 , que se designarán por $\hat{\beta}_0$ y $\hat{\beta}_1$, deben satisfacer

$$\left. \frac{\partial S}{\partial \beta_0} \right|_{\hat{\beta}_0, \hat{\beta}_1} = -2 \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0 \quad (22)$$

y

$$\left. \frac{\partial S}{\partial \beta_1} \right|_{\hat{\beta}_0, \hat{\beta}_1} = -2 \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) x_i = 0. \quad (23)$$

Se simplifican estas dos ecuaciones y se obtiene

$$n \hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^n x_i = \sum_{i=1}^n y_i, \quad (24)$$

$$\hat{\beta}_0 \sum_{i=1}^n x_i + \hat{\beta}_1 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n y_i x_i. \quad (25)$$

Las ecuaciones (24 y 25) son llamadas ecuaciones normales de mínimos cuadrados, su solución es la siguiente:

$$\hat{\beta}_0 = \bar{y}_i - \hat{\beta}_1 \bar{x} \quad (26)$$

y

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n y_i x_i - \frac{\left(\sum_{i=1}^n y_i\right)\left(\sum_{i=1}^n x_i\right)}{n}}{\sum_{i=1}^n x_i^2 - \frac{\left(\sum_{i=1}^n x_i\right)^2}{n}} \quad (27)$$

En donde:

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i \quad y \quad \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (28)$$

Son los promedios de y_i y x_i respectivamente.

El modelo de regresión lineal simple es, entonces:

$$\bar{y} = \hat{\beta}_0 + \hat{\beta}_1 x \quad (29)$$

La ecuación (29) produce un estimado puntual, de la media de y para una determinada x .

Como el denominador de la ecuación (27) es la suma corregida de cuadrados de las x_i y el numerador es la suma corregida de los productos cruzados de x_i y y_i , estas ecuaciones se pueden escribir en una forma más compacta como sigue:

$$S_{xx} = \sum_{i=1}^n x_i^2 - \frac{\left(\sum_{i=1}^n x_i\right)^2}{n} = \sum_{i=1}^n \left(x_i - \bar{x}\right)^2 \quad (27)$$

y

$$S_{xy} = \sum_{i=1}^n x_i y_i - \frac{\left(\sum_{i=1}^n y_i\right)\left(\sum_{i=1}^n x_i\right)}{n} = \sum_{i=1}^n y_i \left(x_i - \bar{x}\right). \quad (28)$$

Otra forma de escribir la ecuación (27) es

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}. \quad (29)$$

La diferencia entre el valor observado y_i y el valor ajustado correspondiente \hat{y}_i se llama residual.

Una vez que la calibración ha sido realizada, la predicción [53] (determinación de concentraciones en muestras desconocidas por regresión) de futuras muestras con concentraciones, es meramente una multiplicación de matrices. El método de regresión parcial de mínimos cuadrados es un método muy popular en aplicaciones industriales, investigaciones en ciencias e ingeniería que involucran el uso de variables controladas y/o fáciles de medir (factores) que explican, regulan o predicen el comportamiento de otras variables (respuestas). PLS es un método útil para la construcción de modelos de predicción cuando los factores son muchos y altamente colineales, el énfasis es sobre la predicción de las respuestas y no necesariamente en tratar de entender las relaciones existentes entre variables.

Así mismo, es una herramienta útil cuando el objetivo es la predicción y no existe una necesidad práctica de limitar el número de factores medidos. El análisis de regresión múltiple es una técnica estadística que puede utilizarse para analizar la relación entre una única variable criterio y varias variables independientes (predictores), el objetivo del análisis de regresión múltiple es usar las variables independientes cuyos valores son conocidos para predecir la única variable criterio seleccionada para el estudio.

PLS y PCA son técnicas de análisis multivariante, que se utilizan cuando las variables son muchas y altamente colineales [54]. PLS se utiliza para la construcción de modelos de predicción y es capaz de relacionar la concentración de la muestra, con los espectros medidos. En esta tesis PLS se utilizó para obtener la predicción de la concentración de glucosa, también para determinar agrupamientos entre muestras positivas o negativas de suero sanguíneo. El método PLS comprime las variables originales en un número de variables latentes o componentes principales (scores y loadings). La primer componente (PC1) tiende a describir la varianza sistemática en los datos (por ejemplo la relación entre espectros y concentraciones) y en menor extensión el ruido. Entonces cuando el número de “scores” se incrementa, la contribución al ruido se incrementa también, por lo cual podemos decir que no se está modelando la variancia sistemática en los datos, sino que se está modelando el ruido, resultando un modelo con una pobre capacidad de predicción. Por lo cual es de gran importancia seleccionar el número óptimo de componentes principales o “scores”. El número óptimo de PCs es calculado a partir de la raíz cuadrada del error cuadrático medio de la predicción o RMSEP (Root Mean Square error of Prediction) para las diferentes PCs usadas en el modelo PLS. Este valor del RMSEP es una medida de lo bueno del ajuste entre las concentraciones conocidas y las predichas.

$$RMSEP = \sqrt{\frac{\sum_1^N (Y_{pred} - Y_{ref})}{N}}. \quad (30)$$

Cuando el valor de RMSEP alcanza un mínimo, se dice que el número de componentes principales o “scores” correspondientes es óptimo. Existe un número límite de componentes principales, que pueden ser derivadas de una matriz X . El máximo número de componentes principales debe ser de $(n-1)$ o p número de variables, dependiendo de cuál sea menor.

El cálculo del RMSEP comúnmente se puede realizar mediante una validación cruzada LOOCV (Leave One Out Cross Validation) [55], o a través de la separación de un conjunto de datos. La técnica LOOCV es generalmente utilizada cuando solamente un número limitado de muestras está disponible para construir y validar el modelo. PLS es construido basado en las muestras remanentes.

Una vez construido, la muestra excluida es predicha y la desviación almacenada, este procedimiento se realiza hasta que todas las muestras son excluidas una vez, finalmente el valor RMSEP es la medida de las desviaciones almacenadas. En el caso del conjunto separado de datos, el modelo PLS es construido del llamado conjunto de entrenamiento y usado para predecir las muestras del llamado conjunto de predicción. Además del valor RMSEP, el coeficiente de correlación (r) es otro valor importante que nos indica la medida o grado de relación lineal entre dos variables, el cual es un número absoluto sin unidades.

Así, cuando la covarianza, $Cov(x, y)$ de dos variables es próxima a -1 o 1, se dice que existe una correlación lineal entre ellas. Los casos en donde la $Cov(x, y) = 1$, representa una correlación perfecta positiva, y para $Cov(x, y) = -1$, representa una correlación perfecta negativa. El concepto de correlación es fundamental en el análisis de regresión y describe la relación entre dos variables. Se dice que dos variables están correlacionadas si los cambios en una variable están asociados

con los cambios en la otra variable. Otra parte muy importante del análisis multivariante es el preprocesamiento de datos para la calibración de modelos de predicción, los cuales se describen a continuación.

1.4.3 Preprocesamiento de datos

A cualquier grupo de operaciones matemáticas que se le apliquen a los datos espectroscópicos, previo al análisis, es llamado preprocesamiento. El preprocesamiento de datos mejora la exactitud y robustece nuestros análisis cuantitativos, haciendo que los datos analíticos sean más fáciles de entender y procesar. El preprocesamiento puede separar la señal del ruido, tomando ventaja de las diferencias en sus propiedades.

Las propiedades utilizadas comúnmente para estos propósitos son: La información contenida de la transformada de Fourier del espectro, la intensidad de la variancia en la réplica del espectro, la intensidad de la variancia en un grupo de espectros teniendo diferentes concentraciones del analito y conocimiento o patrones de intensidad predecible en los espectros [56].

Centrado

El centrado de una variable consiste en la sustracción del valor medio a todos sus elementos. En el caso de los datos espectrales el centrado consiste en restar al espectro de cada una de las muestras el espectro promedio. Este pre tratamiento pone de relieve las diferencias entre espectros, al haber eliminado la tendencia común, es decir el espectro promedio [57].

Corrección de la línea base

Aparte del ruido contenido en el espectro o señal medida, también puede contener otras fuentes de variación no relacionadas con la propiedad de interés, a estas se les denominan variaciones de la línea base. El modelado explícito de la línea base, es un método que consiste en aproximar la línea base del espectro, con una

función polinomial, que se resta posteriormente al espectro, para obtener el espectro corregido.

El caso más sencillo sería el polinomio de grado 0 (constante), también denominada offset corrección y consiste en restar a cada espectro la absorbancia medida a una determinada longitud de onda. La selección de esta longitud de onda se puede hacer en base al conocimiento químico de la muestra, tomando la longitud de onda de menor variabilidad entre muestras, o la longitud de onda que al realizar un modelo PLS tiene un valor en “loadings” próximo a cero. Se puede utilizar polinomios de grado superior, para adaptar la corrección a las características de la variación de la línea base.

Derivadas

El uso de las derivadas permite diferenciar mejor los picos solapados y elimina desplazamientos lineales y cuadráticos de la línea base. La primera derivada se obtiene como la diferencia entre absorbancias, a dos longitudes de onda consecutivas. La derivada reduce la correlación entre variables, presenta el inconveniente de magnificar el ruido en la señal, por lo que habitualmente se utilizan métodos de suavizado de la señal. Derivadas de orden superior se obtienen aplicando sucesivamente la derivada, sin embargo no suelen utilizarse a orden superior a 2, ya que decrece la relación señal-ruido [58].

Normalizado

Es una familia de transformaciones, con el propósito de dar a los espectros, la misma escala, la normalización utilizada en esta tesis es la *mean normalization* (*normalización media*), donde el área bajo la curva de cada espectro es la misma.

Selección del conjunto de muestras de calibración

Para la selección del conjunto de calibración, también denominado conjunto de entrenamiento, se deben contemplar todas las fuentes de variabilidad del sistema [58], las cuales pueden ser físicas o químicas. Para este efecto se utilizan técnicas

de selección de muestras basadas en el análisis de agrupaciones, con el fin de seleccionar los espectros que representen en mayoría la población total.

La presencia de agrupaciones importantes en el conjunto de calibración puede conducir a errores si no son tomadas en cuenta, por lo que es importante analizar los datos con el fin de detectar alguna agrupación importante, y definir si es factible crear un modelo en general para todas las muestras o si es necesario dividir en grupos las muestras y crear modelos de calibración para los diferentes grupos. La detección de agrupaciones puede realizarse mediante simple inspección de los “scores”, al utilizar el análisis de componentes principales (PCA).

Establecimiento del modelo

En los modelos multivariantes, la concentración o cualquier otra propiedad de la muestra se modela en función de la respuesta, [59] en nuestro caso los espectros NIR, mediante la siguiente expresión.

$$c_k = \sum_{j=1}^p x_j b_{j,k} + e_k \quad (31)$$

Donde c_k es la concentración del analito k en la muestra, x_j es la respuesta de la muestra variable j , $b_{j,k}$ es el coeficiente que relaciona la variable p con la concentración del analito y e_k es el término del error no modelado por el modelo.

Utilizando una notación matricial, en la etapa de calibración con n muestras, puede escribirse como:

$$c = XB + e \quad (32)$$

Donde c es el vector de las concentraciones del analito o medio turbio, para las n muestras de calibración, X es la matriz de espectros de las n muestras en las p variables, B es el vector de los coeficientes de regresión y “ e ” es el error de los residuales.

Los métodos de regresión como PLS, solucionan los problemas de colinealidad que suelen ser común en los datos espectroscópicos, debido a que la matriz de respuestas X es descompuesta en factores ortogonales entre sí.

Detección de “Outliers”

Una de las ventajas del análisis multivariante, frente a otras técnicas de análisis univariante, es la capacidad de detectar “outliers” o mediciones inconsistentes con el resto de las mediciones. La detección de los “outliers” se lleva a cabo mediante información obtenida en la etapa del establecimiento del modelo, utilizando la información obtenida del “leverage” y los residuales, la detección de los “outliers” es importante, ya que estos degradan la capacidad predictiva del modelo. El “Leverage” es una medida de la influencia de una muestra en relación al modelo, motivo por el cual, muestras con “leverage” alto están muy alejadas del centro del modelo y tienen gran influencia en él. Este valor se calcula mediante la siguiente expresión

$$h_n = \frac{1}{n} + t_n^T (T^T - T)^{-1} t_n . \quad (33)$$

Donde t_n representa el vector de “scores” de la muestra n , T es la matriz de “scores” del modelo y N el número de muestras de calibración.

Los residuales son reflejo de la falta de ajuste al reconstruir el espectro con un cierto número de factores o componentes principales ($x = TP^T$), es decir

$$E = X - TP^T . \quad (34)$$

Validación del modelo

Debido a que los modelos de calibración como PLS no se apoyan directamente en un modelo teórico, y pueden incorporar variabilidad de los datos, no necesariamente relacionada con la propiedad de interés, deben ser validados cuantitativamente o cualitativamente. Esta validación puede realizarse analizando un grupo de muestras independientes al utilizado en la calibración, con el fin de que no exista un error sistemático entre las predicciones que realiza el modelo y los valores proporcionados por el método de referencia. También puede realizarse mediante una validación cruzada (Leave One Out Cross Validation). El valor del coeficiente de correlación (r) y el error cuadrático medio de la medición (RMSEP), son cantidades que nos indican el grado de concordancia entre las predicciones del modelo y los valores del método de referencia, por lo cual nos dan conocimiento de la validez del modelo para la predicción.

Importancia de los fluidos y muestras biológicas utilizadas en esta tesis

Glucosa en sangre:

La glucosa o dextrosa es un azúcar encontrada en las frutas y en la miel, es una hexosa, como muestra la figura 7, es decir que contiene 6 átomos de carbono ($C_6H_{12}O_6$), y es una aldosa, esto es, el grupo carbonilo está en el extremo de la molécula.

