

Escuela Superior Politécnica del Litoral

Facultad de Ciencias Naturales y Matemáticas

Diseño de una estrategia espectroscópica para el estudio de mecanismos de reacción en sistemas biológicos.

INGE 2700

Proyecto Integrador

Previo la obtención del Título de:

Ingeniería Química

Presentado por:

Cindy Mishell Campos Ramírez

Karen Deyaneira Parra Sánchez

Guayaquil - Ecuador

Año: 2024

Dedicatoria

A mi esposo y a mi hija, el centro de mi universo. A ustedes, que hacen que cada día valga la pena, les dedico con todo mi corazón este logro. Y a mi madre, por ser parte esencial de mi historia.

Karen Parra Sánchez

A mi madre, la persona que merece y representa cada uno de mis esfuerzos y logros.

¡Lo hicimos, mami!

Cindy Campos R.

Agradecimientos

A mi esposo y a mi hija, el corazón de mi vida. Su amor es mi mayor inspiración. Todo lo que hago es por ustedes.

A mi familia, por apoyarme y estar siempre a mi lado.

A mis amigos, por cada risa compartida y por siempre tener una palabra de apoyo cuando más lo necesitaba.

Al PhD. Joan Vera y al Ing. Julio De la Paz, por su guía y su confianza en mí a lo largo de este viaje.

Karen Parra Sánchez

A mis padres y familiares, por su apoyo incondicional. A los profesores que, con su enseñanza y motivación, marcaron cada etapa de mi vida académica, en especial al PhD. Joan Vera, por su confianza y guía constante. A mis amigos, quienes fueron un pilar esencial tanto emocional como académico, ofreciéndome su apoyo en los momentos más desafiantes. Y, finalmente, mis agradecimientos a mi niña interior, que siempre lo deseó y se aferró hasta conseguirlo.

Cindy Campos R.

Declaración Expresa

Nosotros Campos Ramírez Cindy Mishell y Parra Sánchez Karen Deyaneira acordamos y reconocemos que:

La titularidad de los derechos patrimoniales de autor (derechos de autor) del proyecto de graduación corresponderá al autor o autores, sin perjuicio de lo cual la ESPOL recibe en este acto una licencia gratuita de plazo indefinido para el uso no comercial y comercial de la obra con facultad de sublicenciar, incluyendo la autorización para su divulgación, así como para la creación y uso de obras derivadas. En el caso de usos comerciales se respetará el porcentaje de participación en beneficios que corresponda a favor del autor o autores.

La titularidad total y exclusiva sobre los derechos patrimoniales de patente de invención, modelo de utilidad, diseño industrial, secreto industrial, software o información no divulgada que corresponda o pueda corresponder respecto de cualquier investigación, desarrollo tecnológico o invención realizada por mí/nosotros durante el desarrollo del proyecto de graduación, pertenecerán de forma total, exclusiva e indivisible a la ESPOL, sin perjuicio del porcentaje que me/nos corresponda de los beneficios económicos que la ESPOL reciba por la explotación de mi/nuestra innovación, de ser el caso.

En los casos donde la Oficina de Transferencia de Resultados de Investigación (OTRI) de la ESPOL comunique los autores que existe una innovación potencialmente patentable sobre los resultados del proyecto de graduación, no se realizará publicación o divulgación alguna, sin la autorización expresa y previa de la ESPOL.

Guayaquil, 28 de octubre del 2024.



Karen Deyaneira Parra Sánchez



Cindy Mishell Campos Ramirez

Evaluadores

MSc. Nadia Flores

Profesor de Materia

PhD. Joan Vera

Tutor de proyecto

Resumen

La glicación no enzimática de proteínas es un proceso clave en el desarrollo de enfermedades crónicas. Este proyecto busca optimizar la interpretación de datos espectroscópicos obtenidos mediante FTIR y técnicas complementarias como 2DCOS y PCA, utilizando proteínas modelo (BSA y hemoglobina bovina) en condiciones de glucosilación y congestiónamiento molecular. La hipótesis plantea que la combinación de estas herramientas permite identificar cambios conformacionales clave y resalta el papel protector de la metformina frente a la glicación.

El proyecto se desarrolló en el laboratorio universitario, utilizando materiales y equipos disponibles como el espectrómetro FTIR y reactivos de alta pureza. Se diseñó un aplicativo web que integra PCA, facilitando la visualización y análisis dinámico de los datos. Este procedimiento permitió analizar sistemas con distintas perturbaciones, evaluando las interacciones proteína-ligando en medios con y sin PEG-8 como agente de congestiónamiento.

Los resultados mostraron que la metformina protege la estructura proteica al evitar modificaciones inducidas por la glicación, incluso en condiciones de congestiónamiento molecular. Se concluye que esta estrategia optimiza el análisis espectral, acelerando la identificación de perturbaciones relevantes en sistemas biológicos complejos.

Palabras clave: Espectroscopía FTIR, glicación, PCA, metformina.

Abstract

Non-enzymatic glycation of proteins is a key process in the development of chronic diseases. This project seeks to optimize the interpretation of spectroscopic data obtained by FTIR and complementary techniques such as 2DCOS and PCA, using model proteins (BSA and bovine hemoglobin) under conditions of glycosylation and molecular congestion. The hypothesis is that the combination of these tools allows the identification of key conformational changes and highlights the protective role of metformin against glycation.

The project was developed in the university laboratory, using available materials and equipment such as FTIR spectrometer and high purity reagents. A web application was designed that integrates PCA, facilitating the visualization and dynamic analysis of the data. This procedure allowed the analysis of systems with different perturbations, evaluating protein-ligand interactions in media with and without PEG-8 as a congeneric agent.

The results showed that metformin protects the protein structure by preventing glycation-induced modifications, even under conditions of molecular congestion. It is concluded that this strategy optimizes spectral analysis, accelerating the identification of relevant perturbations in complex biological systems.

Keywords: FTIR spectroscopy, glycation, PCA, metformin.

Índice general

Resumen	I
<i>Abstract</i>	II
Índice general	III
Abreviaturas.....	V
Simbología.....	VI
Índice de figuras	VII
Capítulo 1	1
1.1 Introducción	2
1.2 Descripción del problema.....	2
1.3 Justificación del problema.....	3
1.4 Objetivos.....	6
1.4.1 <i>Objetivo general</i>	6
1.4.2 <i>Objetivos específicos</i>	6
1.5 Marco teórico	6
1.5.1 <i>Glicación no enzimática de proteínas (Reacción de Maillard)</i>	6
1.5.2 <i>Rol de la metformina en la inhibición de AGEs</i>	7
1.5.3 <i>Efectos de la congestión molecular en sistemas biológicos</i>	7
1.5.4 <i>Fundamentos de la espectroscopía FTIR en el análisis de proteínas</i>	8
1.5.5 <i>Análisis espectroscópico bidimensional (2DCOS)</i>	8
1.5.6 <i>Métodos estadísticos para el preprocesamiento de datos espectroscópicos</i>	9
Capítulo 2	12
2. METODOLOGÍA	13
2.1 Selección de método estadístico.....	13
2.2 Sistemas Biomoleculares perturbados	14
2.2.1 Materiales	15
2.2.2 Mediciones	17
2.2.3 Procesamiento de datos	18
2.3 Implementación de análisis PCA-2DCos.....	19

2.4	Desarrollo de la interfaz para análisis PCA	21
2.5	Costos asociados	23
Capítulo 3		24
3. Resultados y análisis		25
3.1 Análisis del Diseño Experimental		25
3.2 Estudio comparativo de la BHb con y sin PEG-8		25
3.3 Influencia del uso de PEG-8 en sistemas perturbados.....		50
3.5 Análisis de costos		64
Capítulo 4		65
4. Conclusiones y recomendaciones.....		66
4.1 Conclusiones		66
4.2 Recomendaciones		67
Referencias.....		69
Anexos.....		75

Abreviaturas

2DCOS	Espectroscopía Bidimensional de Correlación.
AGEs	Productos Finales de Glicación Avanzada (Advanced Glycation End-products).
ATR	Reflectancia Total Atenuada (Attenuated Total Reflectance).
BHb	Hemoglobina Bovina.
BSA	Albúmina Sérica Bovina.
FTIR	Espectroscopía Infrarroja por Transformada de Fourier.
HSA	Albúmina Sérica Humana.
KH ₂ PO ₄	Fosfato Monopotásico de Potasio.
NaCl	Cloruro de Sodio.
Na ₂ HPO ₄	Fosfato Dibásico de Sodio.
NaOH	Hidróxido de Sodio.
PCA	Análisis de Componentes Principales (Principal Component Analysis).
PEG	Polietilenglicol.
RSM	Metodología de Superficies de Respuesta (Response Surface Methodology).
UV-Vis	Espectroscopia Ultravioleta-Visible.

Simbología

°C	Grados Celsius
atm	Atmósfera (unidad de presión)
C	Carbono
cm ⁻¹	Centímetro inverso (unidad de frecuencia en espectroscopia)
H	Hidrógeno
HR	Humedad Relativa
mg/ml	Miligramos por mililitro
ml	Mililitros
μl	Microlitros
N	Nitrógeno
O	Oxígeno
pH	Potencial de Hidrógeno

Índice de figuras

Figura 1.1 <i>Reacción entre la glucosa y un grupo amino de una proteína para formar un producto de Amadori.(Zhang et al., 2009)</i>	7
Figura 2.1 <i>Representación de matriz de diseño experimental</i>	15
Figura 2.2 <i>Diagrama de flujo del procedimiento del uso del equipo FTIR.</i>	18
Figura 2.3 <i>Formato de datos preprocesados</i>	19
Figura 2.4 <i>Código en Python donde se realiza el PCA con sklearn leyendo datos de una hoja en Excel.</i>	20
Figura 2.5 <i>Ventana del software Py2DCos.</i>	21
Figura 2.6 <i>Mockup inicial de aplicación web desarrollada</i>	22
Figura 3.1 <i>Estrategia desarrollada para el análisis de sistemas complejos perturbados.</i>	26
Figura 3.2 <i>Sistema unitario de BHb a 10 mg/ml</i>	29
Figura 3.3 <i>Sistema unitario de BHb a 10 mg/ml con PEG-8</i>	30
Figura 3.4 <i>PCA del sistema unitario de BHb a 10 mg/ml</i>	31
Figura 3.5 <i>PCA del sistema unitario de Bhb a 10mg/ml con PEG-8</i>	32
Figura 3.6 <i>Sistema binario de BHb a 10 mg/ml y Metformina 40 mg/ml.</i>	35
Figura 3.7 <i>PCA del sistema binario de BHb a 10 mg/ml y Metformina 40 mg/ml</i>	36
Figura 3.8 <i>Sistema Binario BHb 10 mg/ml con Glucosa 15mg/ml.</i>	37
Figura 3.9 <i>PCA del sistema binario de BHb a 10 mg/ml y Glucosa 15 mg/ml.</i>	38
Figura 3.10 <i>Sistema Ternario BHb 10 mg/ml con Glucosa 15 mg/ml y Metformina 40 mg/ml</i>	39
Figura 3.11 <i>PCA del sistema ternario de BHb a 10 mg/ml, Glucosa 15 mg/ml y Metformina 40 mg/ml</i>	40
Figura 3.12 <i>2DCOS BHb 10 mg/ml por días.</i>	41
Figura 3.13 <i>2DCOS BHb 10 mg/ml y Glucosa 15 mg/ml por días.</i>	41
Figura 3.14 <i>2DCOS BHb 10 mg/ml y Metformina 40 mg/ml por días.</i>	41
Figura 3.15 <i>2DCOS BHb 10 mg/ml, Glucosa 15 mg/ml y Metformina 40 mg/ml por días</i> ...	42
Figura 3.16 <i>Sistema binario BHb 10 mg/ml con Metformina 40 mg/ml con PEG-8</i>	43
Figura 3.17 <i>PCA del sistema binario de BHb a 10 mg/ml y Metformina 15 mg/ml con PEG-8</i>	44
Figura 3.18 <i>Sistema Binario BHb 10 mg/ml con Glucosa 15mg/ml con PEG-8.</i>	45
Figura 3.19 <i>PCA del sistema binario de BHb a 10 mg/ml y Glucosa 15 mg/ml con PEG-8.</i> 46	
Figura 3.20 <i>Sistema Ternario BHb 10 mg/ml con Glucosa 15 mg/ml y Metformina 40 mg/ml con PEG-8</i>	47
Figura 3.22 <i>2DCOS PEG BHb 10 mg/ml por días</i>	49

Figura 3.23 2DCOS PEG BHb 10 mg/ml y Glucosa 15 mg/ml por días	49
Figura 3.24 2DCOS PEG BHb 10 mg/ml y Metformina 40 mg/ml por días	49
Figura 3.25 2DCOS PEG BHb 10 mg/ml, Glucosa 15 mg/ml y Metformina 40 mg/ml por días	50
Figura 3.26 Sistema unitario BSA 10 mg/ml	55
Figura 3.27 Sistema binario BSA 10 mg/ml y Metformina 40 mg/ml	56
Figura 3.28 Sistema binario BSA 10 mg/ml y Glucosa 15 mg/ml	57
Figura 3.29 Sistema ternario BSA 10 mg/ml con Glucosa 15 mg/ml y Metformina 40 mg/ml	58
Figura 3.30 PCA de sistema unitario de BSA 10 mg/ml con PEG 8.....	59
Figura 3.31 PCA de sistema binario con BSA 10 mg/ml y glucosa 15 mg/ml con PEG 8 ...	60
Figura 3.32 PCA de sistema binario de BSA 10 mg/ml y metformina 40 mg/ml con PEG 8	61
Figura 3.33 PCA de sistema ternario BSA 10 mg/ml, glucosa 15 mg/ml y metformina 40 mg/ml con PEG 8	62
Figura 3.34 2DCOS PEG BSA 10 mg/ml por días	63
Figura 3.35 2DCOS PEG BSA 10 mg/ml y glucosa 15 mg/ml por días.....	63
Figura 3.36 2DCOS PEG BSA 10 mg/ml y metformina 40 mg/ml por días	63
Figura 3.37 2DCOS PEG BSA 10 mg/ml, glucosa 15 mg/ml y metformina 40 mg/ml por días	64
Figura A. Validación con Statgraphics XIX para el PCA. a) Tablas de pesos obtenida con Python. b) Tabla de pesos obtenidos con Statgraphics XIX.....	74

Índice de tablas

Tabla 2.1 Método de Pugh para selección de método estadístico	14
Tabla 2.2 Diseño de experimentos	15
Tabla 2.3 Preparación de Tampón fosfato salino para sistemas con BSA y hemoglobina. .	16
Tabla 2.4 Concentraciones preparadas para diferentes sistemas.....	17
Tabla 3.1 Resultados de análisis PCA para el sistema unitario, binario y ternario de BHb .	27
Tabla 3.2 Resultados de análisis PCA de sistema unitario, binario y ternario de BHb con PEG	27
Tabla 3.3 Frecuencias de onda de estructuras para hemoglobina en estado fundamental .	28
Tabla 3.4 Resultados de las frecuencias con mayor carga en PCA para sistemas binario y ternario en BHb.....	34
Tabla 3.5 Resultados de análisis PCA de sistema unitario, binario y ternario de BSA con PEG	51
Tabla 3.6 Resultados de las frecuencias con mayor carga en PCA para sistemas binario y ternario en BSA.....	52
Tabla 3.7 Comparación con otra herramienta de análisis PCA	53

Capítulo 1

1.1 Introducción

La glicación no enzimática de proteínas, conocida como la reacción de Maillard, es un proceso químico clave en el desarrollo de enfermedades crónicas, como las cardiovasculares y neurodegenerativas. Este fenómeno, impulsado por la interacción entre aminoácidos y azúcares reductores, altera las estructuras proteicas y genera productos avanzados de glicación (AGEs).

Para estudiar este proceso, la espectroscopía FTIR es una herramienta valiosa, aunque a menudo requiere de ensayos complementarios para detectar cambios conformacionales complejos. Métodos complementarios como el análisis bidimensional de correlación espectroscópica (2DCOS) ofrecen mayor resolución espectral, pero su interpretación sigue siendo limitada sin el apoyo de estrategias estadísticas.

Esta investigación propone integrar FTIR, 2DCOS y el uso de técnicas estadísticas para estudiar los efectos de la glucosa y la metformina sobre proteínas modelo como la albúmina sérica bovina (BSA) y la hemoglobina bovina (BHb), simulando condiciones biológicas realistas mediante el uso de agentes de congestión molecular. Se busca agilizar la interpretación de los datos espectrales y revelar detalles clave sobre los cambios estructurales en las proteínas bajo perturbaciones específicas.