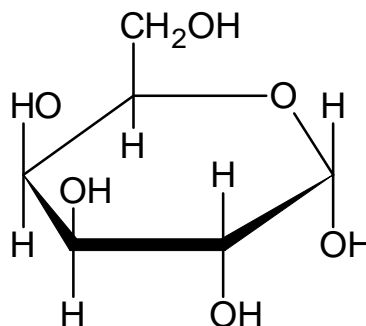


Figura 7. Molécula de Glucosa.

Todas las células del organismo necesitan que diariamente se les aporte energía para que puedan trabajar de manera adecuada, regular funciones vitales como: latido cardiaco, movimientos digestivos, respiración, y control de temperatura corporal, etc. [59]. La glucosa es una de las principales fuentes de energía, y es un elemento perteneciente al grupo de los carbohidratos, los cuales están formados por carbono, hidrógeno y oxígeno. La glucosa es el compuesto orgánico más abundante en la naturaleza, después de la celulosa [60]. El monitoreo de glucosa resulta de gran importancia para el control de funciones básicas, así como para el control de enfermedades relacionadas con el aumento o disminución de glucosa en sangre. Con la diabetes es de suma importancia el monitoreo frecuente de glucosa en sangre, es por ello la importancia de considerar variables importantes en la medición por medio de espectroscopias, para la disminución del error en la lectura, en esta tesis se consideró la variable de la temperatura como una variable que incide en las mediciones ya que influye en los resultados de predicción de cambios en concentración de glucosa.

Anticuerpos:

Actualmente la detección de anticuerpos es realizada por métodos indirectos como las pruebas serológicas, las cuales identifican el anticuerpo generado debido a la presencia del antígeno. Los anticuerpos o inmunoglobulinas son un grupo de glucoproteínas presentes en la sangre u otros fluidos corporales de los vertebrados, disponiendo de una forma idéntica que actúa como receptor de los linfocitos B y son empleados por el sistema inmunitario para identificar y neutralizar elementos extraños tales como bacterias, virus o parásitos.

Los anticuerpos son sintetizados por un tipo de leucocito denominado linfocito B, existen distintas modalidades de anticuerpo, basadas en la forma de cadena pesada que posean como se muestra en la figura 8, y se identifican como inmunoglobulinas: IgG, IgM, IgA, IgD, e IgE, estas 5 clases desempeñan funciones diferentes contribuyendo a dirigir la respuesta inmune adecuada para cada distinto tipo de cuerpo extraño que encuentren.

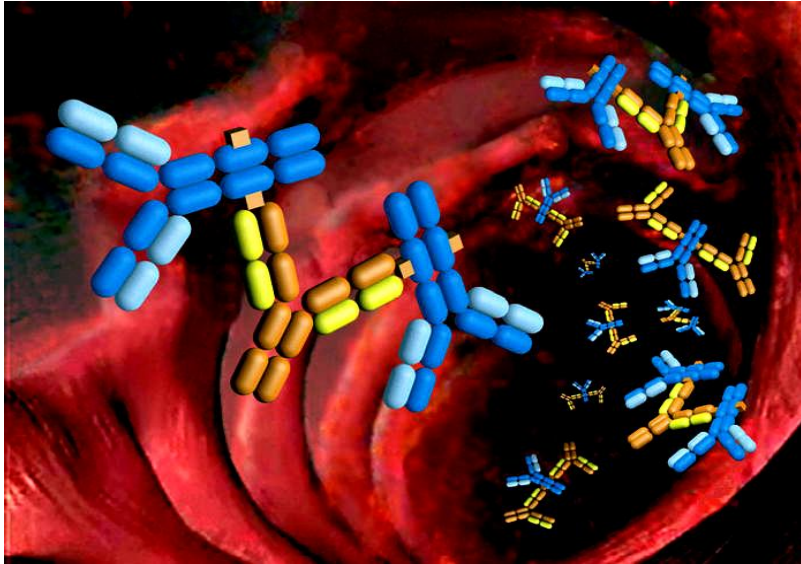


Figura 8. Anticuerpo presente en el sistema inmune que responde a la presencia de un antígeno.

Por otro lado, la espectroscopia Raman ha demostrado ser una herramienta poderosa en el estudio de muestras biológicas, ya que permite la caracterización de tejidos y fluidos de forma más rápida comparado con los métodos tradicionales de diagnóstico, debido a la mínima preparación de la muestra [61-64].

Exopolisacáridos (EPS)

Los EPS pueden ser considerados como el componente primario del biofilm, pueden estar asociados con iones metálicos y cationes bivalentes, tener carga neutra y carga polianiónica según el tipo de exopolisacárido [66], lo que le permite interactuar con diferentes antimicrobianos, de forma que estos pueden quedar atrapados en la matriz con capacidad para actuar sobre las bacterias.

La arquitectura de la matriz no es sólida por completo, las bacterias del biofilm viven en torrentes celulares que se extienden en forma tridimensional desde la superficie a la cual están adheridas. Estos torrentes están compuestos de diferentes células bacterianas, tanto aeróbicas como anaeróbicas, englobadas por exopolisacáridos, y separadas unas de otras por espacios intersticiales huecos, llamados canales de agua, como muestra la figura 9 que permiten el flujo de

líquido y actúan como sistema circulatorio primitivo para el transporte y difusión de nutrientes y oxígeno a las bacterias ubicadas en su interior, así mismo, contribuyen a un mecanismo para la remoción de productos de desecho metabólico [67-68].

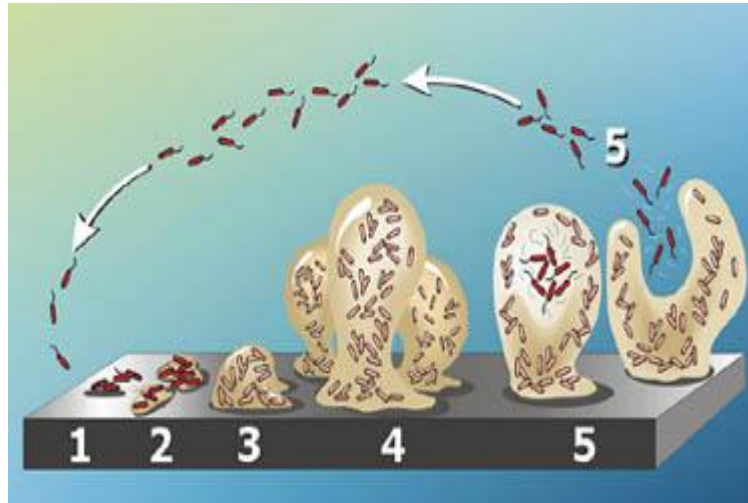


Figura 9. Formación del Biofilm

Es por esto que es de suma importancia el estudiar este microorganismo cuando se presenta en una comunidad tan especializada como lo es el biofilm, sobretodo conocer los grupos funcionales de la barrera de exopolisacáridos por el mismo microorganismo, ya que es esta misma la que impide que los antimicrobianos entren en contacto con la bacteria.

En el siguiente capítulo se describe la parte experimental realizada en esta tesis para las diferentes aplicaciones, como la calibración de modelos de predicción para la detección de concentración de glucosa en un phantom utilizando espectroscopia NIR, la metodología llevada a cabo para la detección de anticuerpos al toxoplasma Gondii, y los pasos llevados a cabo para la caracterización del biofilm, utilizando espectroscopía Raman, se describe también la instrumentación utilizada y los métodos de análisis multivariantes utilizados en estas aplicaciones.

2. METODOLOGÍA DE PREPARACIÓN DE MUESTRAS Y MEDICIONES ESPECTROSCÓPICAS.

2. METODOLOGÍA DE PREPARACIÓN DE MUESTRAS Y MEDICIONES ESPECTROSCÓPICAS

2.1 Instrumentación NIR

Se utilizó la espectroscopía de absorción en el cercano a infrarrojo (NIR) para determinar el error “in vitro” de la predicción de concentración de glucosa en un intralípido, debido al incremento de la temperatura, utilizando el espectrómetro Ocean Optics NIR512, con una lámpara de Tungsteno-Halógeno Ocean Optics LS-1 de 6.5 W, un medio turbio, y un termómetro, el diagrama del arreglo experimental se muestra en la figura 10.

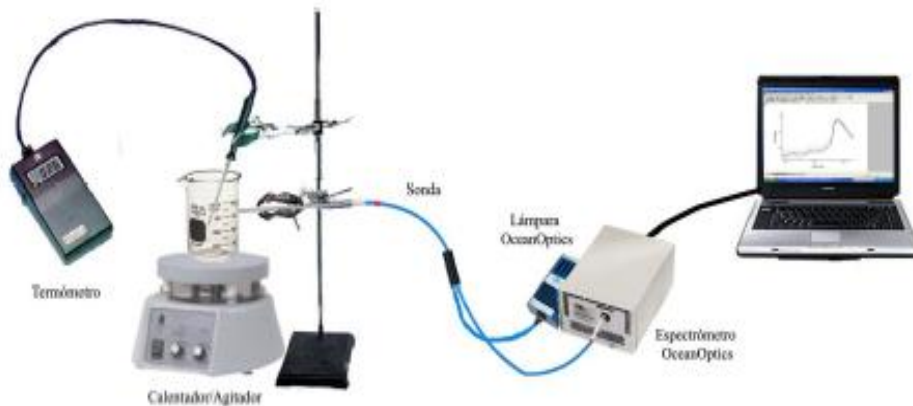


Figura 10. Diagrama esquemático del arreglo experimental para medición de cambios en concentración de glucosa y cambios en temperatura en phantom.

2.2 Instrumentación Raman

El equipo de espectroscopía Raman consiste en una fuente de Luz de radiación láser de 785nm a 300mW de potencia, la cual va a excitar a la muestra, la luz dispersada por la muestra es colectada por un espectrómetro, donde las frecuencias Raman son detectadas por una cámara CCD (de 256x1024 pixeles), la cual convierte las señales luminosas en señales eléctricas, para ser procesadas y finalmente observadas en una computadora, como se muestra en la figura 11. Este equipo se utilizó para la medición del Biofilm causado por la bacteria *E. faecalis*.



Figura 11. Fotografía del arreglo experimental para la espectroscopia Raman

Instrumentación Raman para la detección de anticuerpos

El equipo Raman utilizado en esta etapa de trabajo consiste en una fuente de radiación lineal polarizada de 514nm a 2.6W, la luz dispersada por la muestra es colectada por un espectrómetro, el cual se conecta a un filtro holográfico “Notch” (Kaiser Optical Systems HSPF-31453) para eliminar la radiación Rayleigh, posteriormente la radiación Raman se hace pasar por una rejilla de difracción o monocromador, para ser separada en distintas longitudes de onda, donde son detectados por una cámara CCD (de 256 x 1024 pixeles), la cual convierte las señales luminosas en señales eléctricas, para ser procesadas y finalmente observadas en una computadora, el diagrama esquemático se muestra en la figura 12.

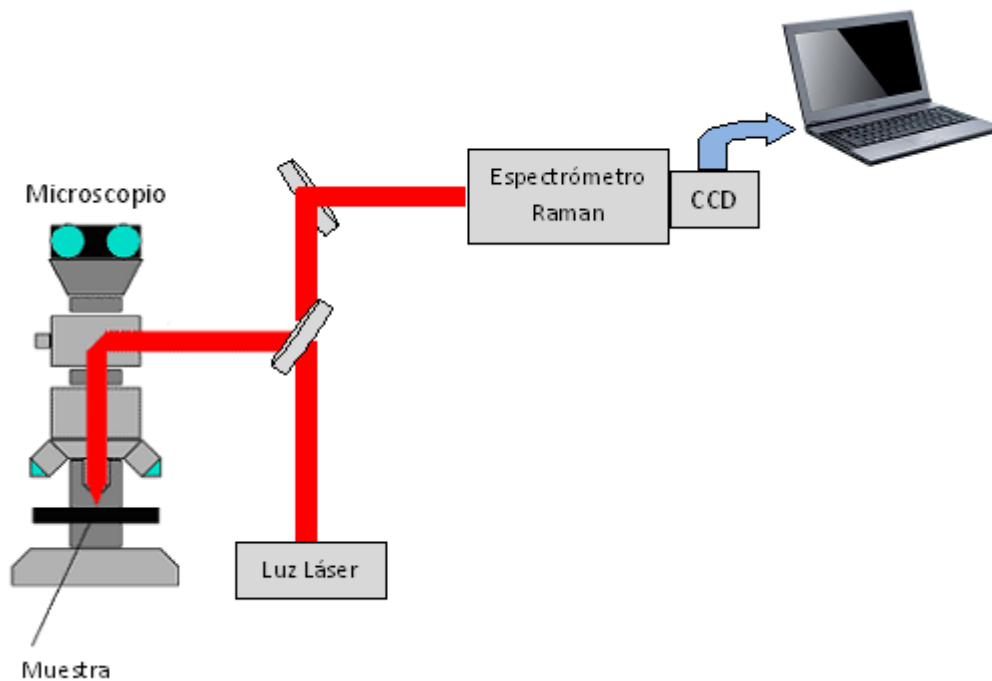


Figura12. Diagrama esquemático del sistema experimental para la obtención de los espectros con micro-Raman.

Para esta etapa de la tesis, además del espectrómetro Raman, se utilizó el equipo utilizado para la prueba serológica ELISA, para detectar los anticuerpos IgG, e IgM.

2.3 Mediciones con espectroscopia en el cercano a infrarrojo (NIR) para la determinación del error

Desarrollo experimental

Primera Etapa

Se preparó una solución base de intralípido, a la que se le tomaron varios espectros NIR modificando su concentración desde 0 a 5000mg/dL en intervalos de 500mg/dL, estos espectros se analizaron con la técnica de análisis multivariante PLS (*Partial Least Square*) y de componentes principales PCA (Principal component Analysis) para verificar que los cambios espectrales eran debido solamente a cambios en la concentración de glucosa.

Segunda Etapa

Se obtuvieron los espectros “in vitro” de un intralípido, variando solamente la temperatura desde 30°C a 50°C en intervalos de 2°C, los resultados se analizaron en PLS y PCA para verificar que los cambios espectrales son debidos únicamente a cambios en la temperatura.

Tercera Etapa

Se construyó un modelo de predicción de la concentración de glucosa, para una solución base intralípido a la que se le fue variando su concentración, desde 0 a 3500mg/dL, en intervalos de 500mg/dL, a temperatura constante de 30°C, a la última concentración de 3500mg/dL se le incrementó la temperatura hasta alcanzar los 48°C, en intervalos de 2°C, este último procedimiento se repite 5 veces, para después estimar la predicción de la concentración de la glucosa con los espectros obtenidos, utilizando el modelo de predicción calibrado para una temperatura constante.

2.4 Mediciones Raman para la detección de anticuerpos

Las 12 muestras de suero utilizadas en este estudio se obtuvieron de mujeres embarazadas del Hospital Materno Infantil y del Hospital Regional del área de ginecología y Obstetricia de la Cd. de León Gto., de las cuales 7 muestras resultaron negativas y 5 positivas a la presencia del anticuerpo IgG, anticuerpo específico contra la presencia del *Toxoplasma Gondii*, de acuerdo con los resultados de la prueba de ELISA (Diagnostic Automation, Inc. USA).

Cabe mencionar, que las muestras fueron obtenidas de un grupo de mujeres embarazadas que cumplía con las siguientes características: un nivel socio económico similar, ninguna resultado positiva a los anticuerpos IgM, IgA contra el *toxoplasma Gondii*. Y las que resultaron positivas, estuvieron en contacto con el parasito antes del embarazo, siendo por ello inmunizadas.

Por otro lado, el número de muestras utilizadas en este estudio es común dentro de las investigaciones de este tipo de enfermedades [68-70]. Para colectar el espectro Raman, una gota de la muestra es colocada en un substrato de aluminio, hasta tener una muestra sólida, se colectan 15 espectros por muestra desde diferentes zonas de la muestra, todos los espectros fueron tomados el mismo día bajo las mismas condiciones experimentales, los espectros Stokes fueron colectados en un intervalo de $450-1850\text{cm}^{-1}$, con una resolución de 2cm^{-1} .

Un total de 180 espectros Raman fueron colectados de las 12 muestras, de las cuales 105 espectros son negativos, y 75 a positivos. La adquisición de datos se realizó con el software GRAMS (versión 3.04, Thermo Galactic, USA), el cual también controlaba el instrumental. El preprocesamiento de datos como la corrección de la línea base, el suavizado, y normalización se realizó con el Origin (Versión 6, Microcal, USA), y las derivadas y el análisis de componentes principales se realizó con el software de Unscrambler TM (Versión 8. CAMO, Norway).

La corrección de la línea base se aplicó a los espectros Raman para eliminar la contribución de fluorescencia, utilizando la función polinomial de segundo grado con 15 puntos. Después los espectros fueron suavizados utilizando el método de suavizado por promedio adyacente, con 10 puntos para promediar. Los espectros fueron normalizados aplicando la transformación de máxima normalización. Finalmente, la segunda derivada de los espectros Raman se aplico utilizando el algoritmo Savitzky- Golay de función polinomial de segundo orden con 5 puntos de ventana, y previo al análisis multivariante los espectros Raman se centraron y escalaron con $1/SDev$.

Se utilizó análisis de componentes principales (PCA) para identificar el agrupamiento entre muestras positivas y negativas, así como el análisis de los loadings y PC's que nos muestran las regiones espectrales en donde se están dando las mayores intensidades, y conocer los grupos funcionales a que pertenecen estas frecuencias.

2.5 Mediciones Raman para el análisis del Biofilm

El equipo de espectroscopía Raman utilizado consiste en una fuente de Luz de radiación láser de 785nm a 300mW, que es la que va a excitar a la muestra, la luz dispersada por la muestra es colectada por un espectrómetro, donde las frecuencias Raman son detectadas por una cámara CCD (256 x 1024 pixeles), la cual convierte las señales luminosas en señales eléctricas, para ser procesadas y finalmente observadas en una computadora, la resolución del equipo es de $15cm^{-1}$

Se tomaron muestras de pacientes que acudieron a la clínica de la Maestría en Endodoncia de la UASLP, los cuales presentaban sintomatología dolorosa después del tratamiento de conductos. Las muestras se introdujeron en una cámara de anaerobiosis (MAIC/2000 COY Laboratory) se incubaron por 48 horas y se inocularon en placas de Agar de CDC anaeróbico y en placas de agar CDC anaeróbico con 5% de sangre de carnero (BBL Becton Dickinson de Mexico, SA de CV). Con las colonias desarrolladas se prepararon muestras para asegurar que se trataban de *E. faecalis*, se tiñeron con tinción de Gram y se observaron al microscopio de luz. A través de estas observaciones se pudo apreciar la presencia de cocos Gram positivos, que por medio de pruebas bioquímicas por MicroScan fueron identificados como *Enterococcus faecalis*. Una vez identificada la cepa se conservó por medio de resiembras periódicas en tubos de Tioglicolato enriquecido y placas de CDC anaeróbico en condiciones de anaerobiosis a 37 °C.

De esa manera se mantuvo en condiciones óptimas para su posterior utilización. Se utilizó un sistema de flujo continuo por goteo, el cual se incubó en cámara de anaerobiosis para simular las condiciones del sistema de conductos y observar el comportamiento del microorganismo sobre la dentina en condiciones *in vitro*. En la siguiente parte de esta tesis, se presentan los resultados experimentales obtenidos para la obtención del “error” en la predicción de concentración de glucosa debido al incremento de temperatura, así como las agrupaciones de las muestras que resultaron positivas o negativas al anticuerpo del toxoplasma gondii, también se muestra los resultados de las asignaciones encontradas para el biofilm formado por la bacetria *E. faecalis*.

3. ANÁLISIS MULTIVARIANTE Y DISCUSIÓN DE RESULTADOS

3 ANÁLISIS MULTIVARIANTE Y DISCUSIÓN DE RESULTADOS

3.1 Determinación del error “*in vitro*” con espectroscopia NIR

Como hemos mencionado anteriormente, el monitoreo continuo de glucosa en sangre es importante en pacientes con diabetes, y actualmente las mediciones de cambios en concentración de glucosa se realizan utilizando métodos electroquímicos o enzimáticos por medio de la extracción de sangre utilizando agujas o lancetas, razón por la cual se han venido desarrollando nuevos métodos menos invasivos, como las espectroscopias, que nos presentan un método no invasivo y sin el uso de muchos reactivos, sin embargo ninguno de los métodos ha tenido éxito con la exactitud requerida para uso clínico, ya que es importante considerar varios factores que influyen en la medición, tales como posición de la sonda, grosor de piel, temperatura, etc. En esta parte de la tesis se describen los experimentos realizados para tomar en cuenta uno de estos factores importantes a considerar al momento de realizar una medición con espectroscopias para la predicción de cambios en concentración de glucosa, este factor es la temperatura, variable importante en las mediciones con espectroscopia NIR.

Resultados de la primera etapa

Se utilizó un phantom de tejido, ya que puede usarse como medio dispersor de luz en NIR, ya que tiene las propiedades ópticas muy similares a las de tejido humano [72], como se muestra en la figura 13.

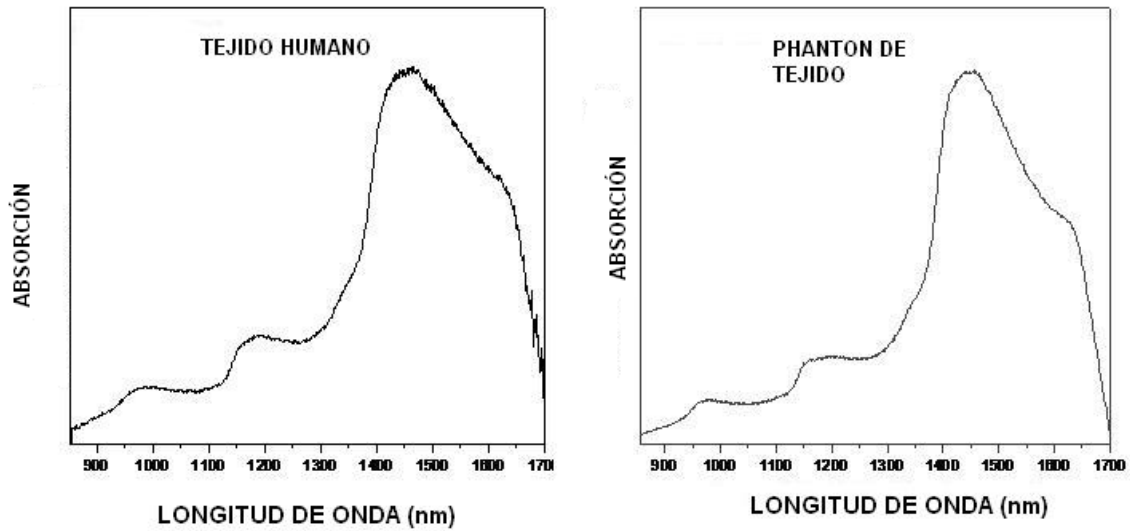


Figura 13. Espectro NIR de tejido humano y del phantom de tejido.

Al phantom de tejido se le cambió la concentración desde los 0mg/dL hasta los 5000mg/dL, en intervalos de 500mg/dL, manteniendo la temperatura constante a 30°C, como se puede apreciar en la figura 14, donde observamos que hay un desplazamiento vertical del espectro debido al incremento de la concentración de la glucosa.

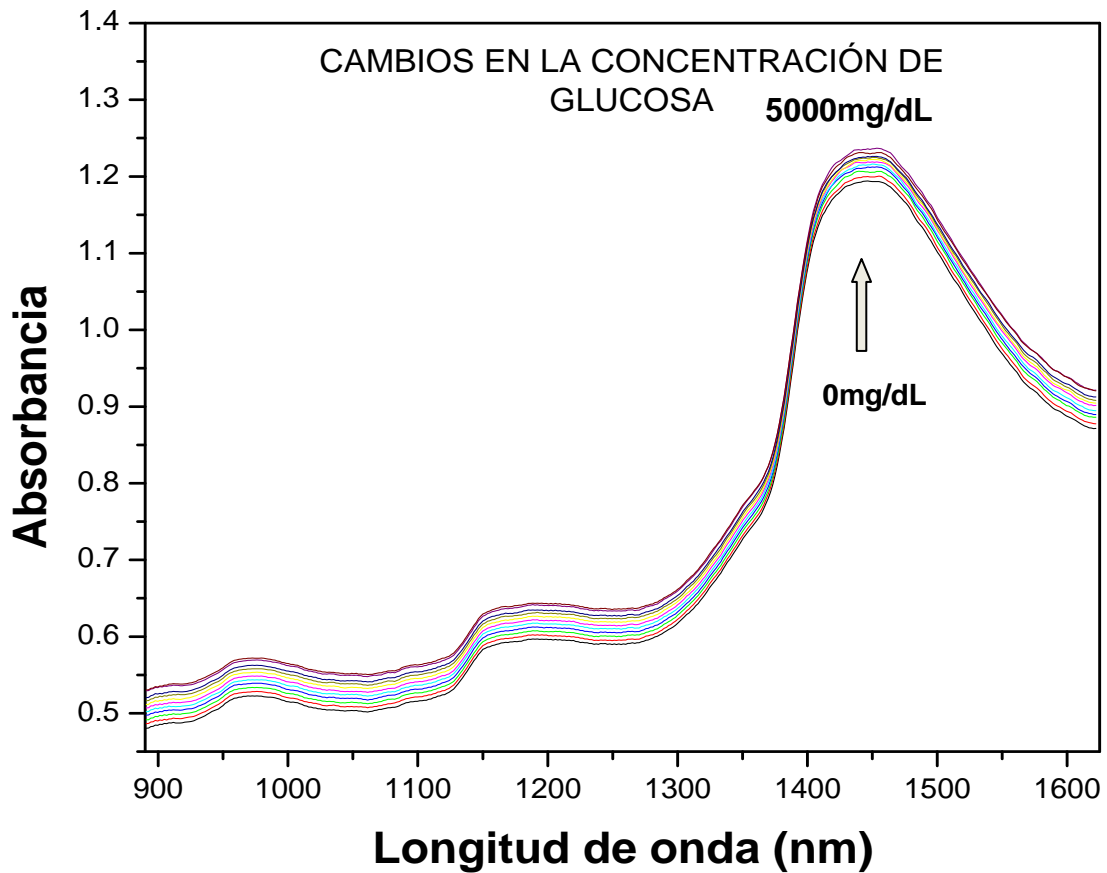


Figura 14. Espectro de absorción NIR para diferentes concentraciones de glucosa.

Utilizando un modelo de regresión PLS para la predicción de la concentración de glucosa, encontramos un coeficiente de correlación r^2 de 0.96 y un RMSEP de 132.43 mg/dL. Este coeficiente de correlación nos indica que los cambios espectrales están fuertemente relacionados con los cambios en la concentración de glucosa, como se aprecia en la figura 15.