1.2 Descripción del problema

La espectroscopia infrarroja por transformada de Fourier (FTIR) es una herramienta útil para observar la glicación no enzimática y se ha utilizado previamente para el estudio de las interacciones entre la BSA, la glucosa y la metformina (De la Paz, 2023). Sin embargo, debido a la complejidad de los sistemas proteicos la técnica clásica de FTIR resulta insuficiente para detectar las interacciones entre ligandos y proteínas, tales como los cambios conformacionales, que pueden afectar la dinámica de la proteína (De La Arada et al., 2011). Diversas investigaciones han empleado diferentes aproximaciones para el tratamiento de la señal FTIR y su interpretación en el análisis de sistemas complejos, espectroscopía FTIR de primera derivada, análisis convolucional, entre otros, aún sin

alcanzar el nivel de detalles que permita desentrañar los complejos mecanismos de modificación conformacional que ocurren en la proteína debido a la dinámica de la interacción ((Yue et al., 2021), (Dong et al., 2022), (Andraka et al., 2012)). Por esta razón, se ha implementado el análisis de correlación bidimensional espectroscópica conocido como 2DCos, que permite observar los cambios provocados por perturbaciones en el sistema. El 2DCos permite correlacionar el acoplamiento de las bandas espectrales ante fluctuaciones en las variables de perturbación (De La Arada et al., 2011) usando los datos obtenidos por FTIR.

El FTIR-2D-COS es una herramienta poderosa para revelar los mecanismos de interacción entre ligandos y proteínas, aunque requiere un dominio profundo de la técnica para obtener el máximo beneficio de los resultados. Por ello, es esencial desarrollar estrategias de tratamiento de datos que permitan al usuario identificar los criterios clave de análisis. En sistemas biológicos, la alta complejidad estructural demanda mayor "resolución" en el análisis. Así, se propone incorporar técnicas estadísticas como PCA, PLS y análisis de superficies de respuesta para eliminar datos irrelevantes, mejorar la calidad espectral y resaltar las variaciones más significativas (Jung, 2003). Estas técnicas servirían para desarrollar un enfoque complementario al FTIR-2D-COS que facilite la interpretación de los datos.

1.3 Justificación del problema

Las interacciones proteína-ligando son fundamentales en procesos biológicos clave como la catálisis enzimática, señalización celular y transporte de moléculas en organismos vivos. Estos procesos se modelan en experimentos *in vitro* para comprender los mecanismos funcionales, estudiar enfermedades y desarrollar fármacos. Estas interacciones dependen de la unión específica y reversible de un ligando a un sitio de la proteína, regulada por la complementariedad estructural y fuerzas intermoleculares (Gholami & Shirmohammadi, 2024).

Las proteínas de transporte, que se unen a ligandos específicos para facilitar su movimiento en el organismo, son esenciales en la investigación de funciones biológicas y el

desarrollo de fármacos. La espectroscopia FTIR permite estudiar estos procesos al revelar detalles sobre los sitios de unión de los ligandos, cambios conformacionales en la proteína y la intensidad de las interacciones. Por ejemplo, (Alhazmi, 2019) estudió la albúmina sérica bovina (BSA), donde la unión de iones metálicos reduce hélices alfa y aumenta láminas beta, observable en la banda Amida I del espectro FTIR ($1700-1600\text{ cm}^{-1}$). Este método, preciso al compararse con electroforesis capilar de afinidad, demuestra el valor de FTIR en el análisis de interacciones proteína-ligando, aunque su interpretación requiere experiencia y tiene limitaciones en resolución espectral.

En este contexto, el 2DCOS aporta valor adicional al resolver picos espectrales superpuestos y facilitar el análisis de cambios en la intensidad de las bandas, revelando si estos ocurren simultáneamente o en secuencia ante una perturbación externa (Foist et al., 2011). La técnica 2DCOS ha sido eficaz en la caracterización de sistemas biológicos complejos y su respuesta a perturbaciones. Por ejemplo, se han estudiado cambios metabólicos en células normales y cancerosas sometidas a radiación mediante micro espectroscopía infrarroja y 2DCOS, mapeando variaciones según la dosis y tiempo post-irradiación (Andraka et al., 2012). Otro estudio usó 2DCOS y FTIR para investigar la distribución de colágeno y proteoglicanos en cartílago humano, distinguiendo tejidos sanos de enfermos (Jiang & Rieppo, 2006). Además, 2DCOS ha sido útil en la desnaturalización de proteínas en condiciones extremas, como variaciones de temperatura y presión, permitiendo identificar cambios estructurales y formación de agregados, además de evaluar la actividad enzimática de tripsina en polvo bajo compresión, correlacionando disminución de actividad con el aumento de presión (Jung et al., 2015).

Estos estudios destacan la capacidad de 2DCOS para revelar detalles estructurales y dinámicos en sistemas complejos, justificando su uso en el análisis de interacciones biomoleculares. Sin embargo, el tamaño de los datos y las características adicionales que se generan debido a las perturbaciones suelen hacer que los resultados sean difíciles de interpretar sin un análisis más detallado, específico y orientado a cambios biomoleculares.

Por ello, se busca desarrollar una estrategia basada en métodos estadísticos, similar a la del estudio de Jung, (2003), en el que se emplearon datos espectrales de una solución de poliestireno/metiletilcetona/tolueno. En ese estudio, se utilizó el análisis de componentes principales (PCA) para crear una matriz de datos sin interferencia de componentes insignificantes, mejorando así la interpretación de sistemas complejos con el 2DCOS, al eliminar datos irrelevantes y redundantes. Otro estudio de Larios et al., (2021) analizó datos de suero sanguíneo canino para detectar diferencias entre muestras infectadas de Leishmaniasis y muestras no infectadas, donde el uso de PCA reveló bandas espectrales sensibles a la infección, proporcionando información útil para métodos diagnósticos.

Para esta investigación, se analizarán sistemas compuestos por una proteína, glucosa y metformina, utilizando polietilenglicol 8000 (PEG) como agente de congestión. Las proteínas de transporte a emplear serán la albúmina sérica bovina y la hemoglobina bovina, que según Khan et al., (2021) y Outman et al., (2023) son estructuralmente similares a la albúmina y hemoglobina humanas, respectivamente. La BSA es una proteína estructuralmente similar a la albúmina sérica humana (HSA), la cual se encuentra en abundancia en el plasma sanguíneo y es susceptible a la glicación no enzimática en condiciones de hiperglucemia (Khan et al., 2021). La hipótesis plantea que el acoplamiento de técnicas de análisis de datos con datos espectroscópicos experimentales permitirá definir criterios relevantes para el análisis FTIR-2DCOS. Se espera que esta estrategia facilite la interpretación y acelere el análisis de datos espectroscópicos, revelando detalles intrínsecos de los cambios conformacionales en estos sistemas.

La justificación de este proyecto establece las bases para su desarrollo, destacando la relevancia del estudio y su impacto potencial en el área científica. A partir de ello, se plantearon objetivos claros que guían el enfoque de la investigación y las estrategias empleadas para alcanzar los resultados esperados. Asimismo, se exploran los fundamentos teóricos que sustentan el proyecto, proporcionando el contexto necesario para comprender su desarrollo y aplicación.

1.4 Objetivos

1.4.1 Objetivo general

Desarrollar una estrategia basada en las técnicas de análisis de datos para el estudio de datos espectroscópicos de sistemas complejos utilizando la herramienta Py2DCos-FTIR.

1.4.2 Objetivos específicos

1.- Evaluar los sistemas proteína-ligando en condiciones de glucosilación no enzimática y congestiónamiento molecular, en presencia y ausencia de metformina para las proteínas BSA y Hemoglobina Bovina bajo perturbación utilizando FTIR.

2.- Caracterizar los sistemas con y sin congestiónamiento molecular utilizando Py2DCos y estrategias estadísticas de análisis de datos para que se establezcan criterios de interpretación de señales.

3.- Diseñar un aplicativo para el acoplamiento de análisis estadístico con FTIR-Py2DCos.

1.5 Marco teórico

1.5.1 Glicación no enzimática de proteínas (Reacción de Maillard)

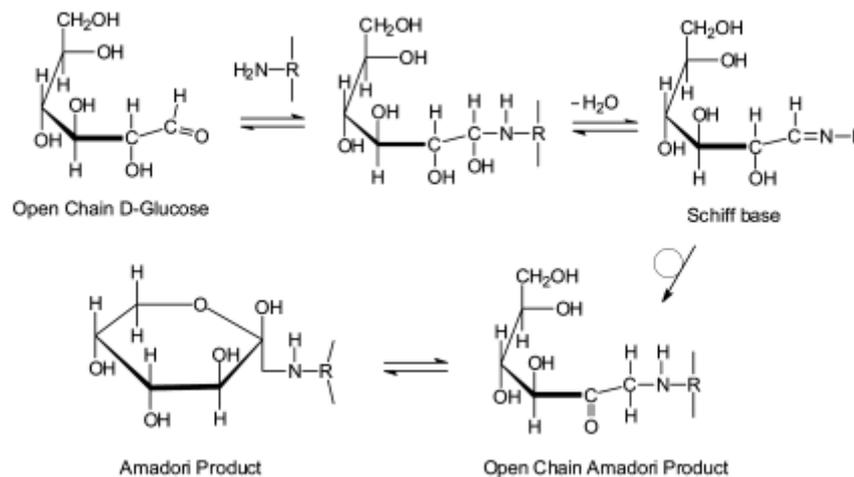
La reacción de Maillard, conocida también como glicación no enzimática de proteínas, fue descrita en 1912 por el científico Louis Camille Maillard, quien caracterizó este proceso como una reacción química entre un grupo amino libre de proteínas y un azúcar reductor. Esta reacción ocurre en varias etapas: en primer lugar, un azúcar reductor como la glucosa reacciona con un grupo amino libre de la proteína, comúnmente en residuos de lisina, formando un intermedio inestable conocido como base de Schiff (Younus & Anwar, 2016). La base de Schiff se reorganiza a través de un proceso espontáneo, generando el producto de Amadori o cetoamina, una molécula más estable que representa un intermedio importante en la reacción.

Este compuesto de Amadori puede luego experimentar una serie de modificaciones químicas, tales como oxidación, deshidratación y ciclación, lo que finalmente da lugar a la formación de productos finales de glicación avanzada (AGEs). Bajo condiciones de

hiperglucemia en pacientes con diabetes mellitus, la formación de AGEs se acelera, lo cual altera la estructura y función de diversas proteínas críticas, como el colágeno, la albúmina sérica humana (HSA), la hemoglobina, entre otras (Khan et al., 2021) La acumulación de AGEs se ha relacionado con la progresión y complicación de diversas enfermedades crónicas, incluyendo la diabetes, la artritis, el Alzheimer, y nefropatías (Khan et al., 2021).

Figura 1.1

Reacción entre la glucosa y un grupo amino de una proteína para formar un producto de Amadori (Q. Zhang et al., 2009).



1.5.2 Rol de la metformina en la inhibición de AGEs

La metformina es un agente anti hiperglucemiante oral utilizado para el manejo de la diabetes mellitus (Khan et al., 2021). Dado su semejanza estructural con las guanidinas, la metformina es objeto de estudios especiales, ya que esta característica le otorga la capacidad de inhibir la glicación al reaccionar con los grupos carbonilo de los azúcares reductores y de los compuestos dicarbonílicos compitiendo con las proteínas por estos grupos carbonilos (Rahbar & Figarola, 2003).

1.5.3 Efectos de la congestión molecular en sistemas biológicos

Según (Caccamo & Magazú, 2022), el PEG estabiliza proteínas en condiciones similares a la congestión celular, reduciendo variaciones espectrales debidas a aumentos de temperatura y mejorando la "resiliencia térmica" del sistema, que mide la capacidad de la proteína para retornar a su estado inicial tras un cambio térmico. En el contexto de la glicación de proteínas, estudios con albúmina sérica bovina (BSA) muestran que PEG 2000

y PEG 4000 afectan la dinámica y estructura de la BSA, promoviendo levemente la formación de productos de glicación. No obstante, el PEG reduce la agregación de BSA y genera cambios estructurales mínimos, limitados a ligeras alteraciones locales. El análisis de fluorescencia tridimensional también confirma que la congestión molecular afecta las interacciones proteína-ligando de forma sutil en el proceso de glicación (Y. Zhang et al., 2023).

1.5.4 Fundamentos de la espectroscopía FTIR en el análisis de proteínas

La espectroscopia infrarroja por transformada de Fourier (FTIR) es una técnica analítica utilizada en química, biología y medicina, que permite identificar y cuantificar enlaces químicos en muestras a través de la interacción de la radiación infrarroja con la materia. Los enlaces químicos vibran a frecuencias específicas cuando son expuestos a esta radiación, presentando características de estiramiento o flexión (Venkatesan et al., 2011).

FTIR tiene diversas aplicaciones, incluyendo la identificación de compuestos mediante espectros que actúan como "huellas dactilares" moleculares. También permite análisis cuantitativos, ya que la intensidad de las bandas de absorción es proporcional a la concentración del analito. Además, es útil para estudiar interacciones moleculares y se aplica en el análisis de materiales biológicos como proteínas, lípidos y carbohidratos (Andraka et al., 2012).

1.5.5 Análisis espectroscópico bidimensional (2DCOS)

La espectroscopia de correlación bidimensional (2DCOS) es una técnica quimiométrica utilizada para analizar datos espectrales, como los obtenidos por FTIR, y visualizar relaciones entre bandas espectrales, revelando información no evidente en espectros unidimensionales (Noda, 2016). Existen espectros 2DCOS síncronos, que muestran correlaciones entre bandas que cambian simultáneamente con una perturbación, y asíncronos, que revelan correlaciones entre bandas que cambian en diferentes momentos (Yue et al., 2021).

La 2DCOS se aplica en el análisis de mecanismos de reacción, ayudando a identificar el orden de los cambios espectrales; en el estudio de transiciones de fase, como la fusión y cristalización de materiales; y en la caracterización de materiales complejos, aportando información sobre la composición y estructura en sistemas como polímeros y biomateriales ((Jung et al., 2015); (Jiang & Rieppo, 2006)).

1.5.6 Métodos estadísticos para el preprocesamiento de datos espectroscópicos

1.4.6.1 Análisis de Componentes Principales (PCA). El PCA es una técnica estadística que se utiliza para reducir la dimensionalidad de los datos y extraer la información más relevante. El PCA transforma un conjunto de datos original en un nuevo conjunto de variables, llamadas componentes principales, que son combinaciones lineales de las variables originales (Venkatesan et al., 2011).

El PCA se utiliza en una variedad de aplicaciones, que incluyen: reducción de la dimensionalidad de los datos ya que se utiliza para simplificar grandes conjuntos de datos identificando las variables más importantes, identificación de patrones en los datos, como grupos o tendencias y visualización de datos multidimensionales en un espacio de menor dimensión, lo que facilita la interpretación (Larios et al., 2021).

1.4.6.2 Regresión de mínimos cuadrados parciales (PLS). El PLS es un método estadístico comúnmente utilizado en el análisis de datos FTIR, especialmente cuando hay muchas variables explicativas en comparación con el número de muestras. A diferencia de la regresión lineal múltiple (MLR), el PLS es más adecuado para datos con variables correlacionadas, ya que encuentra componentes principales (PCs) que capturan la mayor variación en ambas variables, utilizando estas PCs como predictores en el modelo de regresión (Venkatesan et al., 2011). Este método tiene diversas aplicaciones, como la estimación de la concentración de surfactantes en muestras de proteínas a partir de espectros FTIR (Vijayakumar et al., 2024) y la identificación de pacientes con diabetes tipo 2 mediante el análisis de espectros FTIR de saliva (Sánchez-Brito et al., 2021). Entre sus ventajas, el PLS puede manejar conjuntos de datos con variables correlacionadas, identificar las

variables más relevantes y es robusto frente al ruido en los datos. Un estudio demostró que el uso de PLS con espectros FTIR de saliva permitió predecir con precisión los niveles de glucosa en sangre, lo que sugiere que la combinación de FTIR y PLS es útil para el control de la diabetes (Sánchez-Brito et al., 2021).

1.4.6.3 Metodología de superficie de respuesta (RSM). La metodología de superficie de respuesta (RSM) modela la relación entre una o más variables de respuesta y múltiples variables independientes, optimizando resultados al identificar combinaciones óptimas de estas variables. Usada en industrias como la alimentaria y la química, permite maximizar o minimizar resultados y optimizar procesos (Olabinjo et al., 2024). Un estudio aplicó RSM para corregir el efecto del oxígeno en la cuantificación de dióxido de carbono por FTIR, logrando mediciones más precisas de CO₂. Las ventajas incluyen visualización de relaciones complejas, optimización de factores y mejora en la precisión (Seichter et al., 2017).

1.4.7 Diseño de experimentos para el análisis de sistemas complejos. El Diseño de Experimentos (DOE) es una metodología estadística que planifica, ejecuta y analiza experimentos de manera sistemática para identificar y cuantificar relaciones de causa y efecto entre variables en un proceso, aplicándose en la optimización de productos y procesos en distintos campos científicos (Jankovic et al., 2021).