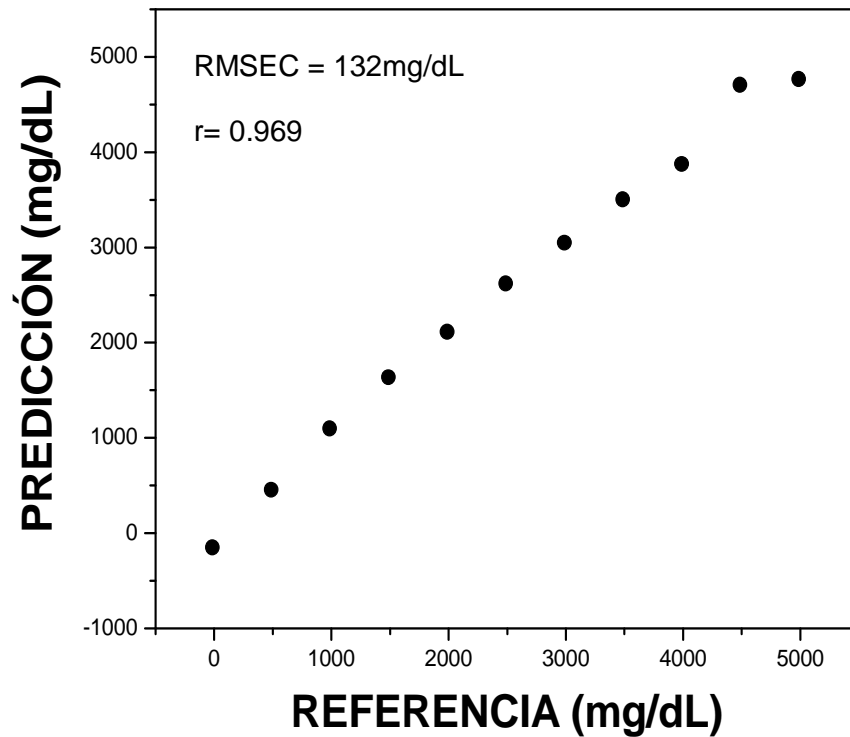


Figura 15. Relación de la predicción de la concentración de glucosa y la de referencia.

Una vez que se obtienen los espectros para diferentes concentraciones de glucosa, se realizó un análisis de componentes principales, este análisis nos indicó que los cambios espectrales están relacionados únicamente con cambios en concentración de glucosa, ya que se obtiene una relación de PC1 y PC2 donde PC1 explica el 100% de la varianza total asociada con el incremento de la concentración de glucosa, como se muestra en la figura 16.

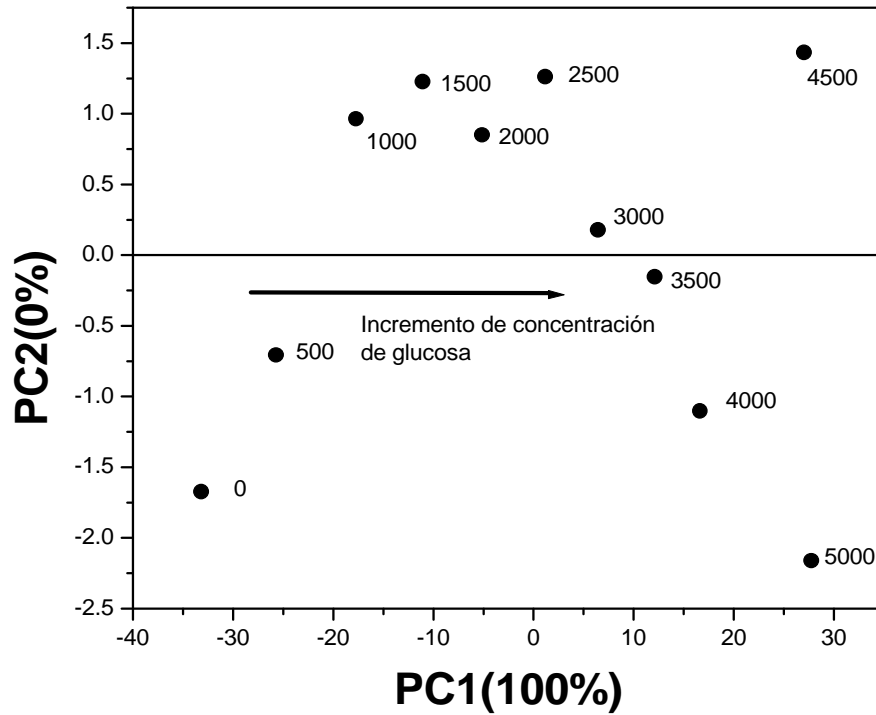


Figura 16. PC1 &PC2 para diferentes concentraciones de glucosa.

Resultados de la segunda etapa

En esta segunda etapa, se busco ver los cambios espectrales debido únicamente a cambios en la temperatura, así que se utilizó un phantom sin azúcar, al que se le fue variando la temperatura desde 30°C a 50°C, en intervalos de 2°C, se observó un cambio en la intensidad de absorbancia, debido únicamente al incremento de la temperatura como se aprecia en la figura 17.

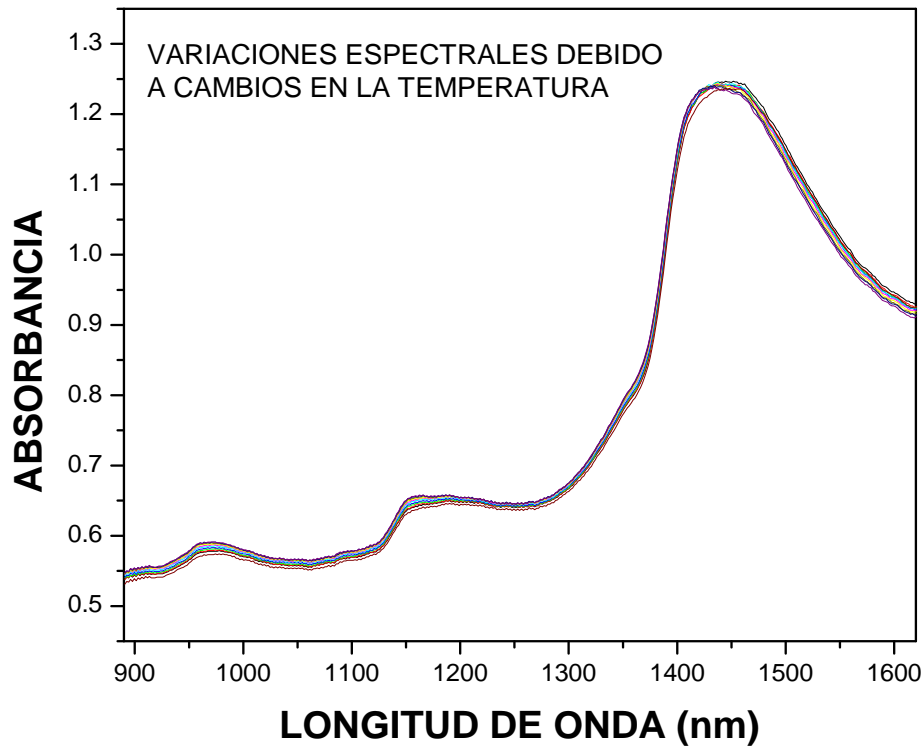


Figura 17. Espectros de absorción NIR de la solución del phantom para diferentes temperaturas.

Con los resultados de los espectros en NIR, para diferentes temperaturas, se realizó un análisis de componentes principales, para identificar que los cambios espectrales se estaban dando debido únicamente al cambio en la temperatura, se graficó la relación de PC1 contra PC2, donde PC1 explica el 91% de la varianza total, y PC2 el 9%, como se muestra en la figura 18, estos resultados nos indican que los cambios espectrales están muy relacionados con los cambios en temperatura del Phantom.

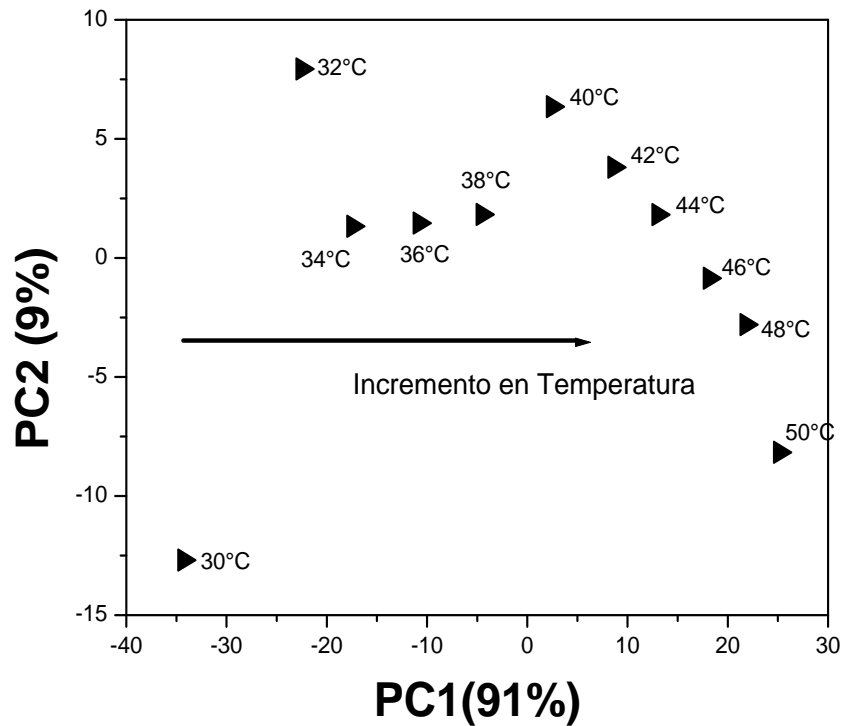


Figura 18. Relación de PC1 versus PC2 para cambios de temperatura en los espectros NIR del phantom.

Para corroborar todavía más estos resultados, se construyó un modelo de regresión PLS con los resultados de los espectros para diferentes temperaturas, encontrando un coeficiente de correlación r^2 de 0.99 y un RMSEC de 0.49°C , este coeficiente nos indica que las variaciones espectrales están fuertemente relacionadas con los cambios de temperatura, como se aprecia en la figura 19.

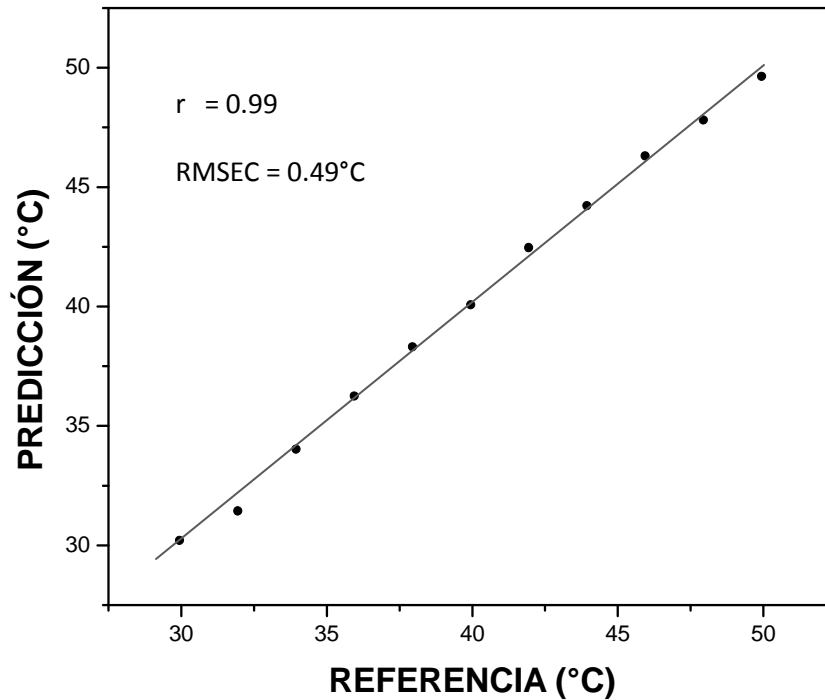


Figura 19. Relación entre las temperaturas de referencia y las de predicción.

Resultados de la tercera etapa

El desarrollo de esta siguiente etapa, nos da respuesta a si la temperatura influye en la predicción de cambios en concentración de glucosa, para ello, se cambió la concentración de glucosa del phantom desde 0 a 3500mg/dL en intervalos de 500mg/dL, a una temperatura constante de 30°C, para 8 mediciones. A la última concentración (de 3500mg/dL) se le incrementó la temperatura en intervalos de 2°C, hasta alcanzar los 48°C, este procedimiento se realizó 5 veces, con la misma concentración de 3500mg/dL, después la concentración de glucosa fue estimada para cada espectro, al que se le incrementó la temperatura, utilizando el modelo de predicción calibrado para una temperatura constante, los resultados de las predicciones de la concentración de glucosa, en base al modelo de la predicción calibrado para una temperatura constante, se muestran en la tabla 2.

Tabla 2. Resultados de la predicción en las concentraciones de glucosa, para diferentes temperaturas.

Temperatura (°C)	Predicción de concentración (mg/dL)				Predicción promedio (mg/dL)
	Experimento 1	Experimento 2	Experimento 3	Experimento 4	
30	4919	5051	4955	4753	5011.33
32	5218	5170	5160	5623	5282.25
34	5338	5475	5097	5537	5366.25
36	5466	5542	5491	5619	5535.75
38	5543	5718	5374	5700	5589.25
40	6255	5796	5439	5892	5851.5
42	6017	5879	5442	5918	5825.75
44	6036	6034	5508	5919	5867.25
46	6170	5793	5532	5916	5585.75
48	6238	6123	5590	5924	5985
50	6422	6117	5771	5849	6007.5

Con los resultados obtenidos en la tabla 2, se calculó el error medio inducido por la temperatura en las predicciones de las concentraciones de glucosa, el cual lo obtuvimos de la siguiente forma:

$$E_{1^{\circ}C} = \sum \left(\frac{1}{\Delta T} \right) \sqrt{(x_i - x_0)^2}$$

Donde $E_{1^{\circ}C}$ es el error inducido por $1^{\circ}C$, X_0 es el valor de la predicción de glucosa en T_0 , X_i es el valor de predicción de glucosa en T_{fin} , y T_{fin} es la temperatura final para la i -ésima medición, T_0 es la temperatura original, y n es el número de predicciones obtenidas.

El error por grado centígrado, se calculó de la siguiente forma:

$$E_{T1^{\circ}C} = \sum \left(\frac{1}{N} \right) \sum_1^N (E_{1^{\circ}C}) \quad (35)$$

Donde N es el número de experimentos realizados, $E_{T1^{\circ}C}$ es el error en la temperatura por grado centígrado.