Dependiendo de la investigación, se seleccionan variables significativas para maximizar la información obtenida con los mínimos recursos (Jankovic et al., 2021). En sistemas biológicos, la complejidad aumenta al incluir interacciones y dinámicas proteicas (Wang & Annunziata, 2007). Jankovic et al., (2021) describen etapas clave en el diseño experimental: primero, se selecciona el sistema de estudio y se define el problema de investigación, lo que orienta la elección de variables y factores a analizar. Luego, se identifican las variables representativas y factores controlables. En la fase de diseño, la metodología de superficies de respuesta es especialmente útil en sistemas complejos, permitiendo optimizar la variable de respuesta; el diseño

compuesto central ayuda a determinar si el modelo debe ser lineal o de orden superior, según la curvatura observada. Finalmente, el análisis de datos, comúnmente con ANOVA, identifica la significancia de los efectos de los factores y evalúa si las diferencias entre muestras son atribuibles a factores específicos.

Capítulo 2

2. METODOLOGÍA

2.1 Selección de método estadístico

El uso del Py2DCos representa gran importancia en el análisis de sistemas biomoleculares complejos estudiados en FTIR. Sin embargo, debido a la extensa magnitud de datos generados resulta imperativo generar selección de la data que mejor simbolice los cambios en las moléculas perturbadas. Por lo tanto, dicha selección debe basarse en un método estadístico que aporte validez a los resultados obtenidos por el analista.

Dentro de las alternativas de métodos estadísticos expuestas en el capítulo 1 se dedujo por medio de una tabla de decisión o método de Pugh usando criterios que se adaptan a los objetivos del presente estudio como manejo de gran cantidad de datos, resultados gráficos, selección de datos más representativos entre otros descritos en la tabla 2.1, dando como resultado que aunque el PLS es una poderosa técnica de modelado predictivo, su enfoque en la relación entre variables independientes y dependientes no es el apropiado para identificar los cambios moleculares inherentes de los sistemas tal como ha sido usado en investigaciones con fines predictivos como pronosticar propiedades químicas de sustancias (L. Zhang et al., 2005) o en la investigación de Brixner Riça et al., 2019 donde se obtuvo un modelo fiable de predicción de cáncer en células.

El RSM es útil para optimizar procesos experimentales, pero no es adecuado para analizar datos espectrales complejos, ya que no identifica patrones o regiones específicas de cambio. En contraste, el PCA genera gráficos de puntuaciones y ponderaciones que permiten destacar las regiones espectrales donde ocurren los cambios más significativos debido a perturbaciones, haciendo de esta técnica una herramienta ideal para el análisis detallado de datos espectrales según (Hniopek et al., 2020), previamente usado en data proveniente de FTIR resaltando información importante en la búsqueda de micro plástico (Wander L et al., 2019)

Tabla 2.1

Método de Pugh para selección de método estadístico

Método	PLS	PCA	RSM
Criterio			
Uso simultáneo en FTIR y 2DCos	✓	✓	✗
Gran cantidad de datos	✓	✓	✗
Identifica cambios significativos	✓	✓	✗
Consideración Porcentual	✓	✓	✗
Resultados Gráficos	✗	✓	✗

2.2 Sistemas Biomoleculares perturbados

Para obtener los datos necesarios para el análisis de componentes principales (PCA), se llevaron a cabo mediciones mediante espectroscopía FTIR. El propósito de estas mediciones fue analizar el mecanismo de reacción de los ligandos glucosa y metformina actuando en las proteínas de transporte, específicamente la albúmina sérica bovina (BSA) y la hemoglobina bovina, con la incorporación de un agente aglutinante para simular condiciones de congestión molecular. El agente seleccionado fue el PEG 8000, un poli éter sintético caracterizado por su naturaleza hidrófila y biocompatible, cualidades que lo hacen ampliamente utilizado en formulaciones farmacéuticas. Además, el PEG 8000 encuentra aplicaciones en diversas industrias debido a sus propiedades como aditivo, emulsionante, lubricante y humectante, ofreciendo versatilidad en múltiples contextos (Ibrahim et al., 2022).

Por otro lado, el diseño experimental detalla la consideración del uso de 2 proteínas, 2 medios (con y sin PEG para la hemoglobina), 3 sistemas (unitario, binario y ternario), 9 y 11 concentraciones para los ligandos y para las proteínas respectivamente con repeticiones para 5 días, como se muestra a detalle en la tabla 2.2 y en la figura 2.1, donde cada término de la matriz representa las diferentes concentraciones (C_n , $n=1, 2, \dots$) para cada día (D_n , $n=1, 2, \dots$). Las condiciones ambientales se mantuvieron a 1 atm de presión, 20 °C y HR entre 50 % a 60%. Estos experimentos se llevaron a cabo en el Laboratorio de Análisis Químico Instrumental (LAQUINS), garantizando un control adecuado de las variables experimentales y permitiendo la generación de datos consistentes para el análisis.

Tabla 2.2

Diseño de experimentos

Sistema	BSA	Hemoglobina	Total
Unitario	$(2 \times 8) + 10 + 1 = 27$	$11 \times 2 = 22$	
Binario	$(8 \times 2) + (10 \times 2) + 1 = 37$	$(8 \times 2) + (10 \times 2) + 1 = 37 \times 2 = 74$	
Ternario	$8 + 1 = 9$	$8 + 1 = 9 \times 2 = 18$	
Total	75	114	198

Figura 2.1

Representación de matriz de diseño experimental

Tiempo	Concentraciones											
	Unitario				Binario				Ternario			
	D1C1	D1C2	D1C3	...	D1C1	D1C2	D1C3	...	D1C1	D1C2	D1C3	...
	D2C1	D2C2	D2C3	...	D2C1	D2C2	D2C3	...	D2C1	D2C2	D2C3	...
	D3C1	D3C2	D3C3	...	D3C1	D3C2	D3C3	...	D3C1	D3C2	D3C3	...
	D4C1	D4C2	D4C3	...	D4C1	D4C2	D4C3	...	D4C1	D4C2	D4C3	...
	D5C1	D5C2	D5C3	...	D5C1	D5C2	D5C3	...	D5C1	D5C2	D5C3	...

2.2.1 Materiales

- Espectrómetro FT-IR Nicolet iS10 provisto de celda ATR de ZnSe multirrebote a 45°, rango de 650-4000 cm^{-1} con resolución de 4 cm^{-1} y 60 barridos

- Micropipetas de 100 μ l y de 1000 μ l
- Viales plásticos de 1.5 ml
- Balones aforados de 500 ml y de 100 ml
- Solución Buffer Fosfato Salino 20mM a pH neutro
- Glucosa
- Metformina aislada de la comercial
- Albúmina sérica bovina y Hemoglobina Bovina de alta pureza y de grado biología molecular
- Fosfato dibásico de sodio (Na_2HPO_4)
- Fosfato monopotásico (KH_2PO_4)
- Cloruro de sodio (NaCl)
- Polietilenglicol 8000 ($\text{H}(\text{OCH}_2\text{CH}_2)_n\text{OH}$)

Tabla 2.3

Preparación de Tampón fosfato salino para sistemas con BSA y hemoglobina.

		BSA		Hemoglobina	
		Sistema Unitario y Ternario	Sistema Binario	Unitario, Binario, Ternario	
				Sin PEG-8	Con PEG-8
Compuesto	Fórmula	Cantidad (g)	Cantidad (g)	Cantidad (g)	Cantidad (g)
Fosfato dibásico de sodio	Na_2HPO_4	1.2021	1.4683	1.2035	1.2058
Fosfato monopotásico	KH_2PO_4	0.5068	0.5213	0.5015	0.5070
Cloruro de sodio	NaCl	1.6028	1.6654	1.6005	1.6004
PEG-8	$\text{H}(\text{OCH}_2\text{CH}_2)_n\text{OH}$	25.0033	25.0008	-	25.0088
pH final		6.99	7.02	7.01	6.99

Para regular el pH, se usó hidróxido de sodio NaOH y ácido fosfórico H₃PO₄. Usando el buffer se preparó las soluciones madre con concentración de 100 mg/ml para la proteína, glucosa y metformina, a partir de estas soluciones se prepararon en viales de 1.5 ml las siguientes disoluciones:

Tabla 2.4

Concentraciones preparadas para diferentes sistemas

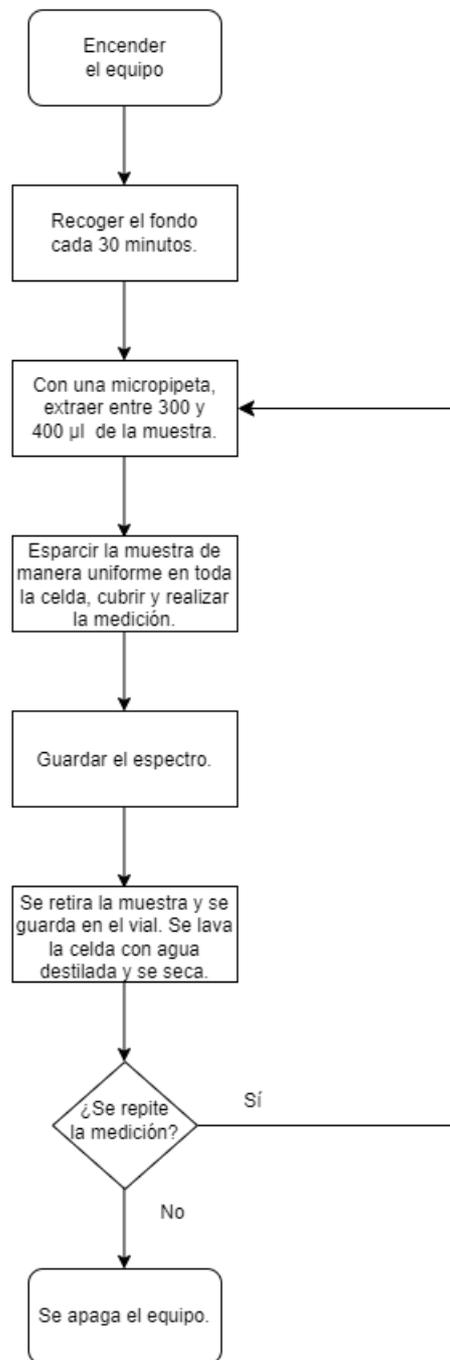
Sistemas	Unitario	Binario	Ternario
Perturbación de concentraciones	BSA/BHb 0 a 10 mg/ml (de 1 en 1)	BSA fija 10 mg/ml	
		Glucosa variable de 0 a 40 mg/ml	Metformina variable de 0 a 40 mg/ml
	Glucosa 0 a 40 mg/ml (de 5 en 5)	Glucosa fija 15 mg/ml	
		BSA/BHb variable de 0 a 10 mg/ml	
	Metformina 0 a 40 mg/ml (de 5 en 5)	Metformina fija 15 mg/ml	
		BSA variable de 0 a 10 mg/ml	
			BSA/BHb fija 10 mg/ml Glucosa fija 15 mg/ml Metformina variable de 0 a 40 mg/ml

2.2.2 Mediciones

Las lecturas se realizaron de manera transversal y longitudinal, donde los sistemas a varias concentraciones una vez preparados fueron reservados para tomar las mediciones cada día durante 5 días. El procedimiento de uso del equipo FTIR se puede observar en la figura 2.1.

Figura 2.23

Diagrama de flujo del procedimiento del uso del equipo FTIR.



2.2.3 Procesamiento de datos

Los datos obtenidos se guardan en formato “.spa” y se procede a usar

SpectraGryph, un software de espectroscopía óptica intuitivo y gratuito que permite darles

un formato particular a los datos, acotar el rango de interés o sustraer el espectro del buffer correspondiente para mejorar la visibilidad de las señales, minimizar el ruido, cambiar nombres de cada espectro, etc. Posteriormente, los datos se guardan en formato “.csv” y, finalmente, se organizan y archivan en formato “.xlsx” como se muestra en la figura 2.2.

Figura 2.3 4
Formato de datos preprocesados

Observaciones	Muestras									
	Gluc fija prot 1 mg/ml DI	Gluc fija prot 2 mg/ml DI	Gluc fija prot 3 mg/ml DI	Gluc fija prot 4 mg/ml DI	Gluc fija prot 5 mg/ml DI	Gluc fija prot 6 mg/ml DI	Gluc fija prot 7 mg/ml DI	Gluc fija prot 8 mg/ml DI	Gluc fija prot 9 mg/ml DI	Gluc fija prot 10 mg/ml DI
99,217104	99,100158	99,391535	99,176503	99,770425	99,773662	99,591466	105,51857	105,95759	105,3673	
99,403932	99,226936	99,460343	99,42158	99,034746	99,068031	99,867315	105,87112	106,18513	105,82412	
99,548748	99,499463	99,719395	99,580222	99,272619	99,223487	99,071002	106,17363	106,37852	106,28887	
99,281952	99,414147	99,861671	99,448424	99,135527	99,238328	99,868607	105,93925	106,24487	106,15091	
99,005955	99,047115	99,463544	99,156884	99,750907	99,00658	99,573573	105,66839	106,03048	105,81979	
99,164559	99,10496	99,					105,78969	106,00282	105,86874	
99,391736	99,596189	99,					106,01839	106,37206	106,27957	
99,60708	99,941204	100,					106,4619	106,89346	106,838	
99,794749	99,867763	100,					106,64232	107,12095	107,03579	
99,573258	99,482934	99,					106,20699	106,73765	106,66077	
99,697544	99,599316	99,					106,3516	106,83341	106,67551	
100,21073	100,14866	100,35224	99,800536	99,844999	1,00E-02	99,854352	106,66944	107,24097	106,86764	
100,0241	100,09435	100,29649	99,879886	99,852934	99,972843	99,666723	106,68488	107,18249	106,80588	
99,619918	99,939795	99,929932	99,78432	99,608523	99,776089	99,35853	106,51474	106,91018	106,58438	
99,564868	100,07075	99,851002	99,71231	99,496656	99,709731	99,188012	106,62156	107,12882	106,67787	
99,479517	1,00E-02	99,90429	99,640059	99,386259	99,463418	99,260174	106,93266	107,25842	106,90168	
99,722328	99,959425	100,10166	99,805158	99,573258	99,569193	99,694535	107,21113	107,49037	107,13083	
100,04908	100,26757	100,45977	99,938862	99,950632	1,00E-02	99,975018	107,53178	107,93566	107,58048	
1784,547	100,07107	100,51957	100,61052	100,18864	99,896215	100,20741	99,761657	107,48561	107,79048	
1762,619	99,941297	100,46974	100,56477	100,58827	99,851385	100,26037	99,703306	106,96795	107,31086	
1760,69	100,0426	100,54041	100,91163	100,89229	100,3329	100,62764	100,13691	106,88542	107,59215	
1758,762	100,10637	100,55535	101,10692	100,77203	100,52516	100,71066	100,33744	107,11452	107,87176	

2.3 Implementación de análisis PCA-2DCos

Para el análisis de Componentes Principales (PCA), inicialmente se realizaron pruebas utilizando el programa Statgraphics XIX en su versión de prueba gratuita. Este programa permitió analizar los datos obtenidos mediante la técnica de FTIR como se puede ver en la figura 2.2, donde la variación de los porcentajes de transmitancia debido a las perturbaciones de concentración y tiempo son consideradas las muestras y, las frecuencias de onda son las observaciones. El objetivo del análisis fue evaluar cómo las perturbaciones en la concentración y el tiempo contribuían a la varianza de los datos y determinar cuáles factores tenían un mayor impacto en los pesos de los componentes.

Con el fin de superar la restricción de la licencia de prueba de 30 días de Statgraphics XIX y trabajar con una mayor cantidad de datos, se decidió implementar el análisis PCA utilizando la librería sklearn de Python. Este cambio permitió realizar el análisis sin limitaciones, además de ofrecer una herramienta más flexible y adecuada para manejar grandes volúmenes de información. Dado que Statgraphics XIX es un software diseñado

para uso comercial, los resultados obtenidos con este programa son considerados confiables, por lo que se utilizó como herramienta de validación para comparar los resultados generados con Python. La validación realizada utilizando los datos de BSA fija Glucosa variable por concentración, presentada en los anexos, muestra que ambos métodos produjeron los mismos resultados, garantizando la fiabilidad del análisis realizado con Python.