Los resultados muestran que un incremento en la temperatura de $1^{\circ}C$ induce un error en la predicción de la concentración de glucosa de 80.5 mg/dL.

El error sistemático, o el Bias, representa la diferencia entre los valores medidos y los valores de predicción, se calcula:

$$E = \left(\frac{1}{N} \right) \sum_1^N (y_p - y_o) \quad (36)$$

Donde \hat{y}_p es la predicción promedio para cada 2°C de temperatura y y_o es la medición, es decir 5000mg/dL, obteniéndose un error Bias de 627.98 mg/dL contra el error reportado de Masatoshi [73] de 500mg/dL, para el error de predicción de concentración de glucosa por simulación, estos resultados se realizaron de forma experimental y nos indican que si influye el incremento de la temperatura en la medición de cambios en concentración.

Conclusiones particulares del error inducido debido al incremento de la temperatura:

Con los resultados de los cambios espectrales debido a variaciones en la temperatura, encontramos un r^2 de 0.99, por lo que podemos decir que están fuertemente relacionados solamente con los cambios en la temperatura.

De forma similar, con los resultados obtenidos para los cambios en la concentración de glucosa, encontramos un r^2 de 0.99, lo cual nos indica que los cambios espectrales sí están fuertemente relacionados únicamente a los cambios en concentración de glucosa.

Encontramos también que el error en la predicción de la concentración de glucosa debido al incremento de la temperatura de 1°C es de 80.5 mg/dL, por lo que podemos decir que la espectroscopía NIR es sensible a cambios en la temperatura, y que la temperatura representa un factor externo de importancia que debe ser considerado a la hora de calibrar modelos de predicción.

3.2 Utilización de la espectroscopia Raman y análisis multivariante para la detección de anticuerpos

Como se mencionó anteriormente, los diagnósticos clínicos para identificar infecciones en el organismo humano, actualmente se realizan por métodos serológicos, los cuales detectan los anticuerpos específicos a un agente agresor como una bacteria o virus, estas técnicas requieren de algunos reactivos y de tiempo para realizarse, en esta parte de la tesis mostramos los resultados realizados con espectroscopia Raman como una técnica alternativa para la detección de anticuerpos, y el uso del análisis multivariante para identificar en los resultados semejanzas o diferencias entre muestras. Cabe mencionar que los resultados se comparan y se toma como referencia la prueba de ELISA que actualmente se utiliza con este fin.

Resultados

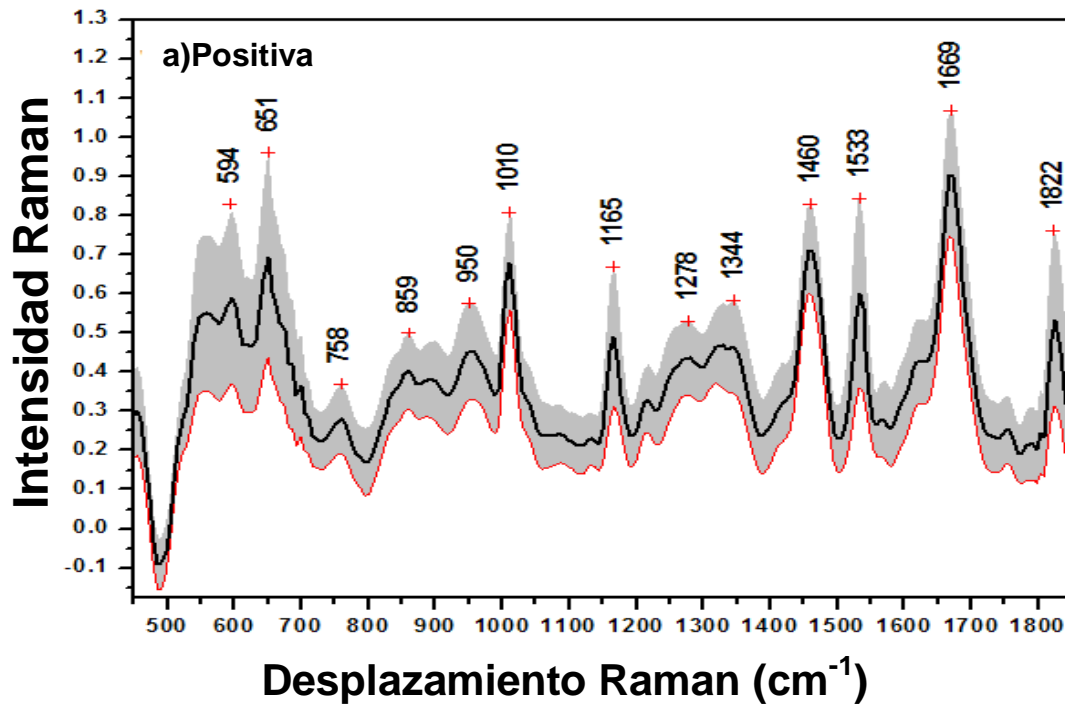
Antes de realizar un análisis multivariante, los espectros Raman colectados fueron analizados visualmente para comparar diferencias espectrales entre las muestras, tales como diferencias entre picos, desplazamiento entre bandas o cambios en intensidad, los cuales podrían indicar la presencia o ausencia de anticuerpos del T. Gondii.

La figura 20 muestra el promedio de los espectros Raman, y las desviaciones estándar de las muestras positivas y negativas, respectivamente.

La diferencia más importante entre las muestras positivas y negativas se relaciona con cambios en la intensidad Raman, ya que no existen bandas adicionales o desplazamientos observados al comparar el promedio de los espectros de muestras positivas con las negativas.

Con respecto a las bandas que se observan en el espectro promedio de las muestras positivas, las regiones de 1533 y 1166 cm^{-1} son aproximadamente 1.5 veces más grande que el promedio espectral de las muestras negativas. La región de 1533 cm^{-1} está asociado con el modo vibracional de la molécula beta-caroteno, la cual puede estar relacionada con la presencia del anticuerpo del *Toxoplasma Gondii*. El pico en la región 1166 cm^{-1} se asigna a modos vibracionales de la tirosina y triptophan, del grupo de los aminoácidos que forman las proteínas, presentes en la estructura de los anticuerpos.

Identificando las regiones en las que se estaban dando cambios en intensidad (1166, y 1533 cm^{-1}), se recorto la región de (450-1850 cm^{-1}) a un rango de (1150-1560 cm^{-1}) para realizar un análisis de componentes principales (PCA).



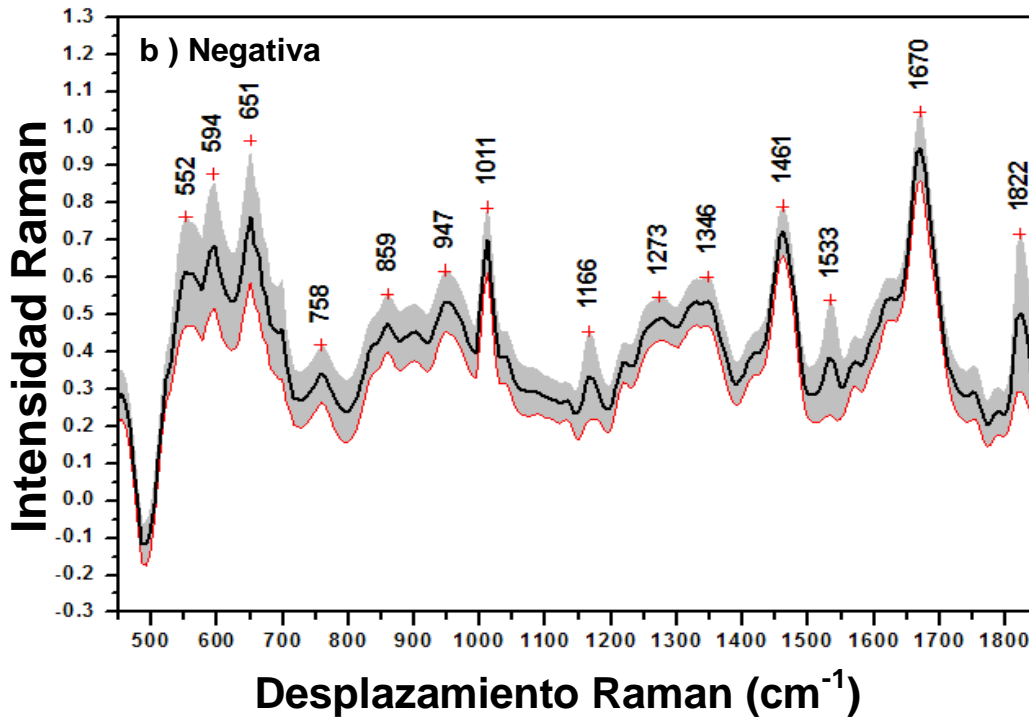


Figura 20. Espectro promedio de los resultados a) positivos y b) negativos al toxoplasma Gondii, la región sombreada representa la desviación estándar.

Como hemos mencionado, PCA es una técnica de análisis multivariante, que reduce la dimensionalidad de los datos, utilizada para determinar agrupaciones, patrones, similitudes o diferencias entre las muestras. La primer componente principal (PC1) explica la mayor cantidad obtenida de la varianza total. El siguiente factor (PC2) explica la mayor cantidad obtenida de la varianza total no contenida en PC1, los sucesivos factores explican cada vez menos información. Los scores capturan la estructura de las filas (muestras) permitiendo establecer relaciones entre muestras, y la detección de muestras discrepantes y agrupaciones entre ellas. El uso de las derivadas sucesivas consigue diferenciar mejor los picos solapados, y elimina desplazamientos lineales, por lo cual utilizamos la segunda derivada de los espectros Raman para obtener una discriminación de las muestras.

Las figuras 21 y 22 muestran los resultados de la 2ª derivada para PC1&PC2, PC1&PC3, PC1&PC4, y PC1&PC5, donde PC1 representa el 10%, PC2 el 8%, PC3 el 5 %, PC4 el 4%y PC5 el 2% de la varianza total. Los bajos porcentajes en la varianza se atribuyen a la complejidad de la información de los espectros.

Las graficas de PCA muestran una clara discriminación entre muestras positivas y negativas. La línea continua indicada en las graficas fue calculada para evaluar la sensibilidad y especificidad del método. Esta línea se obtuvo utilizando el software de Matlab (versión 6.1, Math Works, USA) utilizando discriminación lineal, utilizando la programación cuadrática para dos clases.

Con los resultados obtenidos del análisis de PCA se calculó la sensibilidad y la especificidad de diagnóstico donde:

La sensibilidad es la probabilidad de clasificar correctamente a un individuo enfermo, es decir, la probabilidad de que para un sujeto enfermo se obtenga en la prueba un resultado positivo. La sensibilidad es, por lo tanto, la capacidad del test para detectar la enfermedad.

La especificidad es la probabilidad de clasificar correctamente a un individuo sano, es decir, la probabilidad de que para un sujeto sano se obtenga un resultado negativo. En otras palabras, se puede definir la especificidad como la capacidad de detectar a los sanos.

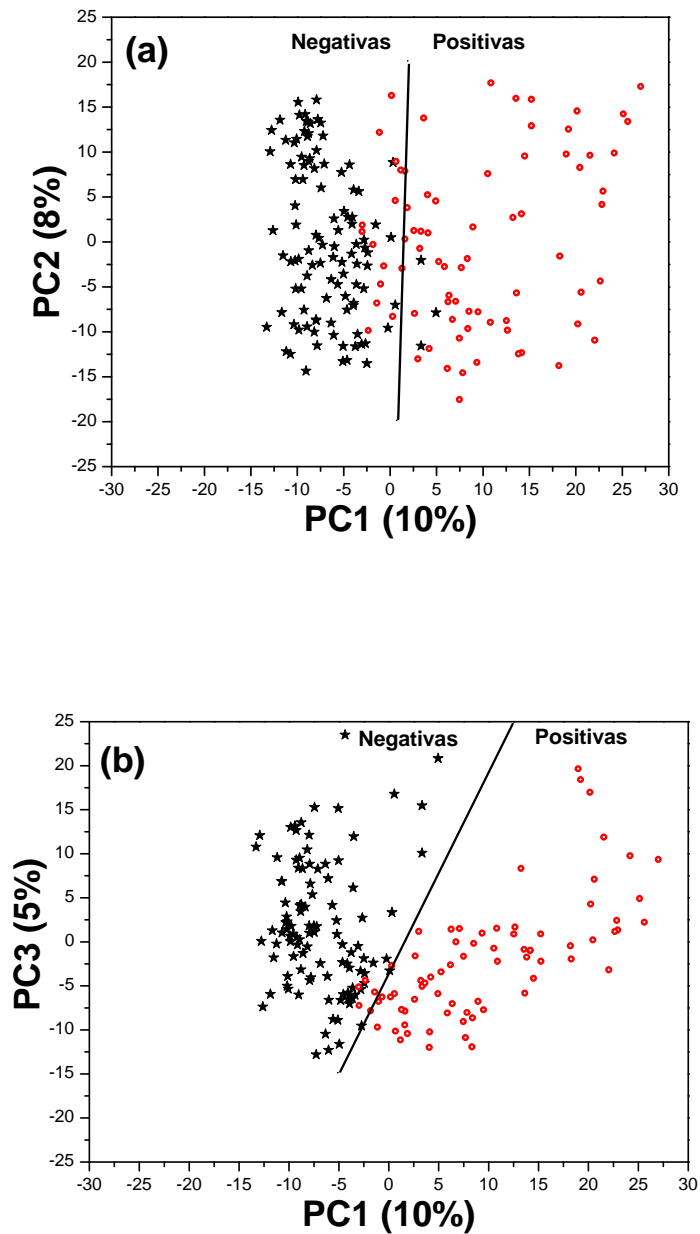


Figura 21. Análisis de componentes principales (PCA) de los espectros de suero con presencia y ausencia del anticuerpo al T.gondii, a) PC1&PC2, b) PC1&PC3.

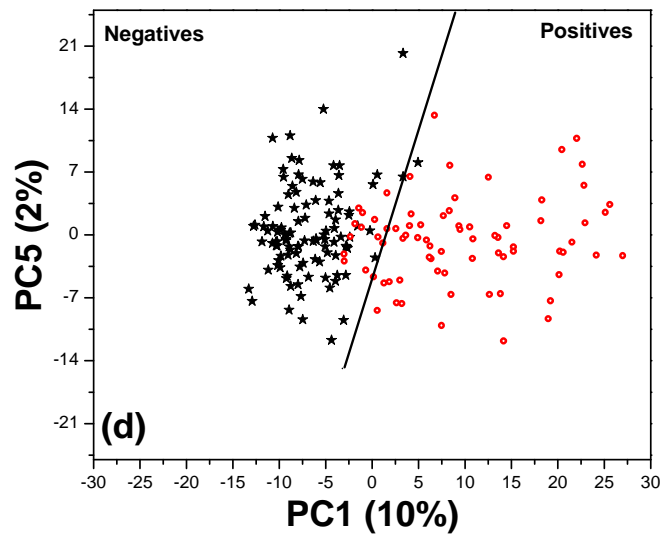
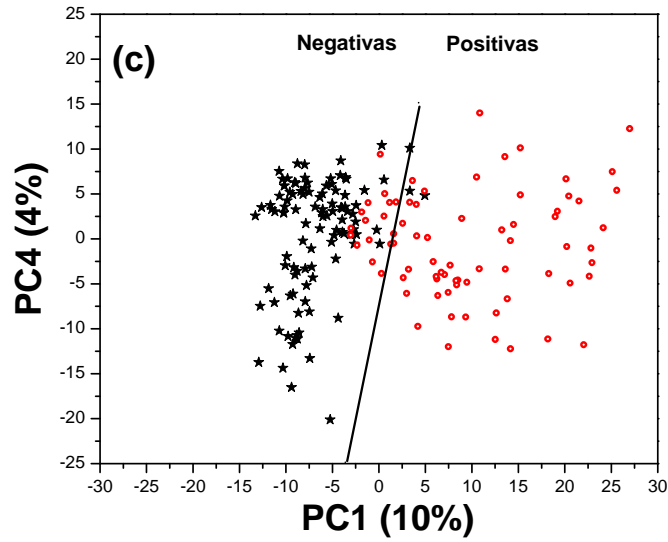


Figura 22. Análisis de componentes principales (PCA) de los espectros de suero con presencia y ausencia del anticuerpo al T.Gondii c) PC1&PC4, d) PC1&PC5.

Los valores de la sensibilidad y la especificidad fueron calculados con las siguientes fórmulas:

$$\text{Sensibilidad} = \left(\frac{T_{(P)}}{F_{(N)} + T_{(P)}} \right) \times 100 \% \quad (37)$$

$$\text{Especificidad} = \left(\frac{T_{(N)}}{F_{(P)} + T_{(N)}} \right) \times 100 \% \quad (38)$$

Donde:

$T_{(P)}$ representa clasificaciones para ambas técnicas (espectroscopia y la prueba de ELISA) como positivas.

$F_{(P)}$ corresponde a la clasificación como positiva por espectroscopia y negativa por la prueba de ELISA.

$F_{(N)}$ corresponde a una clasificación negativa por espectroscopia y positiva por la prueba de ELISA.

$T_{(N)}$ clasificada por ambas técnicas (espectroscopia y la de ELISA) como negativas.

Los resultados de la sensibilidad y la especificidad obtenidos del análisis de componentes principales, se muestra en la tabla 4.

Tabla 4. Valores experimentales de la Sensitividad y especificidad de los PC´s

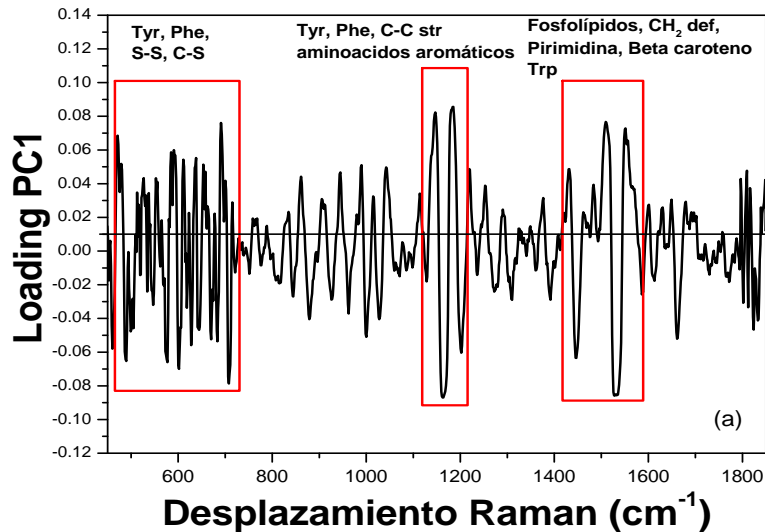
Scores	Sensitividad(%)	Especificidad(%)
PC1&PC2	84	97
PC1&PC3	95	98
PC1&PC4	78	98
PC1&PC5	85	99

Después de analizar las graficas de los componentes principales, con la información relacionada con la discriminación entre muestras, se procedió a realizar un análisis de las graficas de loadings para PC1, con el fin de localizar las regiones espectrales asociadas con la presencia o ausencia de anticuerpos al toxoplasma Gondii. Las graficas de los loadings pueden definirse como una relación entre las variables originales y las componentes principales, que retienen la relación entre las variables (longitudes de onda), por lo que mostramos en donde se dan con mayor intensidad esos picos.

Las figuras 23(a) y (b) son las graficas de los loadings de los PC1 y la media de los espectros Raman. Las tres regiones enmarcadas en rectángulos localizados en: $450-700\text{cm}^{-1}$, $1100-1200\text{cm}^{-1}$, y $1400-1600\text{cm}^{-1}$, fueron identificados como aquellos que tienen la mayor contribución en las variables. Las más altas intensidades en los loadings representan las diferencias significativas para la descrimicación [74].

Estas tres regiones corresponden a las asignaciones de Tirosina(Tyr), de fosfolípidos (Phe), S-S- y C-S, C-C, aminoácidos aromáticos, CH₂, pirimidina, beta-carotenos y Tryptophan (Trp), de acuerdo a lo reportado con las frecuencias Raman para anticuerpos por Darwin y Pézolet, [75-76]. Estas tres zonas corresponden a frecuencias vibracionales asociados con aminoácidos presentes en las cadenas de los anticuerpos de acuerdo a lo reportado por Tsuboi [77].

Estos resultados sugieren que la información contenida en PC1 está fuertemente relacionada con la información espectral de los anticuerpos presentes en las muestras de suero, y la influencia de estos grupos para poder discriminar entre muestras positivas y negativas al T. Gondii.



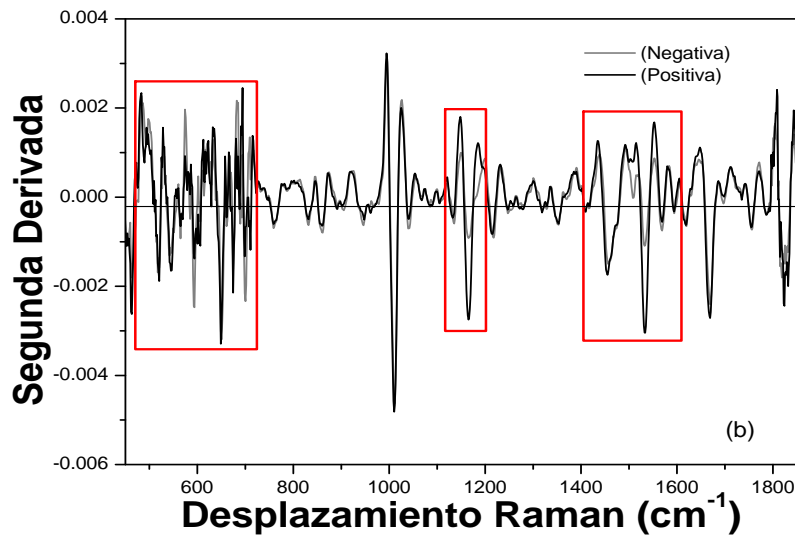


Figura 23. Relación entre los loadings de los espectros Raman derivados, y los valores de PC1. (a) Los picos con mayor amplitud representan la ubicación de las diferencias significativas entre las muestras de suero positivas y negativas, (b) espectro promedio derivado de las muestras de suero positivas y negativas, donde las regiones encerradas en los recuadros representan las frecuencias donde se ubicaron los más importantes cambios para el modelo y discriminar entre las muestras de suero

Algunos de los desplazamientos Raman que obtenemos de los loadings, nos muestran amplitudes en donde se dan las mayores intensidades Raman y representan las asignaciones mostradas en la tabla número 5, las bandas obtenidas en los espectros Raman corresponden a carbohidratos, proteínas y aminoácidos contenidos en los anticuerpos.

Tabla 5. Asignación de las frecuencias para los Loadings.

Desplazamiento Raman (cm-1)	Asignación
713 (650-730)	Hidrocarburos [78]
851 (789-861)	Proteínas de ácidos nucleicos [79]
1037 (100-1300)	C – C [78]
1154 (1140-1300)	P=O [79]
1174 (1162-1184)	Typ, Phe, Ácidos amino aromáticos, C- C [79]
1445 (1250-1500)	CH ₂ , CH ₃ , Betacaroteno, Pirimidina, Fosfolípidos [80]
1548 (1500-1690)	Triazinas [78]
1660 (1656-1690)	Amida I [78]

Conclusiones específicas para la detección de anticuerpos:

La espectroscopia Raman es una técnica con la que es posible detectar la presencia de anticuerpos en suero sanguíneo, para los antígenos específicos al toxoplasma Gondii. Con PLS es posible predecir muestras de diferentes concentraciones, y con PCA la primera componente describe el mayor porcentaje de la varianza total de los datos, y tanto con PLS y PCA es posible distinguir entre muestras positivas y negativas.

Obtuvimos también una alta sensibilidad y efectividad con los resultados para PLS y PCA, lo que indica una distinción entre muestras positivas y negativas.

Con los resultados de los Loadings, observamos que los mayores cambios en intensidad se dan en los desplazamientos Raman correspondientes a polisacáridos, betacarotenos, citosinas, guaninas, y lípidos C-O.

Podemos decir que Raman es una técnica alternativa para la detección de anticuerpos y que junto con las técnicas de análisis multivariante de PLS y PCA podemos diferenciar entre muestras positivas y negativas, además de poder predecir diferentes concentraciones de IgG, Igm, e IgA.

3.3 Espectroscopia Raman para el análisis del Biofilm

Como habíamos mencionado, los biofilms bacterianos son en sí comunidades bacterianas englobadas en una matriz de exopolisacáridos, y hasta el momento los resultados obtenidos para identificar esta matriz de exopolisacáridos en el área de endodoncia era por medio de microscopia electrónica de barrido, por lo cual se introduce en esta parte de la tesis el uso de la espectroscopia Raman como una herramienta para identificar los grupos funcionales que forman dicha matriz. El siguiente espectro mostrado en la figura 24 nos muestra el resultado del espectro Raman del Biofilm, con algunos picos interesantes, correspondientes en general a la presencia de estructuras químicas de azúcares, como se esperaba y se confirmó también con los resultados del microscopio electrónico de barrido en la figura 25. Cabe mencionar como relevancia que muchos estudios del Biofilm se han basado en observaciones al MEB (Microscopio Electrónico de Barrido)) sin definir químicamente la estructura de la matriz, en este estudio se utilizó la espectroscopia Raman para definir los grupos funcionales que forman el Biofilm, ya que su utilización en el área odontológica ha sido muy escasa a pesar de sus bondades que ofrece esta técnica, Spencer et al [25] utilizó este sistema para leer estructura dentinaria posterior a la aplicación de adhesivos dentinarios y Kwon et al [24] para ver cambios en la composición del hidróxido de calcio después de la colocación intraconducto en casos de necrosis.

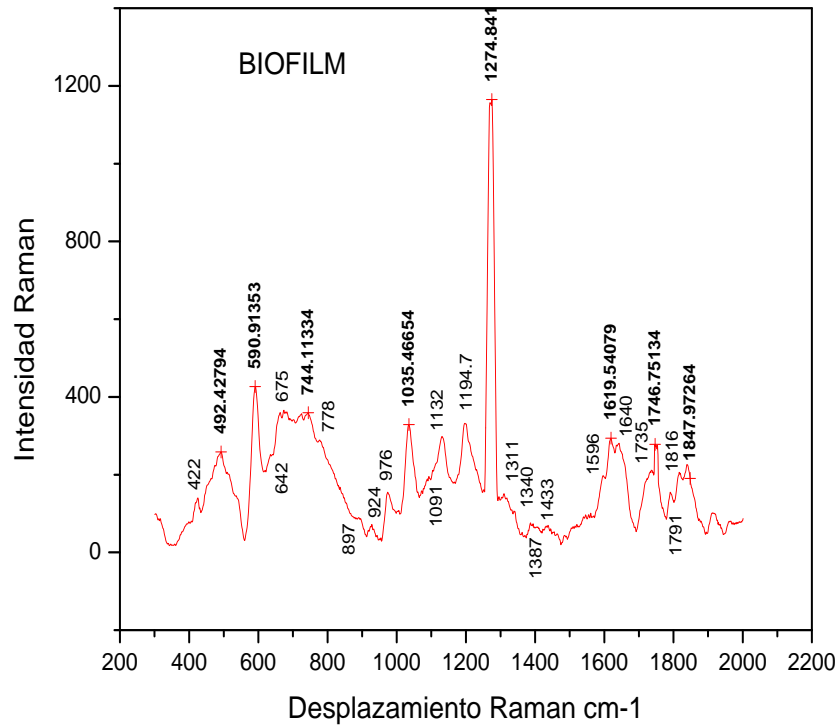


Figura 24. Espectro Raman del Biofilm

Los resultados del espectro Raman del Biofilm nos muestran los grupos funcionales que producen las bacterias *E. faecalis*.