Figura 2.4 5

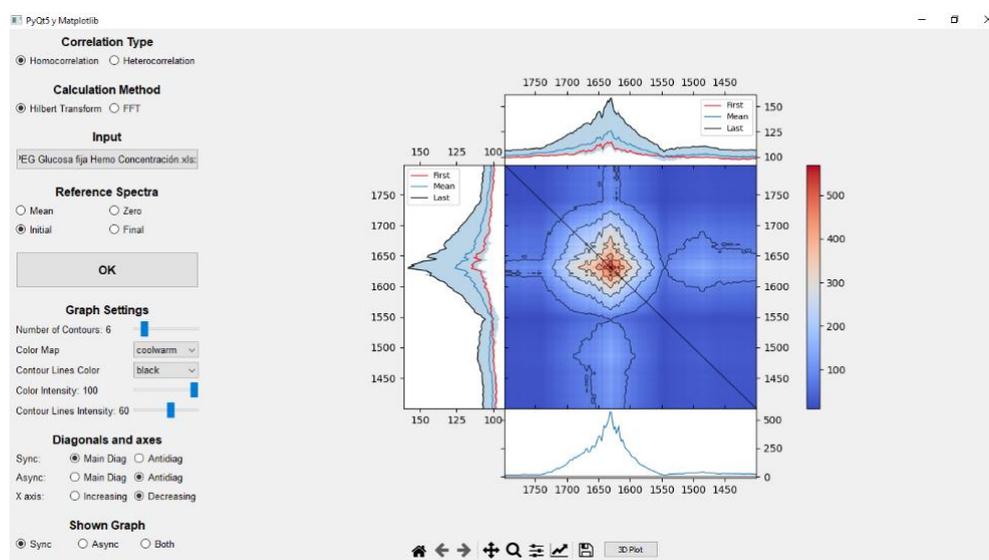
Código en Python donde se realiza el PCA con sklearn leyendo datos de una hoja en Excel.

```
data = excel_data.parse(sheet_name, header=1, decimal=",")
# Extraer frecuencias de onda y datos
wavelengths_column = data.columns[0]
frecuencias_onda = data[wavelengths_column].astype(float).values
data_cleaned = data.drop(columns=[wavelengths_column]).astype(float)
# Centrar los datos en la media sin estandarizar
data_centered = data_cleaned - data_cleaned.mean()
# Realizar el PCA
pca = PCA()
pca.fit(data_centered)
```

Una vez realizado el análisis PCA, también se realizó el 2DCos utilizando el software Py2DCos de libre acceso desarrollado por De la Paz, (2023). Se utilizaron los documentos de Excel utilizados para el PCA para ver la correlación de las frecuencias de onda que presentaron perturbaciones a través de la concentración y del tiempo.

Figura 2.5

Ventana del software Py2DCos.



2.4 Desarrollo de la interfaz para análisis PCA

La aplicación desarrollada está diseñada para analizar datos espectroscópicos obtenidos mediante FTIR, el diseño de la interfaz prioriza la interacción intuitiva, brindando a los usuarios un acceso dinámico y flexible al análisis de componentes principales. El procedimiento de uso de la aplicación inicia con la carga de archivos en formato Excel (.xlsx), desde los cuales se extraen y procesan los datos. Todas las hojas del archivo se analizarán por default, sin embargo, existe la opción de des seleccionar las hojas que se desee descartar. Estos datos son sometidos a un análisis PCA, lo que permite identificar las principales fuentes de variación y determinar las longitudes de onda clave asociadas a los grupos funcionales afectados. Inmediatamente, la aplicación ofrece opciones para visualizar resultados en tiempo real mediante gráficas interactivas y finalmente, se habilita el botón de descargar los análisis en distintos formatos, como Excel y PDF, adaptándose a las necesidades de los usuarios.

Figura 2.6 7
Mockup inicial de aplicación web desarrollada

Análisis PCA desde archivos Excel

Sube un archivo Excel para realizar el análisis PCA.

Cargar archivo Excel

 Drag and drop file here
Limit 200MB per file • XLSX
Browse files

 PEG Glucosa fija Hemo Concentración - copia.xlsx 164.3KB ✕

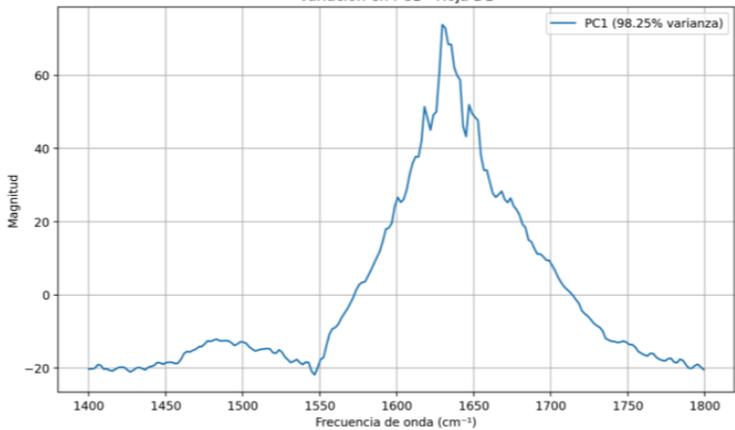
Seleccionar hojas a analizar

D1 ✕
D2 ✕
D3 ✕
D4 ✕
D5 ✕
✕ ▾

Resultados para la hoja: D1

	Componente	Varianza acumulada (%)	Longitud de onda (cm ⁻¹)	Grupo funcional
0	PC1	98.2513	1,785.76	Desconocido
1	PC2	99.4323	1,789.617	Desconocido
2	PC3	99.6781	1,797.331	Desconocido
3	PC4	99.7639	1,785.76	Desconocido
4	PC5	99.8362	1,799.259	Desconocido
5	PC6	99.8862	1,795.403	Desconocido
6	PC7	99.9336	1,795.403	Desconocido
7	PC8	99.9667	1,799.259	Desconocido
8	PC9	99.9872	1,787.689	Desconocido
9	PC10	100	1,793.474	Desconocido

Variación en PC1 - Hoja D1



Descargar Gráfica de D1 (PNG)

.

Descargar Tablas en Excel

Descargar Tablas en PDF

Reiniciar para un nuevo análisis

La aplicación incluye módulos integrados para la visualización de gráficos PCA y para la generación de resultados descargables en diferentes formatos, PNG para cada imagen disponible a la selección del usuario y un archivo final que incluye las tablas de resultados presentado en PDF y .xlsx. Tecnologías como Python, Streamlit, y librerías especializadas (NumPy, Pandas, scikit-learn, matplotlib, y pdf) se emplearon para desarrollar esta solución. Además, se diseñó con un enfoque modular que permite realizar múltiples análisis consecutivos sin necesidad de recargar la página.

2.5 Costos asociados

Se emplearon tanto licencias gratuitas de software de pago como un software de libre acceso, lo que permitió realizar los análisis computacionales sin costos adicionales. En cuanto a la parte experimental, los gastos estuvieron relacionados principalmente con la adquisición de los reactivos necesarios y el uso del equipo de espectroscopía. Sin embargo, dado que la mayoría de las instituciones de investigación y universidades cuentan con equipos de espectroscopía, como IR o UV-Vis, los costos experimentales se limitan exclusivamente a los reactivos utilizados. Esta combinación de herramientas gratuitas y recursos institucionales contribuyó a reducir significativamente los gastos asociados al desarrollo del caso de estudio.

Capítulo 3

3. Resultados y análisis

3.1 Análisis del Diseño Experimental

Se estudió la glicación no enzimática de proteínas inducida por un exceso de glucosa, evaluando el efecto protector de un fármaco y la influencia del agente congestionante, además se buscó comprender los mecanismos químicos involucrados. Para ello, se realizaron 198 experimentos espectroscópicos descritos en el capítulo 2, generando aproximadamente 44,200 datos de medida para un sistema en un medio específico de una proteína. La gran cantidad de datos obtenidos hizo necesaria la creación de un algoritmo de toma de decisiones (figura 3.1) que combinara FTIR-2DCOS con el análisis de componentes principales (PCA), una herramienta estadística capaz de identificar únicamente las variaciones más significativas en el sistema ante las perturbaciones de concentración y tiempo causadas por los ligandos añadidos. La estrategia desarrollada se ve en figura 3.1 donde se señalan los pasos a seguir para el análisis de sistemas perturbados usando PCA-FTIR-2DCOS.

3.2 Estudio comparativo de la BHb con y sin PEG-8

Aunque el rango de actividad característico de la hemoglobina se encuentra entre 1500 y 1700 cm^{-1} (Chen & Spiro, 2002), el pretratamiento de datos incluyó un rango ampliado desde 1400 cm^{-1} para mejorar la visualización. En las Tablas 3.1 y 3.2 se presentan los pesos obtenidos para cada sistema de hemoglobina sin y con PEG, respectivamente, lo que permitió seleccionar los archivos relevantes para el análisis. En los casos donde las concentraciones mostraron un peso representativo similar, indicando una contribución equitativa a los cambios observados, se optó por mantener la misma concentración de los sistemas posteriores.

Figura 3.1

Estrategia desarrollada para el análisis de sistemas complejos perturbados

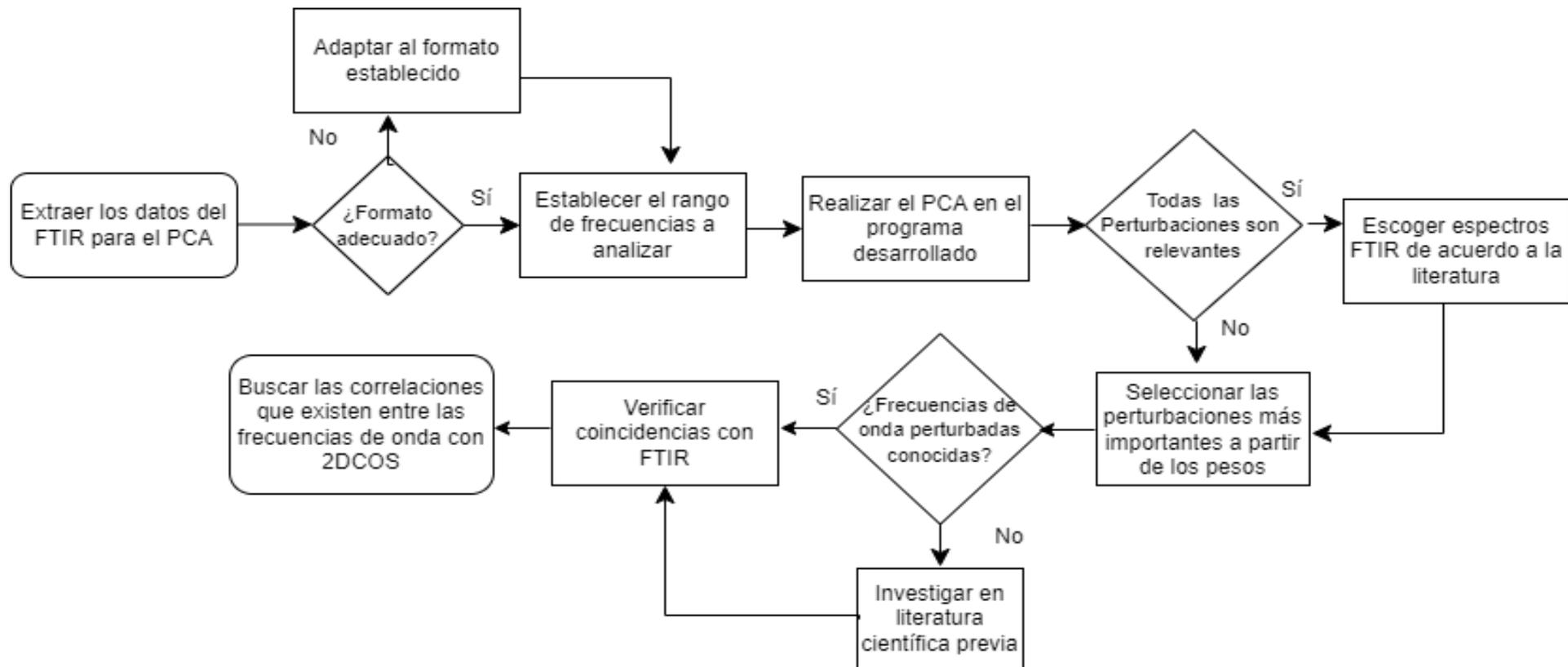


Tabla 3.1

Resultados de análisis PCA para el sistema unitario, binario y ternario de BHb

UNITARIO		BINARIO			TERNARIO	
Tiempo		Tiempo			Tiempo	
mg/ml	PC1	mg/ml	PC1		mg/ml	PC1
1	0,35	5	0,33		5	0,15
2	0,32	10	0,31		10	0,21
3	0,34	15	0,35		15	0,27
4	0,34	20	0,34		20	0,31
5	0,30	25	0,36		25	0,37
6	0,32	30	0,36		30	0,41
7	0,31	35	0,39		35	0,46
8	0,30	40	0,37		40	0,50
9	0,31	-			-	
10	0,28					
Concentración		Concentración			Concentración	
d	PC1	d	PC1	PC2	d	PC1
1	0,29	1	-0,095	0,44	1	0,48
2	0,58	2	-0,027	0,51	2	0,43
3	0,51	3	-0,007	0,42	3	0,44
4	0,42	4	0,992	0,10	4	0,45
5	-0,40	5	-0,074	0,60	5	0,43

Tabla 3.2

Resultados de análisis PCA de sistema unitario, binario y ternario de BHb con PEG

UNITARIO		BINARIO			TERNARIO	
Tiempo		Tiempo			Tiempo	
mg/ml	PC1	mg/ml	PC1		mg/ml	PC1
1	0,07	5	0,35		5	0,15
2	0,28	10	0,34		10	0,22
3	0,31	15	0,39		15	0,26
4	0,35	20	0,37		20	0,31
5	0,33	25	0,32		25	0,36
6	0,36	30	0,33		30	0,41

7	0,37		35	0,36		35	0,45
8	0,29		40	0,34		40	0,49
9	0,31		-			-	
10	0,35		-			-	
Concentración			Concentración			Concentración	
d	PC1	PC2	d	PC1	PC2	d	PC1
1	0,19	0,57				1	0,44
2	0,29	0,39	1	0,67	-0,053	2	0,44
3	0,39	0,22	2	0,68	-0,039	3	0,45
4	0,18	0,49	3	-0,055	0,92	4	0,45
5	0,83	-0,48	4	0,29	0,38	5	0,45

Las figuras 3.2 y 3.3 muestran el espectro de la hemoglobina en estado fundamental, sin y con PEG-8, respectivamente. En ambos casos, se identifican los picos característicos de las bandas amida I y amida II, siendo la banda amida I la más sensible a los cambios en la estructura secundaria de las proteínas. La tabla 3.3 presenta las frecuencias de onda referenciales características de la proteína (Ioannou & Varotsis, 2019) y las observadas para el sistema unitario. Las Figuras 3.4 y 3.5 muestran los análisis de componentes principales (PCA) realizados para la BHb, las cargas más alejadas de cero representan las mayores variabilidades a lo largo del tiempo (Bro & Smilde, 2014). Este comportamiento refleja los cambios estructurales principales de la proteína en función de la perturbación de tiempo.

Tabla 3.3

Frecuencias de onda de estructuras para hemoglobina en estado fundamental

Estructuras	Referencia (cm-1)	Sin PEG 8000 (cm-1)	Con PEG 8000 (cm-1)
Hélices alfa	1649-1660	1647-1658	1653
Láminas beta	1610-1638	1631	1628-1639
Giros beta	1660-1680	-	-
Amida II	1500-1600	1548	1548