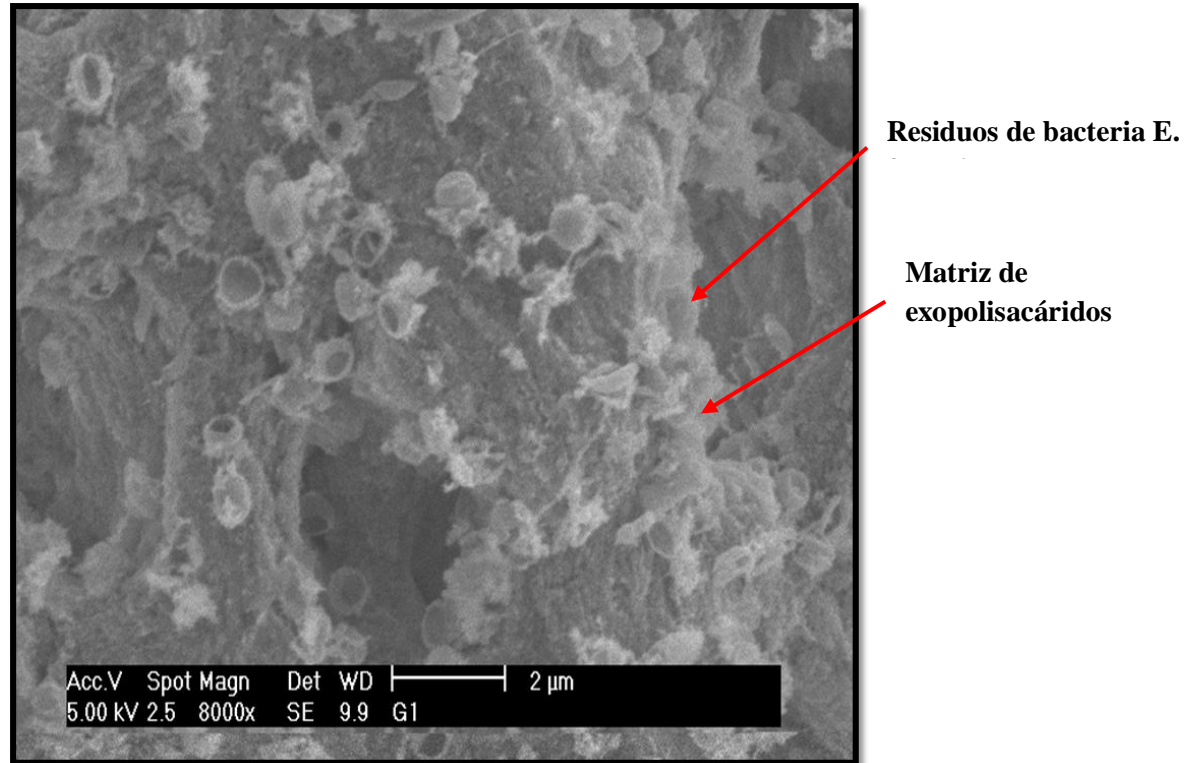


Figura 25. SEM del Biofilm construido por la bacteria *E. faecalis* a 2μm.

Los resultados del SEM (Scanning Electron Microscopy) muestran la presencia de residuos de bacterias *E. Faecalis*, así como también se ve claramente la formación del biofilm y los canales de irrigación típicos en una matriz de exopolisacáridos.

En la tabla núm. 5 se muestran los resultados de las asignaciones correspondientes a algunas frecuencias del espectro Raman de la matriz de exopolisacáridos producidos por la bacteria *E. faecalis*, los cuales coinciden con los resultados que muestra el SEM.

Tabla 5. Asignaciones Raman para el espectro Raman del biofilm producido por la bacteria *E. faecalis*.

FRECUENCIA RAMAN EN cm-1	ASIGNACIONES RAMAN
422	(C-C-O) + (C-C-C) [81]
492	C-C-C, C-C-O, C-C [82]
514	(C-C-O) + (C-C-C) [82]
541	En 540cm-1 esta C2- C(1)- O1 [82]
590	De 200 a 600 esta la presencia de C-C-C, y de C-C-O, C-C y C-O Ref. Manual pag 416 o bien Grupo OH [82]
642	Movimiento CH o (hidrocarburos) [78]
675	(500-700) movimiento de (triple enlace) CH (650-730) hidrocauburos 578-700 acetilenicos con triple ligadura [78]
744	(700-900CM-1)Movimiento CH aromáticos [78]
778	Movimiento CH aromáticos, o de

	(776-790cm ⁻¹ esta la citosina, uracil) [78]
897	Movimiento CH [78]
924	Esta banda puede ser asignada como un modo par que consiste en la contribuciones C-O stretching, C-C-H, C-C, C-C-O o bien C(1) H y C-O-H [82]
976	Hidrocarbonos [78]
1035	P-O-grupo alquilo [78]
1091	C-C, o bien Acido nucleico O-P-O [79]
1132	C-O [78]
1194.7	C-O [78]
1274	C(6)-OH y C(1)-OH Se ha reportado banda de glucosa en

	1273 C(6)-OH [82]
1311	CH ₂ , proteínas, o ácidos nucleicos, o bien CH ₂ y movimiento CH [78]
1340	Vibración CH ₂ . [82]
1387	CH ₂ y CH ₃ [79]
1433	(1400-1500)se encontró CH ₂ y CH ₃ con movimiento asimétrico, o bien hidrocarbonos [78]
1596	C=C y C=N o bien olefinas [78]
1619	1630 1690)Contiene componentes C=N aunque algunos tipos que absorben en esta región son guaninas [78]
1640	C=C y C=N,[77]
1735	C=O, muy aproximadamente los grupos carbonilos en esteres saturados absorben cerca de 1740 y los aldehídos en 1725

	[78]
1746.7	(1550-1900cm-1) grupos carbonilos dentro de esteres saturados y 1725 para aldehídos [78]
1791	(1150-1900cm-1) C=O componentes carbonilos [78]
1816	(1150-1900cm-1) C=O componentes carbonilos [78]
1847cm-1	(1150-1900cm-1) C=O componentes carbonilos [78]

Conclusiones específicas para el análisis de biofilm

En los resultados del espectro del Biofilm, la región más importante está entre los 700 y 1700 cm^{-1} , porque los grupos característicos de los azúcares se presentan en esta región, y como parte relevante el pico de 1274 cm^{-1} corresponde a el grupo funcional C_6OH , descrito como una cetohexosa, característico de azúcares, de tal forma que las asignaciones encontradas para la matriz del Biofilm formado por la bacteria *E. faecalis* corresponden a los grupos funcionales de los azúcares como se esperaba, así como la presencia de citosina y uracil en 778 cm^{-1} indicadores de la presencia de ADN propio de la bacteria, y en 924 cm^{-1} , C-C, C-H, C-C-O característico de carbohidratos, resultados que se corroboran con las imágenes del microscopio Electrónico de barrido, donde se observa la matriz de exopolisacáridos y residuos de bacterias.

En el siguiente capítulo se describen las conclusiones generales de este trabajo de investigación doctoral, en función de los múltiples resultados obtenidos para las diferentes aplicaciones, en la parte final se anexan perspectivas para este trabajo, el título de un artículo publicado durante el transcurso y elaboración de la tesis, anexos y las referencias bibliográficas.

CONCLUSIONES

La espectroscopia Raman y el análisis multivariante son herramientas útiles para el análisis de muestras biológicas, de tal forma que hemos podido identificar el error introducido en el incremento de 1°C en temperatura, para la determinación de concentración de glucosa es de 80.5mg/dL. Cabe mencionar que este es un resultado experimental in vitro, utilizando un phantom como base para cambios en concentración de glucosa. El error sistemático o Bias de estos resultados experimentales es de 627.9 (mg/dL) contra lo reportado en un sistema de simulación por Masatoshi, que reporta un error de 500 (mg/dL) por 1°C. Este resultado es de relevancia cuando se quiere predecir los cambios en concentración de glucosa, ya que incluir el factor de temperatura es importante en los sistemas de espectroscopias.

Con respecto a los resultados de identificación de anticuerpos por medio de espectroscopia Raman, fue posible identificar diferencias en intensidades Raman entre muestras positivas y negativas de anticuerpos específicos al *Toxoplasma Gondii*, comparando los resultados con los análisis bioquímicos como referencia de oro, encontramos con espectroscopia Raman un 96% de sensibilidad y un 100%. De acuerdo con estos resultados, la Espectroscopia Raman es una herramienta alternativa para la determinación de presencia o ausencia de anticuerpos de una forma directa y sin la utilización de muchos reactivos como los que se utilizan actualmente en análisis bioquímicos, ya que los resultados de los espectros Raman de las muestras nos muestran información de los grupos funcionales encontrados en los anticuerpos como aminoácidos, fosfolípidos, carbohidratos, betacarotenos, etc. Los cuales fueron utilizados para identificar la presencia o ausencia de anticuerpos mediante análisis de componentes principales (PCA).

Con respecto a los resultados del estudio del biofilm, hasta ahora muchos estudios se han basado solo en la observación al MEB (Microscopio Electrónico de Barrido), sin definir químicamente la estructura de la matriz, en este estudio, se utilizó la espectroscopia Raman con la que fue posible identificar los grupos funcionales que conforman al Biofilm, donde las regiones de las asignaciones Raman encontradas corresponden a carbohidratos y residuos de ADN debido a la presencia de bacterias en la superficie del biofilm. Estos resultados se corroboraron con los resultados obtenidos con el microscopio electrónico de barrido, donde se observan los residuos de bacterias *E. faecalis*, y el biofilm ya maduro formando una capa de exopolisacáridos.

Podemos decir que las técnicas espectroscópicas y los métodos de análisis multivariante son técnicas confiables para la identificación o cuantificación de fluidos biológicos, en base a los resultados obtenidos, comparando las referencias de oro utilizadas en los análisis bioquímicos y los resultados de los datos espectroscópicos de las muestras biológicas.

PERSPECTIVAS DEL TRABAJO

En futuros trabajos se evaluará la factibilidad de la espectroscopia Raman y análisis multivariante para discriminar entre muestras biológicas con anticuerpos para diferentes antígenos, Por ejemplo el de Citomegalovirus, o para discriminar entre diferentes isotipos de anticuerpos, como el IgG, IgM e IgA, en muestras biológicas. Cabe mencionar que ya se inició a manera de prueba, un análisis entre diferentes cepas para la bacteria *E. coli*, y un análisis para biofilms formado por diferentes bacterias, con la *E. faecalis* y con *Candida*, utilizando espectroscopia Raman, y donde se perciben diferencias en los resultados de componentes principales.

REFERENCIAS BIBLIOGRAFICAS:

1. Cesar Mello, *Rapid differentiation among bacteria that cause gastroenteritis by use of low-resolution Raman spectroscopy and discriminant analysis*. Anal Bioanal Chem., 2005, 383:701-706.
2. Kerstin Rebrin, Garry M. Steil, William P. Van Antwerp, John J. Mastrototaro, *subcutaneous glucose predicts plasma glucose independent of insulin : implications for continuous monitoring*, Am. J. Physiol., E561-E571, 1999.
3. H. Peter Chase, Laura M. Klim, Susie L. Owen, Todd A. Mackenzie, Georgeanna J. Klingensmith, Robert Murtfeldt, Satish K. Garg, *Continuous Subcutaneous Glucose Monitoring in Children with Type 1 Diabetes*, Pediatrics, 107:2, 222-6, 2001.
4. Hopkins George W. and Mauze Ganapati R., *In vivo NIR diffuse reflectance tissue spectroscopy of human subjects, technical report HP laboratories*. Palo Alto 1999.
5. Yoon Gilwon, Amerov Airat K, Jeon Kye *determination of glucose concentration in scattering medium based on selected wave lengths by use of an overtone absorption band*, Appl. Opt. 2002, Vol 41, pp 1469-1475.