Figura 3.2

Sistema unitario de BHb a 10 mg/ml

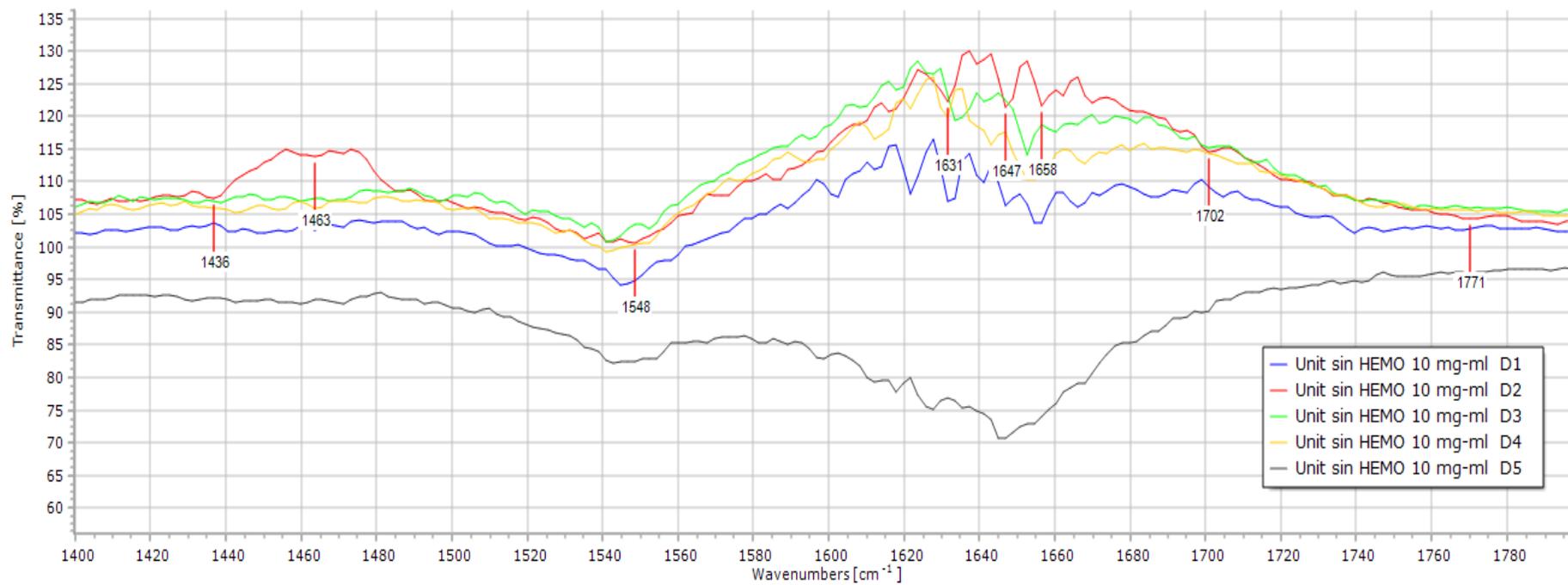


Figura 3.3

Sistema unitario de BHb a 10 mg/ml con PEG-8

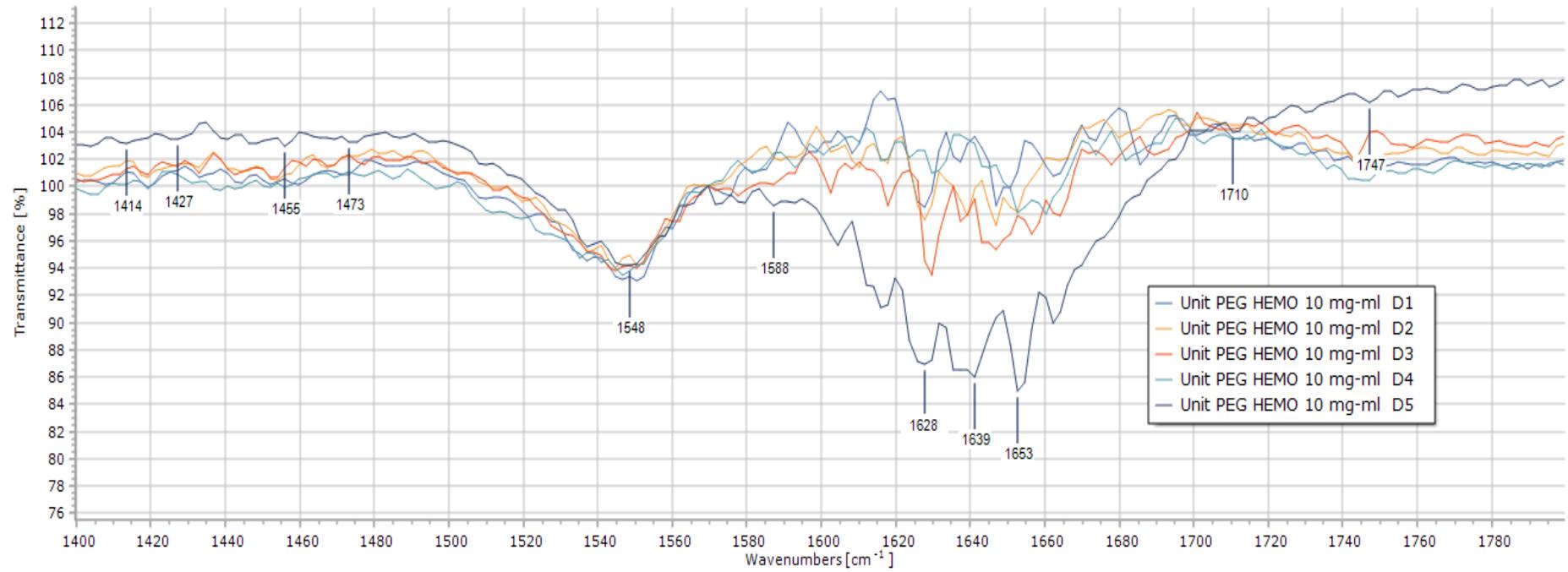


Figura 3.4

PCA del sistema unitario de BHb a 10 mg/ml

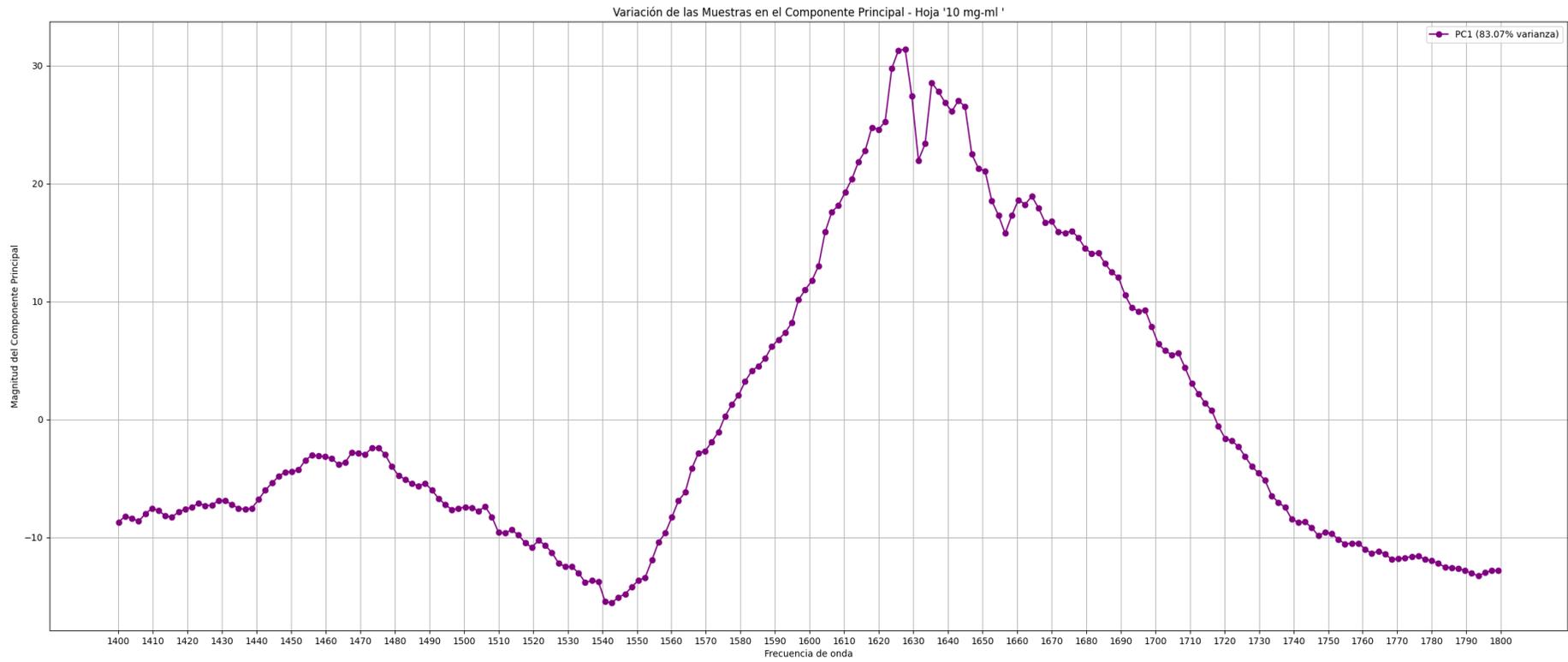
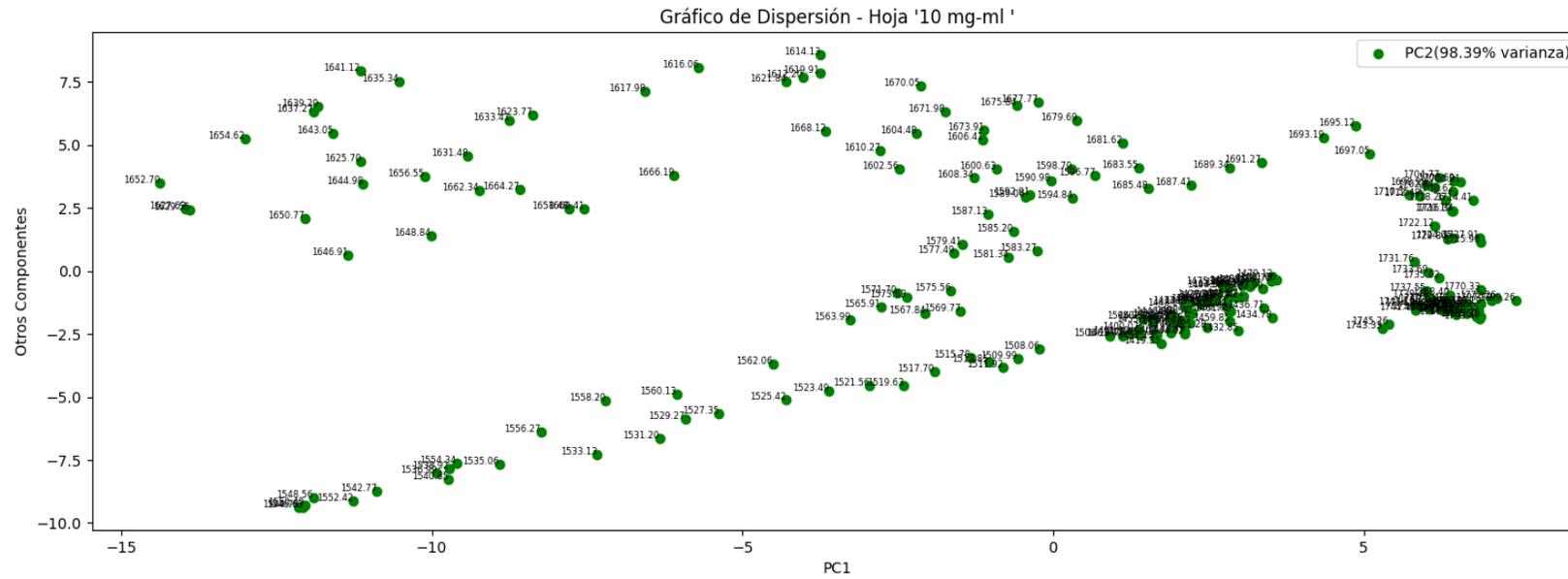


Figura 3.5

PCA del sistema unitario de Bhb a 10mg/ml con PEG-8



En la tabla 3.4 se reunió los resultados del PCA (Figuras 3.7, 3.9 y 3.11) para los sistemas conjugados de la BHb, se observó que en los sistemas binarios la interacción proteína-ligando, como en el caso de BHb con metformina o glucosa, produce cambios significativos en las Amidas I y II, observados principalmente en las frecuencias relacionadas con carbonilos y aminos. Estos cambios estructurales se evidencian en los espectros (Figura 3.6 con metformina y Figura 3.8 con glucosa), donde en el sistema con glucosa las láminas beta disminuyen en intensidad, mientras que las hélices alfa incrementan, comportamiento que coincide con los picos de correlación cruzada en el análisis 2DCOS (Figura 3.13). Aunque generalmente la glicación provoca un efecto contrario, estudios previos en BHb también han demostrado que este patrón puede atribuirse al proceso de glicación (Li et al., 2023).

En el sistema ternario BHb-Glucosa-Metformina, las frecuencias relacionadas con la interacción proteína-ligando muestran comportamientos similares a los del sistema BHb-Metformina donde si bien también se encuentra una dinámica entre las α -hélices y β -láminas, la carga de esta interrelación es aproximadamente 3 veces menor que la de BHb-Metformina lo que sugiere que la metformina actúa como un competidor en sitios de glicación, modulando estas interacciones y disminuyendo los efectos estructurales de la glicación en la hemoglobina. Esto es consistente con los picos observados en el análisis 2DCOS, donde las correlaciones cruzadas resaltan cambios estructurales similares a los del sistema binario BHb-Metformina (Figuras 3.14 y 3.15). Así mismo la Figura 3.12 muestra el análisis bidimensional de la BHb en estado fundamental a través del tiempo donde se ubican los cambios naturales que facilitan la comparación ante la interacción con los ligandos.

Tabla 3.4

Resultados de las frecuencias con mayor carga en PCA para sistemas binario y ternario en BHb

Binario							
Sin PEG-8				Con PEG-8			
BHb-Glucosa		BHb-Metformina		BHb-Glucosa		BHb-Metformina	
Frecuencias (cm ⁻¹)	Cargas						
1633	26			1650	20		
1641	23	1560-1580	65	1648	18	1565-1567	61
1646	19			1629	16		
Ternario							
Sin PEG-8				Con PEG-8			
Frecuencias (cm ⁻¹)		Cargas		Frecuencias (cm ⁻¹)		Cargas	
1560-1580		63		1563-1569		64	
1644		25		1648		22	

En el sistema donde se simula congestión molecular, los análisis PCA muestran una dinámica similar al medio sin congestión, aunque con ligeras variaciones en las frecuencias y principalmente en la disminución de las cargas, esto se complementa al observarlos en conjunto con los espectros FTIR y 2DCos (Figuras 3.15-3.24). Como se observan los mismos cambios se sugiere que el efecto de protección sobre la estructura de la BHb por parte del fármaco, así como la glicación sigue sucediendo en el medio congestionado. Sin embargo, debido a la disminución de intensidad de las cargas que representan los cambios debido a perturbaciones se infiere la disminución de interacción.

Figura 3.6

Sistema binario de BHB a 10 mg/ml y Metformina 40 mg/ml

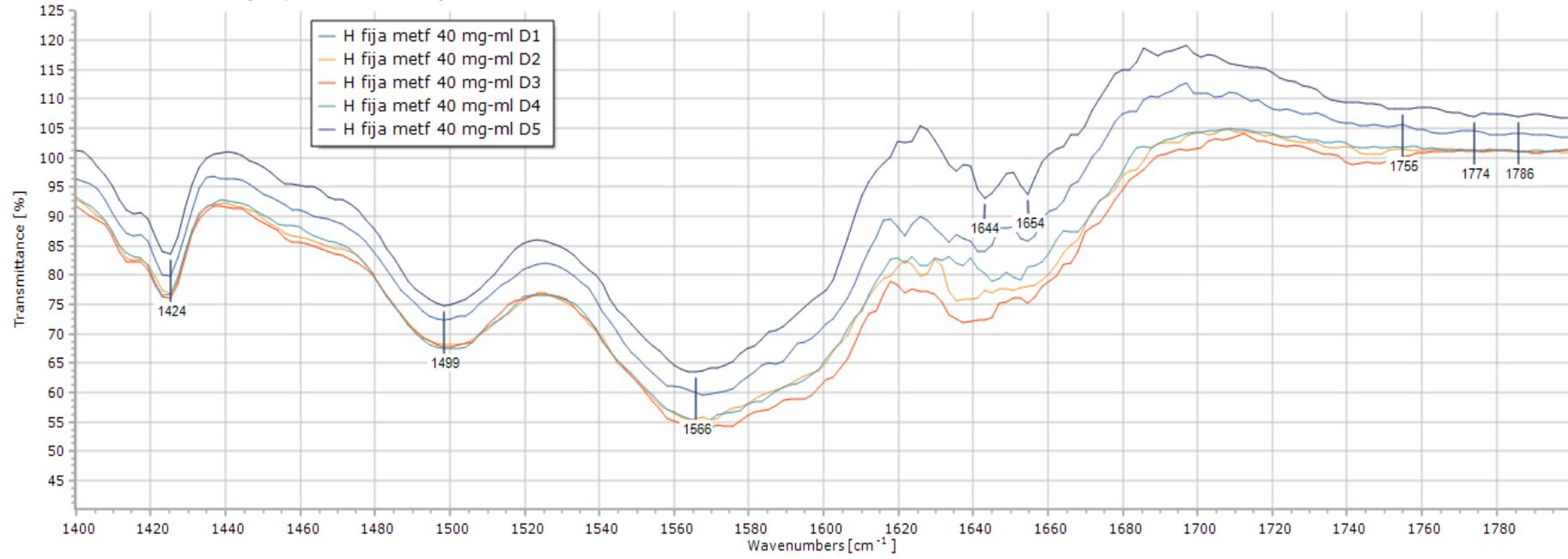


Figura 3.7

PCA del sistema binario de BHB a 10 mg/ml y Metformina 40 mg/ml

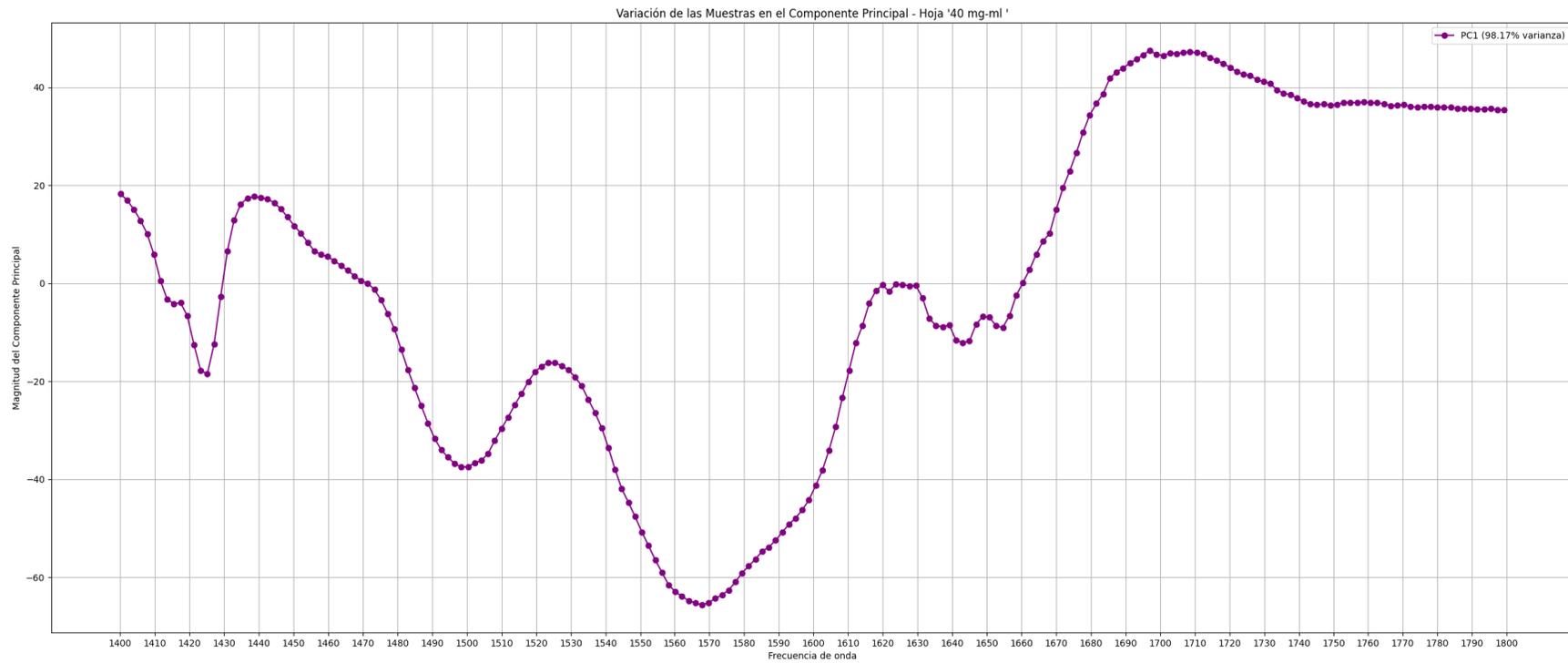


Figura 3.8

Sistema Binario BHb 10 mg/ml con Glucosa 15mg/ml

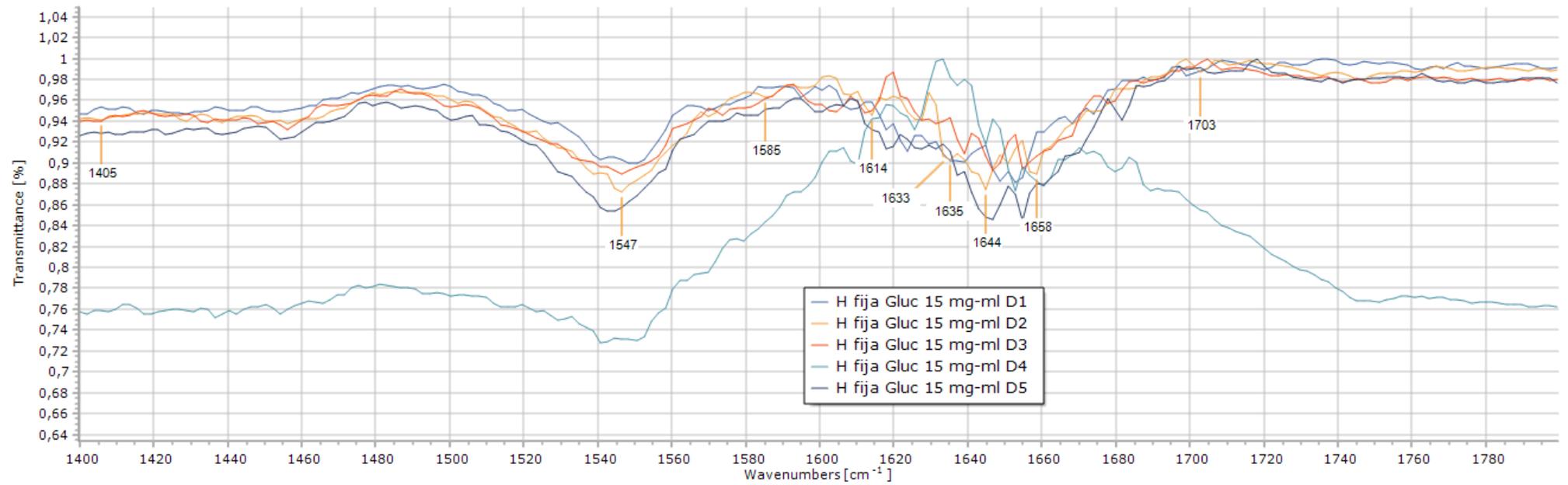


Figura 3.9

PCA del sistema binario de BHB a 10 mg/ml y Glucosa 15 mg/ml

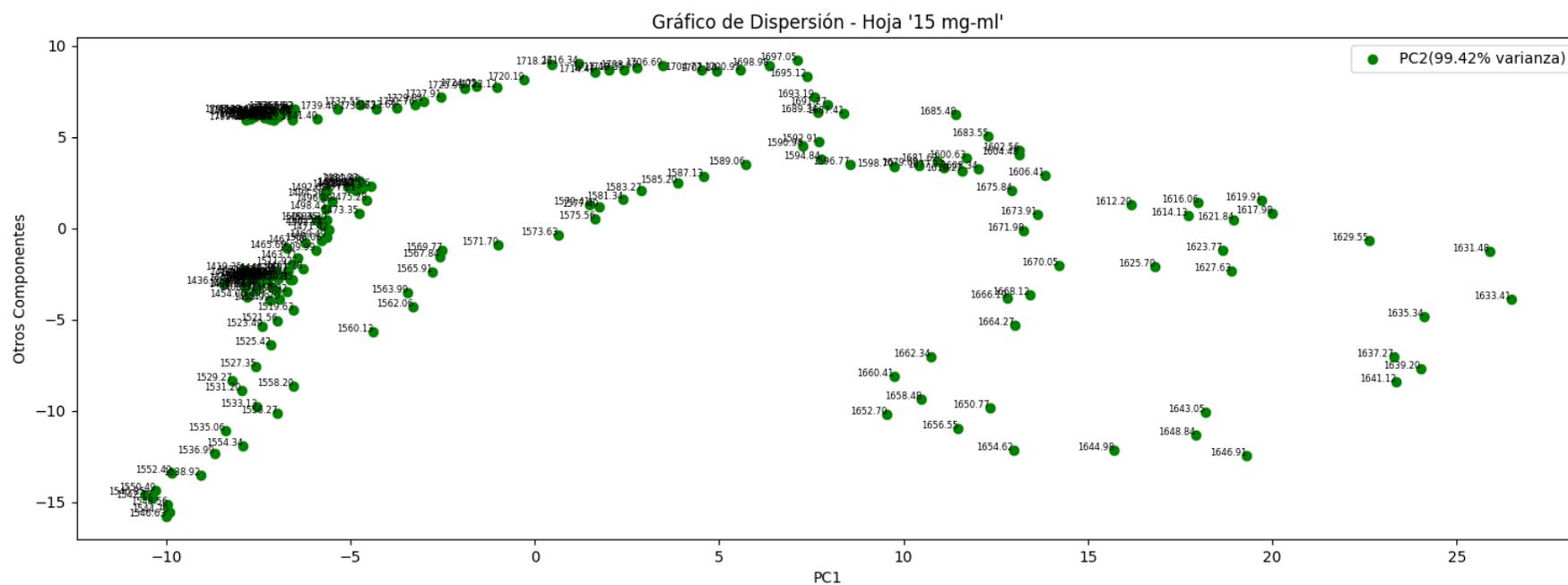


Figura 3.10

Sistema Ternario BHB 10 mg/ml con Glucosa 15 mg/ml y Metformina 40 mg/ml

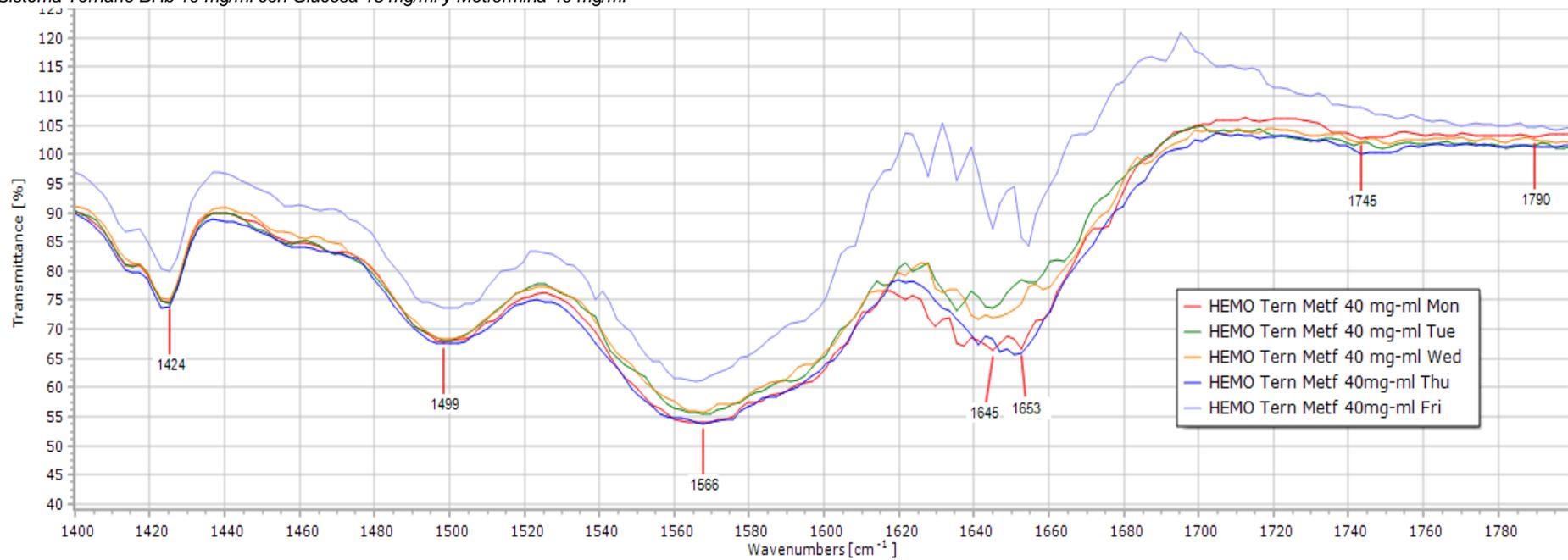


Figura 3.11

PCA del sistema ternario de BHB a 10 mg/ml, Glucosa 15 mg/ml y Metformina 40 mg/ml