6. Malin Sthepen F. Et al. *Non invasive prediction of glucose by near infrared diffuse reflectance spectroscopy*. Clinical chemistry, 1999, 45:1651-1658.
7. Fan XH, Ashara T. Miyata., Ohdan H., Diagnostic value of non invasive near-infrared spectroscopy to asses liver viability of brain-dead donors, Transplant Proc, 1999; 31(7): 2922-2923.
8. Cristopher V. Eddy and Mark A. *Near infrared spectroscopy for measuring urea in hemodialysis fluid*, Clinical Chemestry, 2001.
9. A. Bañsari, G. Poli, V. Molina, M. Dovis, E- Petruzzelli, A. Boniolo, E. Rolleri, J. Clin. Pathol. 33(1980) 640-643.
10. A.B. Sabin, H.A. Feldman, Sciencie 108 (1948) 660-663.
11. B.C. Walton, B.M. Benchoff, W. H. Brooks, Am J Trop Med Hyg 15 (1966)149-52.
12. B.R. Dannemann, W.C. Vaughan, P. Thulliez, J.S. Remington, J. Clin. Microbiol. 28(1990) 1928-1933.
13. Lasa I., *Biofilms Bacterianos. Actualidad, Col 28Num 2, 2004*
14. Biofilms. *Medicina, Buenos Aires, vol-65, No. 4 2005*
15. Porternier I., Haapasalo H, Orstavik D. et al, *Inactivation of antibacterial activity of iodine potassium iodide and chlorhexidine digluconate against enterococcus faecalis by dentin, dentin matrix, type-1 collagen, and heath-killed microbial whole cells, JOE 2002, vol-28(9)634-637.* Nair P.N. *On the causes of persistent apical periodontitis. A review, IEJ, 39, 249-281, 2006.*
16. *Biofilm Bacterianos: Cronificación, 2005*
17. Lasa I., del Pozo. *Biofilms bacterianos e infección, An sist. Saaint Navar, 2005, Vol.28(2): 163-175.*
18. Love R.M., et al *Enterococcus faecalis- a mechanism for its role in endodontic failure, IEJ, 34, 399-405, 2001.*
19. Distel J.W. et al., *Biofilm Formation in Medical Root Canals, JOE 2002, Vol 28, No. 10 689-693*

20. Stuart Charles H., et al., *Enterococcus faecalis: Its Role in Root Canal Treatment Failure and Current Concepts in Retreatment*
21. Chavez de Paz L. *Response to alkaline stress by root canal bacteria in biofilms, IEJ, 40, 344-355, 2007*
22. Craig J., et al, *Comparrison of the antimicrobial efficacy of 1.3% NaOCl/Biopure MTAD to 5.25% NaOCl/15% EDTA for Root canal irrigation, JOE, Vol. 33, No. 1, January 2007.*
23. Kwon T:Y. et al, *FT-Raman spectroscopy of calcium hydroxide medicament in root canals, IEJ, 37, 489-493, 2004.*
24. Spencer P., *Molecular Structure of Acid-etched Dentin smear layers. In situ Study J Dent Res 80(9):1802-1807, 2001*
25. Douglas A. Skoog, F. James Holler, Timothy A. Nieman, *Principios de Análisis Instrumental, Edit Mc Graw Hill Quinta edición, 2001*
26. Raúl Gómez y Rogelio Murillo, *Espectroscopia Infrarroja <http://sistemas.fciencias.unam.mx>*
27. Borman, S.A. *Anal Chem. 933 (1984).*
28. Araujo Andrade Cuauhtémoc, Facundo Ruíz, Martínez Mendoza José Refugio, *Feasibility for non invasive estimation of glucose concentration in newborns using NIR spectroscopy and PLS, Medical Physics: Ninth Mexican symposium on Medical Physics. AIP conference Proceedings, volumen 854, pp 224-226 (2006).*
29. S. Khall, *Spectroscopic and Clinical Aspects of noninvasive glucose measurements, Clinical Chemistry 45:2, 165-177 (1999).*
30. J.S. Remington, P. Thulliez and J.G. Montoya, *“Recent Developments for Diagnosis of Toxoplasmosis”, J. Clin. Microbial. 42, 941-945(2004).*
31. A.B. Sabin H. A. Feldman, *“Dyes as microchemical indicators of a new immunity phenomenon affectings a protozoan parasite (toxoplasma)” Science 108, 660-63 (1948).*

32. Reed, Z.C. Feng, K.t. Yue and T.S. Gansler, "Raman spectroscopic characterization of human breast tissues: implications for breast cancer diagnosis" *Appl. Spectrosc.* 47,787-791(1993)
33. Naes T. and Martens H. *Multivariate calibration II. Chemometric methods. Trends in analytical chemistry*, 3 266-271 (1984).
34. Naes T. and Isaksson T., *Selection of samples for calibration in near infrared spectroscopy, I. General principles illustrated by example.*, *Appl Spectrosc* 43 328-335 (1989).
35. Hobart H. Willard, Lynne L. Merritt, Jr. John A. Dean Frank A Settle, Jr. *Métodos Instrumentales de Análisis. Grupo Editorial Iberoamérica.*
36. José M. Pastor, Tariq Jawhari y Juan Carlos Merino, *Caracterización de Materiales Polímeros mediante la utilización de la microespectroscopia vibracional (Raman, FT- Raman)*, *Revista Iberoamericana de Polimeros*, Vol 4(3), Septiembre de 2003.
37. Ibelings M. et al, *rapid identification of Candida spp. In peritonitis patients by Raman spectroscopy*, *Clinical Microbiology and Infection*, Vol 11(5) May 2005.
38. Venkatakrishna K. Optical pathology oral tissue: A Raman spectroscopy diagnostic method, *Current Science*, Vol 80, (5), March 2001
39. Roberto Y. Sato Berrú, Cirilo Medina Guti *Revista internacional de contaminación ambiental*, enero/marzo 2004, vol. 20 número 001 Universidad Nacional Autónoma de México, Distrito Federal, México pp.17-24.
40. Tsuda H, Arends J., *Raman Spectroscopy In Dental Research: A Short Review Of Recent Studies*, *Adv Dent Res.* 11(4): 539-547 Nov 1997
41. Kwon T:Y. et al, FT-Raman spectroscopy of calcium hydroxide medicament in root canals, *IEJ*, 37, 489-493, 2004.

- 42.Spencer P., Molecular Structure of Acid-etched Dentin smear layers. In situ Study J Dent Res 80(9):1802-1807, 2001
- 43.Chris H. et al., Infection Central, Dental Clinics of North America, Oct. 1984, Vol. 28 Num 4
- 44.Douglas A. Skoog, F. James Holler, Stanley R. Crouch, *Principios de Análisis Instrumental*, Edit. Cengage Learning, sexta edición.
- 45.Thorniley M.S., Wickramasinghe Y.A.P.D., and Rolfe P, *Near infrared spectroscopy: a new technique for non invasive monitoring of tissue and blood oxygenation in vivo*, *Biochem. Soc. Trans* 16 62-63 (1988).
- 46.Alejandro Tomás, Andrea Bellincontro, *Aplicación de la espectroscopia Infrarroja cercana(NIR) al estudio de la deshidratación de Uva CV, Montepulciano, V congreso Iberoamericano de tecnología postcosecha y agroexportaciones 2007, (S3-1007).*
- 47.Angel Dago Morales, Reinaldo Fernández Fernández, María Dolores Ruíz Martínez, *La espectroscopia infrarroja y el método de calibración multivariada de mínimos cuadrados parciales en la predicción del índice de octano experimental en gasolinas*, Revista CENIC ciencias Químicas, Vol 37, No. 1, 2006.
- 48.D A Boas, G Strangman, JP Culver, R D Hoge, G Jaszewski, *Can the cerebral metabolic rate of oxygen be estimated with near-infrared spectroscopy*, *Phys, Med. Biol.* 48 (2003) 2405-2418.
- 49.T.D. Ridder, S.P. Hendee and C.D. Brown, *Noninvasive Alcohol Testing using diffuse Reflectance Near Infrared Spectroscopy*. Applied Spectroscopy, Volumen 59, número 2, 2005
- 50.Laura Inés Nieto, Ana Silvia Donolo, Amado Javier Bava, Empleo de la espectroscopia infrarroja y la transformada de Fourier para diferenciar bacterias de importancia clínica, *Acta Bioquím Clin Latinoam* 2004; 38 (3) 289-95.

51. Frank, L.E & Friedman, J. H. (1993), *A statistical view of chemometrics regression tools Technometrics*, 35, 109-148.
52. Joseph F. Hair, Rolph E. Anderson, *Análisis Multivariante, 5a Edición, Editorial Prentice Hall, 1999*
53. Pelletier M.J., *Quantitative analysis using Raman spectrometry, Applied spectroscopy* 57 20A-42A(2003).
54. Parisi A.F., Nogueiras L. and Prieto H., *Anal Chem. Acta* 238, 95 (1990).
55. Kim H. Esbensen, *An introduction to multivariate and data analysis and experimental design, Aalborg University, Esbjerg, 5th ed. CAMO (2002).*
56. Naes T. and Martens H. *Multivariate calibration II chemometric methods. Trends in analytical chemistry*, 3 266-271(1984).
57. Naes T. and Isaksson T., *Selection of samples for calibration in near infrared spectroscopy. I. General principles illustrated by example, Appl spectrosc* 43 328-335 (1989)
58. Bereton R.G. *Introduction to multivariate calibration in analytical chemistry, Analyst*, 125 2125-2154(2000).
59. Ahonen, Jarmo, *Kinesiología y anatomía aplicada a la actividad física, Paidotribo, 1996, ISBN 8480193077.*
60. Carrillo, Leonor, *Energía de Biomasa, S.S. 2004, ISBN 9874386797.*
61. Sabin, A. B., and Feldman, H.A. 1984, *Science*, 108, 660-663.
62. Walton, B.C., Benchoff, B.M., and Brooks, W. H. 1966, *Am J. Trop. Med. Hyg*, 15, 149-52.
63. Balsarini, A. Poli, G. Molina, V. Dosis, M. Petrosselli 1980 *J. Clin Pathol.*, 33, 640-643.

64. Dannemann, B.R. Vaughan and Remington, J.S. 1990, J. Clin. Microbiol., 28, 1928-1933.
65. Donlan RM., Biofilms: Microbial life on Surface. Emerg Infected, Dis 2002, 8(9):881-890
66. De Beer D. Stoodley et al, Effects of biofilm structure on oxygen distribution and mass transport, Biotechnol Bioeng, 1994: 43: 1131-8
67. Cahi Wen L. et al. Susceptibility of Enterococcus faecalis biofilm to antibiotics and calcium hydroxide, Journal of Oral Science, Vol. 49, No. " 161-166, 2007.
68. Chowdary, M.V.P., Kalyan Kumar, K., Keerthi Thakur, Anand, A., Jacob Kurien, Murali Krishna, c., and Stanley Mathew. 2007, Photomedicine and Laser Surgery 25(4), 269-274.
69. Stone, N., Hart Prieto, M.C., Crow, P., Uff, J., and Ritchie, A.W. 2007, Anal. Bioanal. Chem. 387, 1657-1668.
70. Grazielle, V. Nogueira, Landulfo Silveira, Jr., Airton A. Martín, Renato A. Zangaro, and Marcos T.T Pacheco, Maria C. Chavantes and Carlso A. Pasqualucci. 2005, Journal of biomedical optics, 10(3), 031117.
71. Tzu- Feng Hsieh, Ken-Jen Yu., and Shan- Yang Lin. 2007, Disease Markers, 23, 147-152.
72. Christopher V. Eddy and Mark A. Near Infrared Spectroscopy for measuring urea in hemodialysis fluid, Clinical Chemistry (2001)
73. M. Tarumi, M. Shimada, T. Murakami, Tamura, M. Shimada, H Arimoto and Y.. Yamada, *Simulation study in vitro glucose measurement by NIR spectroscopy and a method of error reduction*, Phys. Med. Biol. 48, 2373-2390, 2003.
74. Esbensen K H (2001) Multivariate Data Analysis in practice 5th Ed. CAMO Process AS, Oslo.

75. P.C. Painter and L. Koenig, *Biopolymers* 14 (1975)457-468.
76. Lambert J L, Pelletier L (1975) *Biopolymers* 14: 457- 468.
77. Tsuboi M, *in vibrational spectroscopy-Modern trends*, A.J. Barnes and W.J. Orville-Thomas(eds.), Elseiver, New York, 1977.
78. Norman B Couthup, *Introduction to Infrared and Raman Spectroscopy*, 6a Edit. 1990.
79. Kurt W. Short, Susan Carpenter. *Biophysical Journal*, June 2005.
80. M. Pézolet, M. Pigeon-Gosselin, L. Coulombe, *Biochim Biophys Acta* 453(1976) 502.
81. Apostol N. Batsoulis, Nikolaos G. FT- Raman Spectroscopic Simultaneous Determination of Fructose and Glucose in Honey, *Journal of Agricultural and Food Chemistry*, 2005, 53, 207-210.
82. Luiz Fernando C. de Oliveira, Roxana Colombana, Fourier Transform Raman spectroscopy of Honey, *Applied Spectroscopy*, volume 56 Number 3, 2002.

ANEXOS

Artículo de investigación

Uso de la espectroscopías Raman y quimiometría aplicada al estudio de muestras de suero con y sin anticuerpos específicos al Toxoplasma Gondii.

R. Ávila-Rodríguez, C. Araujo-Andrade, C. Frausto-Reyes, J. L. Pichardo-Molina, G. Barbosa-Sabanero, J. R. Martínez, Facundo Ruiz and R. Ivanov Tzonchev.

Aceptado para su publicación en la revista *Trends in applied Spectroscopy* en Agosto de 2008.

ANEXO



Original Communication

Raman spectroscopy and chemometrics applied to the study of serum samples with and without antibodies against *Toxoplasma gondii*

R. Ávila-Rodríguez¹, C. Araujo-Andrade², C. Frausto-Reyes³, J. L. Pichardo-Molina³, G. Barbosa-Sabanero⁴, J. R. Martínez^{5,*}, Facundo Ruiz⁵ and R. Ivanov Tzonchev²

¹Programa de Posgrado en Ingeniería Eléctrica, Facultad de Ingeniería, Universidad Autónoma de San Luis Potosí. Av. Salvador Nava S/N, Zona, ²Unidad Académica de Física de la Universidad Autónoma de Zacatecas. Zacatecas, Zac, ³Centro de Investigaciones en Óptica, A. C, Loma del Bosque No.115, Col. Lomas del Campestre, C. P. 37150, León Gto, ⁴Instituto de Investigaciones Médicas, Facultad de Medicina de la Universidad de Guanajuato. 20 de Enero 929, Col. Obregón. C.P. 37320. León, Gto, ⁵Facultad de Ciencias, Universidad Autónoma de San Luis Potosí, 78000 San Luis Potosí, S.L.P., México

ABSTRACT

Clinical diagnosis of Infections generally is realized by means of serological methods, which detect and identify the specific antibodies in tissue fluids associated to a pathogen like bacteria, virus or exotoxins. Antibodies are proteins present in the tissue fluids of our bodies that aid to the elimination or neutralization of pathogens or antigens. On the other hand, the serological methods require chemical reagents, special simple preparation and are time consuming to detect or identify the antibodies. In this work, samples of human serum have been studied by Raman microscopy. Spectra were obtained at excitation wavelength of 514 nm. The Raman spectra of serum contain multiples bands due to serum components such proteins, amino acids, carbohydrates and carotenoids. The results obtained by Multivariate classification, i.e. principal component analysis (PCA) and Fisher's linear discriminant, suggest that is feasible discriminate between serum samples with or without antibodies against *Toxoplasma gondii*.

KEYWORDS: antibodies, human serum, principal component analysis, Raman microscopy, *Toxoplasma gondii*