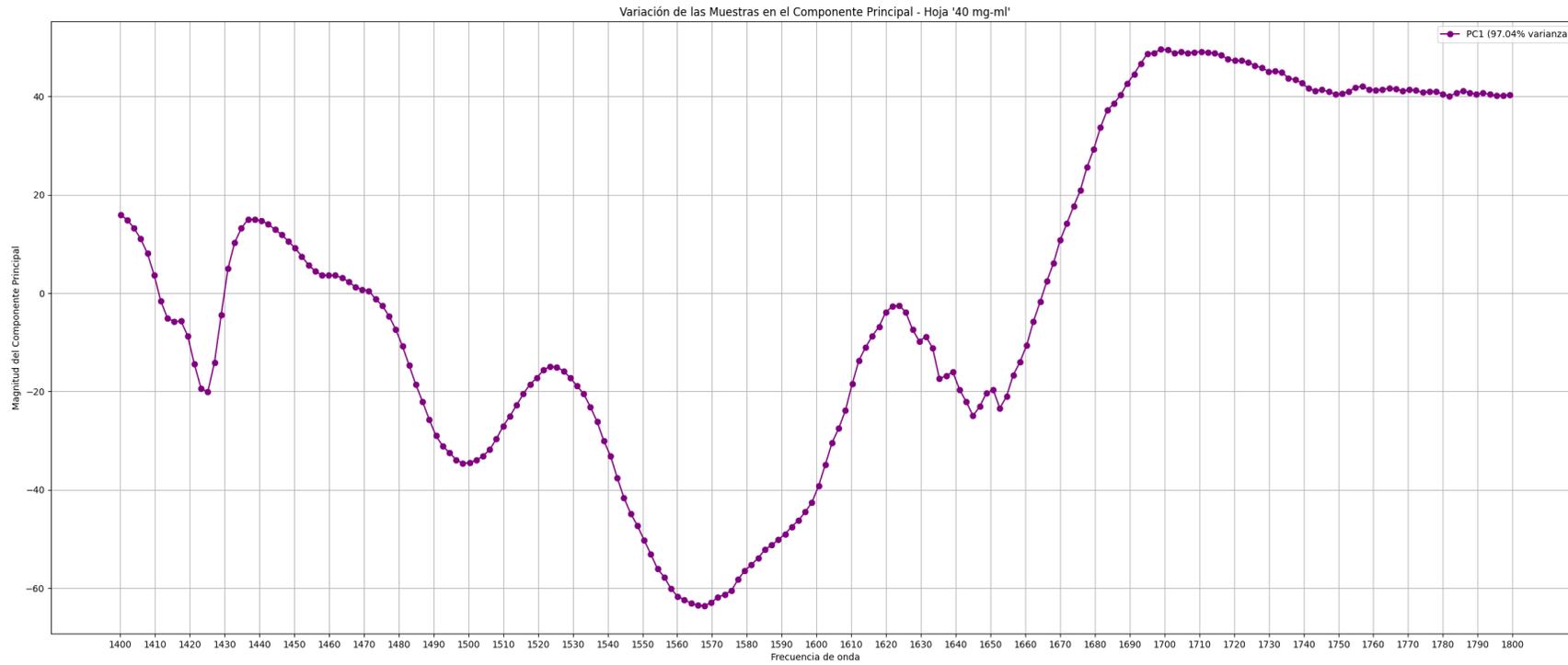


Figura 3.12

2DCOS BHb 10 mg/ml por días

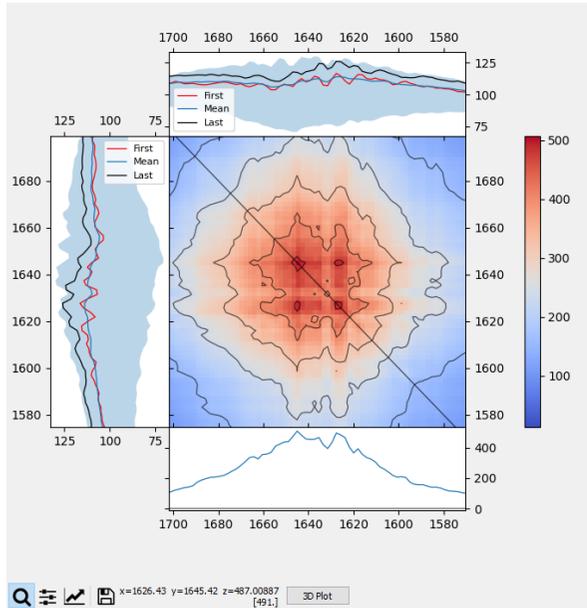


Figura 3.13

2DCOS BHb 10 mg/ml y Glucosa 15 mg/ml por días

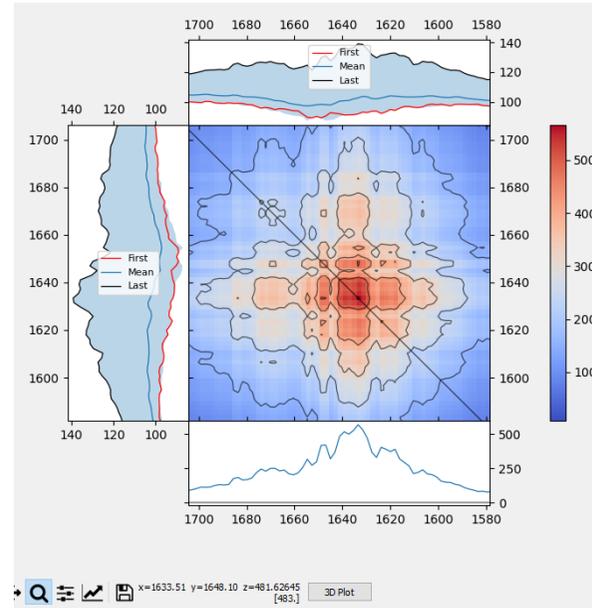


Figura 3.14

2DCOS BHb 10 mg/ml y Metformina 40 mg/ml por días

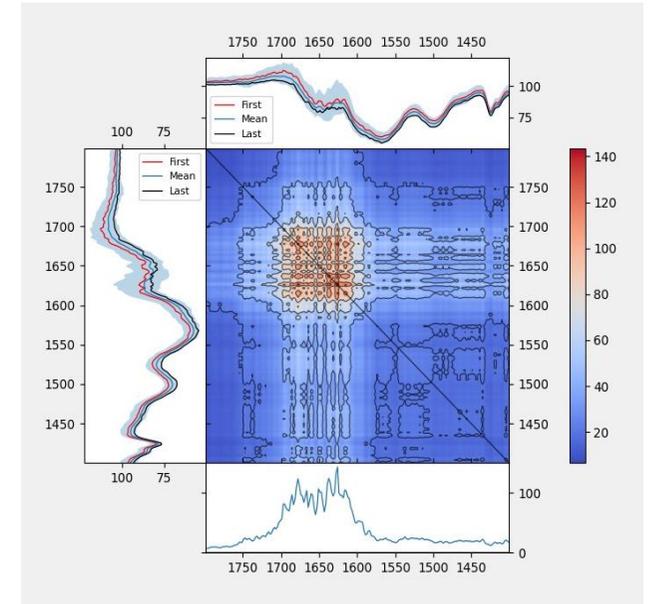


Figura 3.15

2DCOS BHb 10 mg/ml, Glucosa 15 mg/ml y Metformina 40 mg/ml por días

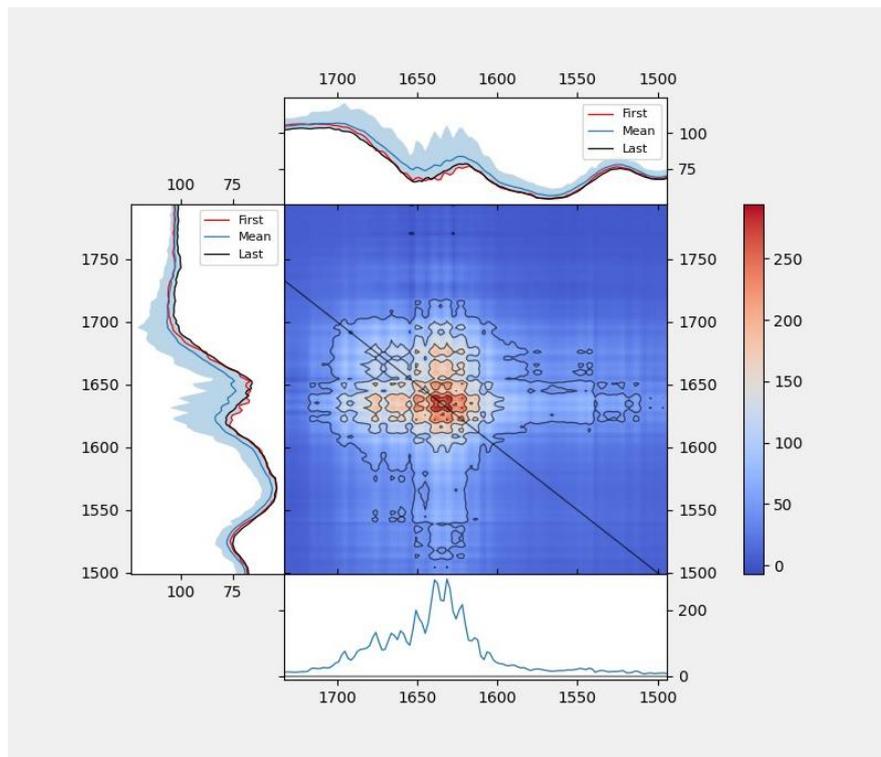


Figura 3.16

Sistema binario Bhb 10 mg/ml con Metformina 40 mg/ml con PEG-8

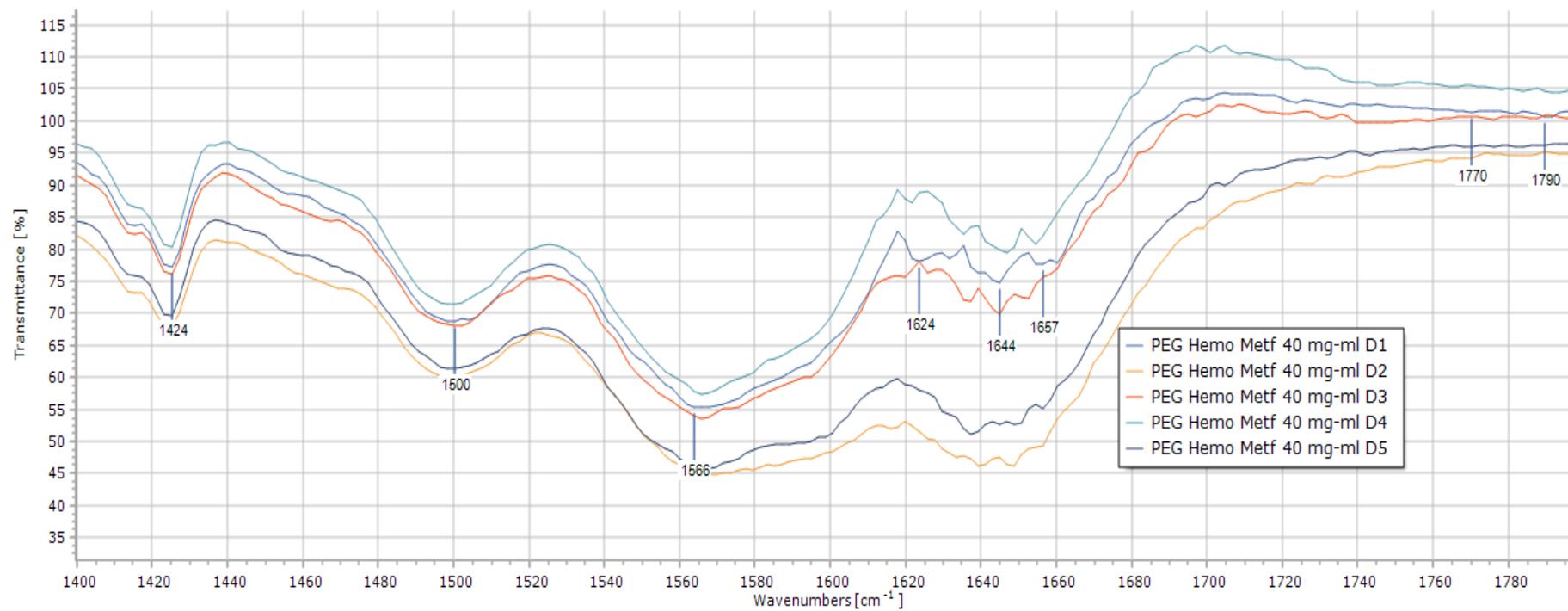


Figura 3.17

PCA del sistema binario de BHB a 10 mg/ml y Metformina 15 mg/ml con PEG-8

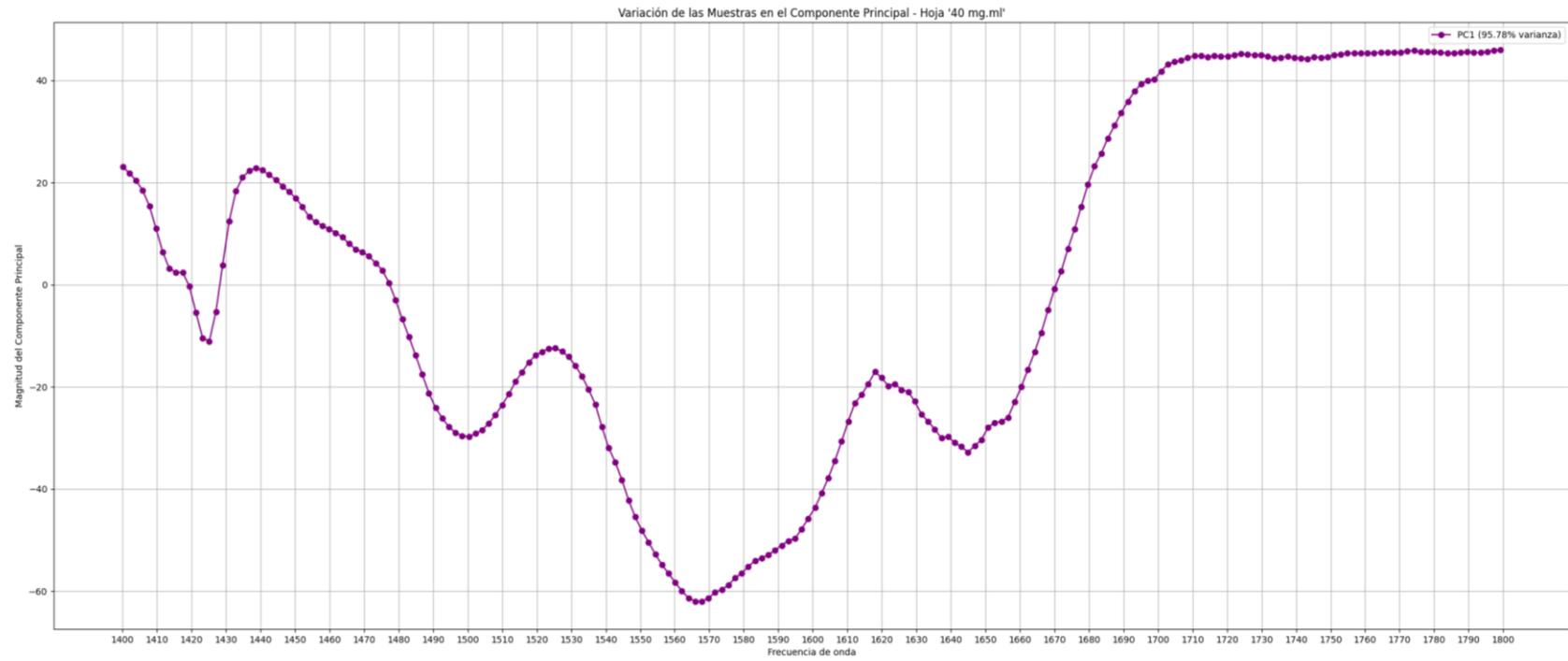


Figura 3.18

Sistema Binario BHb 10 mg/ml con Glucosa 15mg/ml con PEG-8

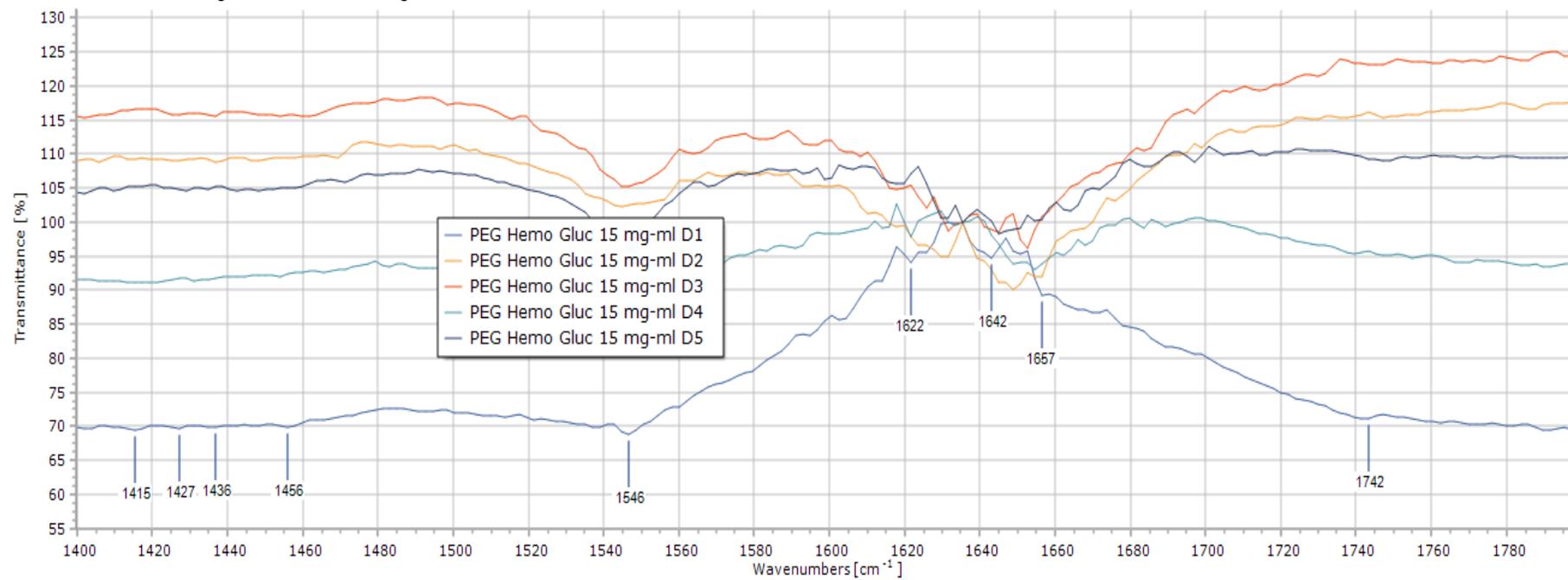


Figura 3.19

PCA del sistema binario de BHB a 10 mg/ml y Glucosa 15 mg/ml con PEG-8

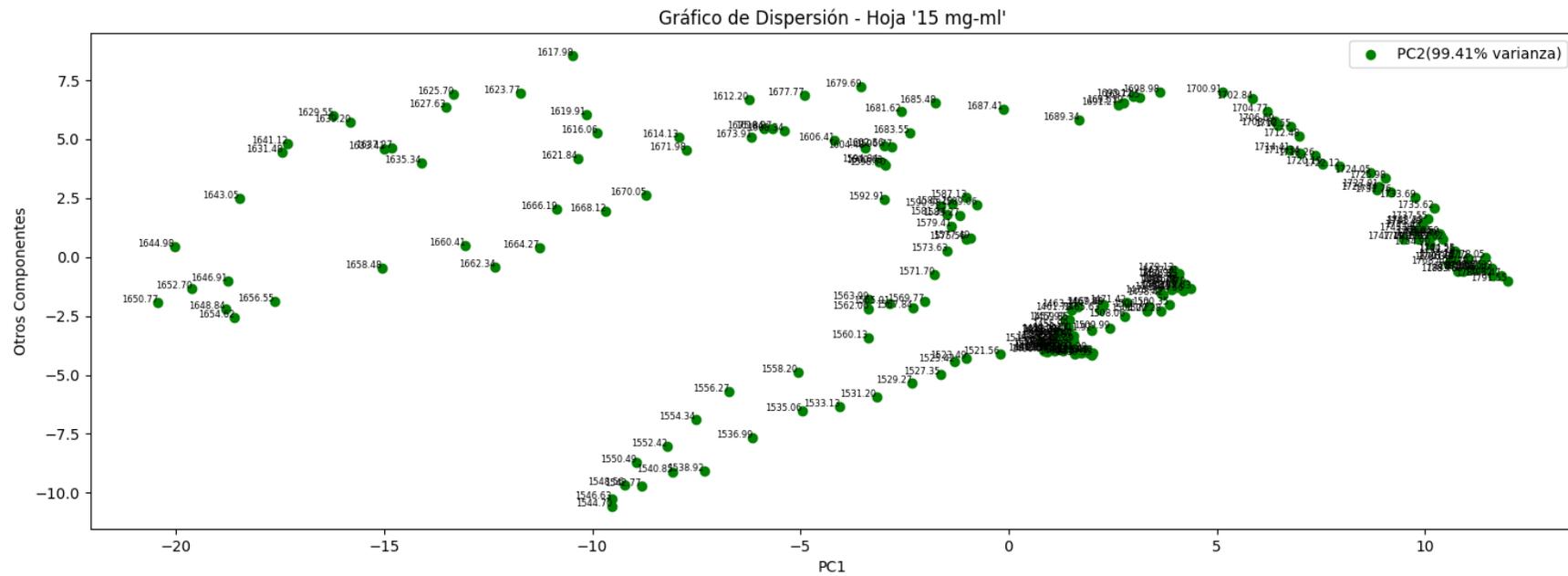


Figura 3.20

Sistema Ternario BHB 10 mg/ml con Glucosa 15 mg/ml y Metformina 40 mg/ml con PEG-8

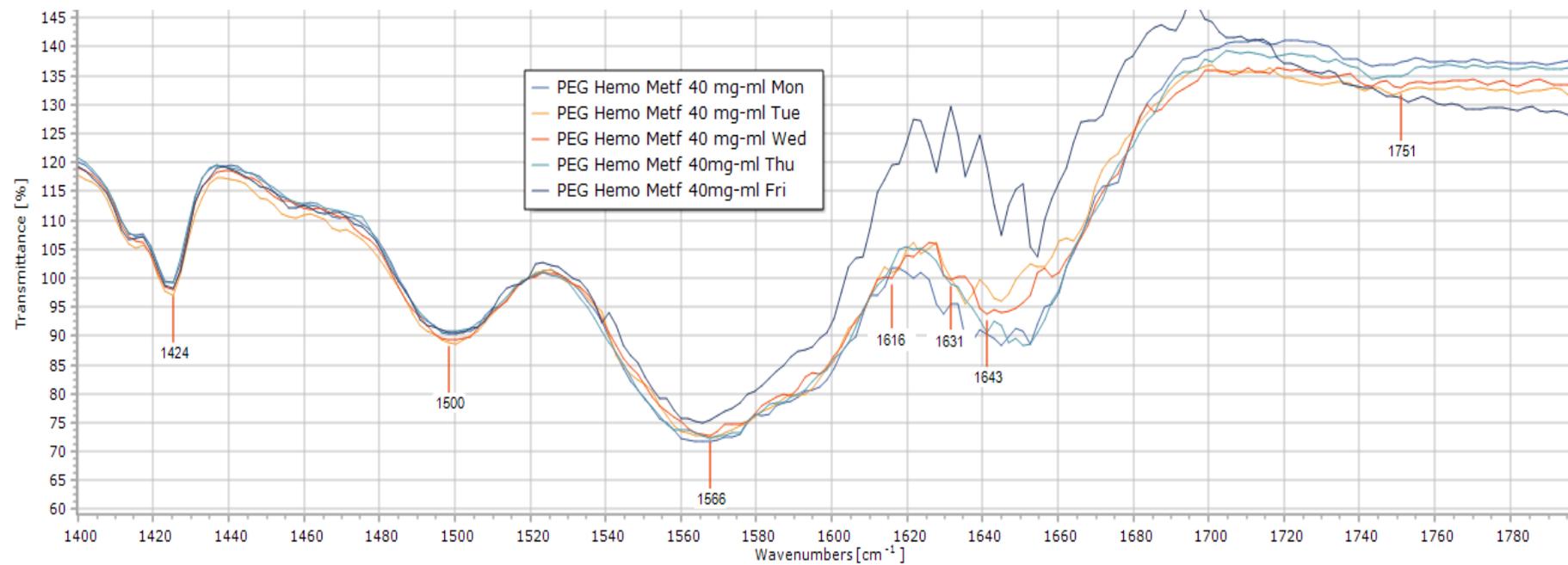


Figura 3.21

PCA del sistema ternario de BHB a 10 mg/ml, Glucosa 15 mg/ml y Metformina 40 mg/ml con PEG-8

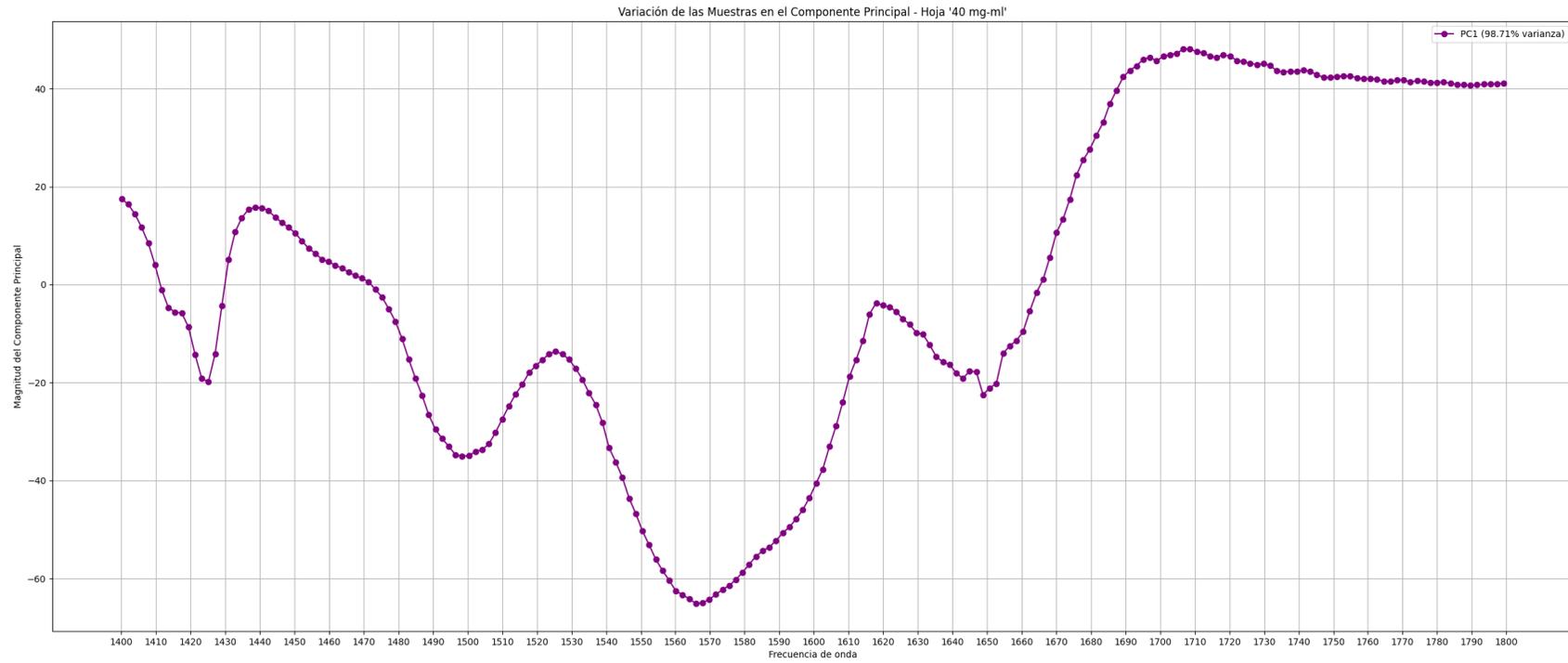


Figura 3.22

2DCOS PEG Bhb 10 mg/ml por días

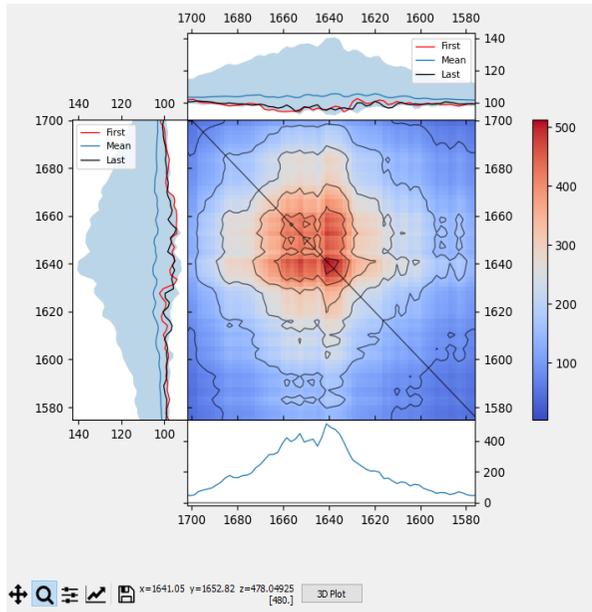


Figura 3.23

2DCOS PEG Bhb 10 mg/ml y Glucosa 15 mg/ml por días

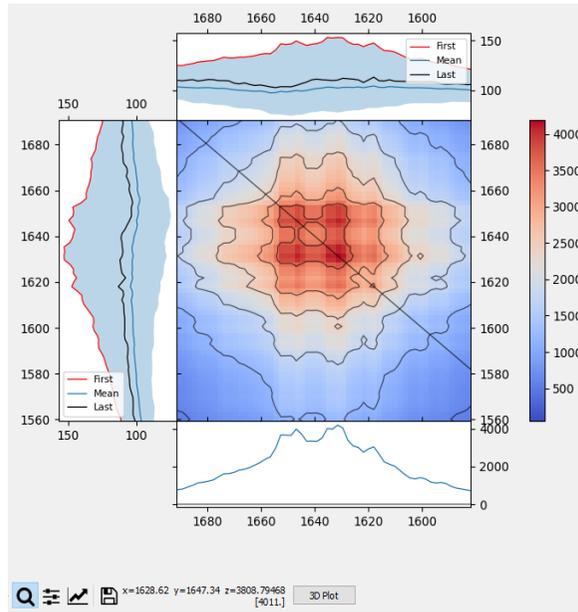


Figura 3.24

2DCOS PEG Bhb 10 mg/ml y Metformina 40 mg/ml por días

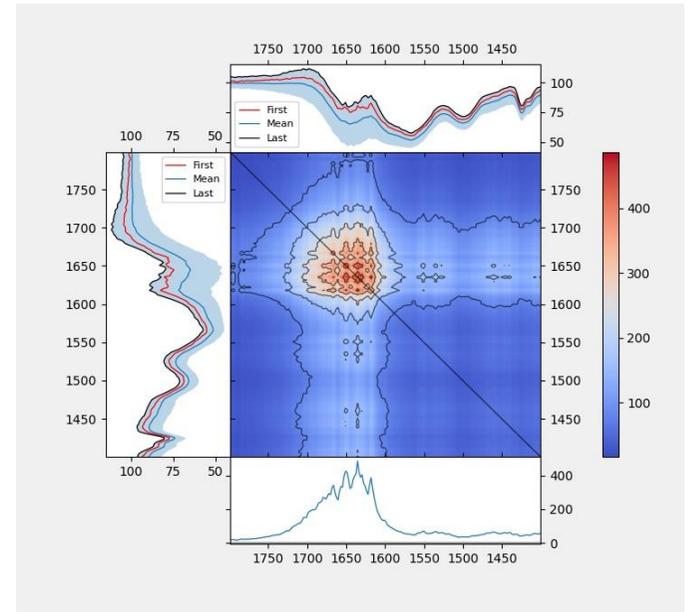
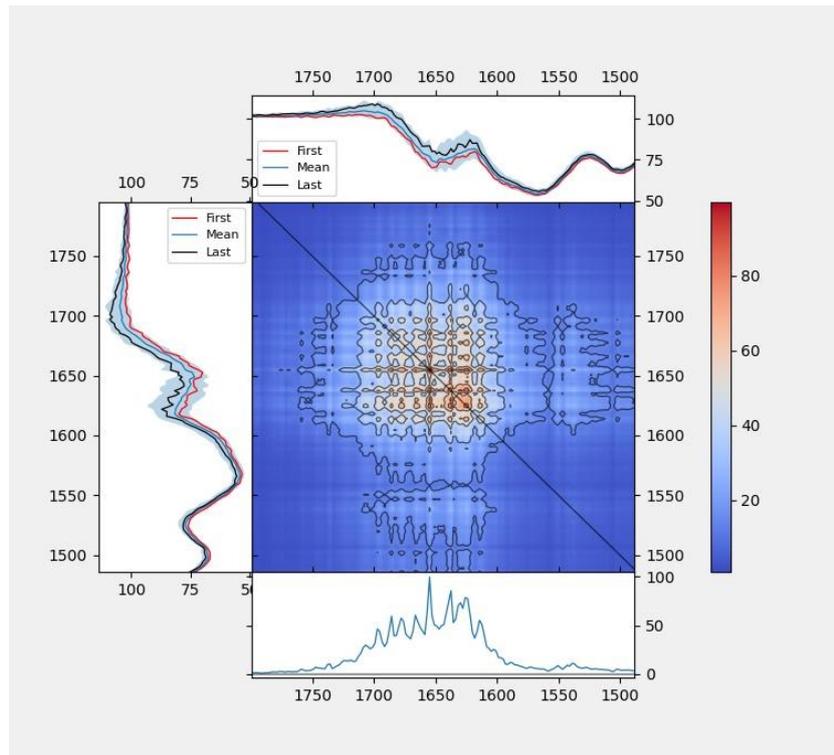


Figura 3.25

2DCOS PEG BHb 10 mg/ml, Glucosa 15 mg/ml y Metformina 40 mg/ml por días



3.3 Influencia del uso de PEG-8 en sistemas perturbados

Se usa los resultados del estudio de la BSA y el sistema de la BHb con PEG-8 para comprobar el papel del medio congestionado en las biomoléculas interactuando con los ligandos, con la cual se realizaron los mismos análisis representados en las tablas 3.5 y 3.6. En la tabla 3.3 se observó ligeras des localaciones de las estructuras características de la BHb en comparación con el sistema sin congestionante. Mientras que, los datos extraídos de las figuras 3.17, 3.19 y 3.21 indicaron la influencia del PEG-8 mediante las cargas, ya que resultaron menores que las obtenidas en el sistema solo con buffer. Para la BHb, se identificaron tanto la glicosación debido a la glucosa y el efecto Quimi-protector del fármaco al haber integrado el PCA con FTIR (Figuras 3.16, 3.18 y 3.20) y Py2DCoS (Figuras 3.22 - 3.25).

Tabla 3.5

Resultados de análisis PCA de sistema unitario, binario y ternario de BSA con PEG

UNITARIO			BINARIO			TERNARIO	
Tiempo			Tiempo			Tiempo	
mg/ml	PC1		mg/ml	PC1		mg/ml	PC1
5	-0,051		5	-0,006		5	0,14
6	-0,054		10	-0,030		10	0,19
7	0,52		15	0,39		15	0,25
8	0,46		20	0,37		20	0,30
9	0,53		25	0,39		25	0,35
10	0,49		30	0,44		30	0,40
			35	0,43		35	0,48
	-		40	0,43		40	0,51
Concentración			Concentración			Concentración	
d	PC1	PC2	d	PC1	PC2	d	PC1
1	0,23	0,66	1	-0,15	0,68	1	0,56
2	-0,044	0,74	2	0,99	0,11	2	0,56
3	0,97	-0,12	3	-0,0014	0,72	3	0,60

Tabla 3.6

Resultados de las frecuencias con mayor carga en PCA para sistemas binario y ternario en BSA

Binario			
BSA-Glucosa		BSA -Metformina	
Frecuencias (cm ⁻¹)	Cargas	Frecuencias (cm ⁻¹)	Cargas
1650	20	1563-1569	62
1648	18	1708-1800	45
1629	16	1644	32
Ternario			
Frecuencias (cm ⁻¹)		Cargas	
1563-1569		64	
1648		22	
1643		19	

El congestionamiento molecular influye significativamente en las propiedades dinámicas y estructurales de las proteínas. En sistemas con PEG-8 como agente congestionante, se observa una mayor complejidad en los datos obtenidos mediante PCA, lo que sugiere que la interacción de la proteína en un medio congestionado puede derivar en una reducción del movimiento libre de la proteína, impidiendo o disminuyendo la formación de complejos proteicos. Este comportamiento ha sido respaldado por estudios teóricos, los cuales demuestran que el congestionamiento macromolecular y el confinamiento espacial afectan los equilibrios y las tasas de plegamiento y unión de proteínas (Zhou, 2004).

La principal diferencia con el estudio previo de la BSA sin medio congestionado es la frecuencia que indica glicación reportada por De la Paz (2023), que fue en 1620 cm⁻¹ mientras que como se muestra en la Tabla 3.6 las variaciones destacadas coinciden en las

frecuencias ubicados en las estructuras α -hélices y láminas β (Barth, 2007) para el sistema BSA-Glucosa, lo que podría estar relacionado al uso del medio congestionante ya que facilita que las moléculas se unen entre sí y, por lo tanto, influir indirectamente en los resultados espectroscópicos, mejorando la observación de las bandas (Y. Zhang et al., 2023). En la investigación previa con la BSA no se realizó PCA, lo que dificultaría en el análisis de las transformaciones de la proteína ocasionadas por los ligandos con variaciones de tiempo y concentración. Mientras que, integrando los espectros mostrados en las figuras 3.26 a 3.29, el PCA (figuras 3.30-3.33) y los 2DCoS (figuras 3.34-3.37)

3.4 Desarrollo de herramienta de análisis intuitivo con PCA para sistemas complejos

Debido la facilidad de análisis proporcionada por el PCA realizado con la data existente en comparación con la investigación previa realizada por De la Paz (2023), y a causa de que las aplicaciones para realizar este análisis estadístico se encuentran disponibles únicamente bajo pago, se desarrolló una aplicación web con interfaz intuitiva que permite al analista subir archivos en formato .xlsx con las características iguales a las trabajadas en el análisis 2DCoS con resultados inmediatos que incluyen diagramas de dispersión, tabla de resultados con detalle de los PC y los pesos que representan las varianzas. Además, con la finalidad de obtener una mejor integración entre el análisis y la complejidad de los sistemas, entre los resultados se agregó el detalle de las frecuencias que representan la mayor variabilidad con sus respectivas cargas, acompañados de los posibles grupos funcionales que esas frecuencias o rango de frecuencias representan químicamente, en la tabla 3.7 se realiza una comparación con la plataforma Statgraphics.

Tabla 3.7

Comparación con otra herramienta de análisis PCA

Característica	Software desarrollado	Statgraphics XIX
Permite archivos .xlsx	✓	✗

Genera gráficos para los componentes principales	✓	✓
Muestra los pesos de las muestras por cada componente principal	✓	✓
Muestra la varianza acumulada de los componentes principales	✓	✓
Muestra la frecuencia de onda en los gráficos de dispersión	✓	✗
Permite descargar las tablas de resultados	✓	✗
Indica los compuestos de las frecuencias de onda más relevantes	✓	✗
Libre acceso	✓	✗

Figura 3.26.

Sistema unitario BSA 10 mg/ml

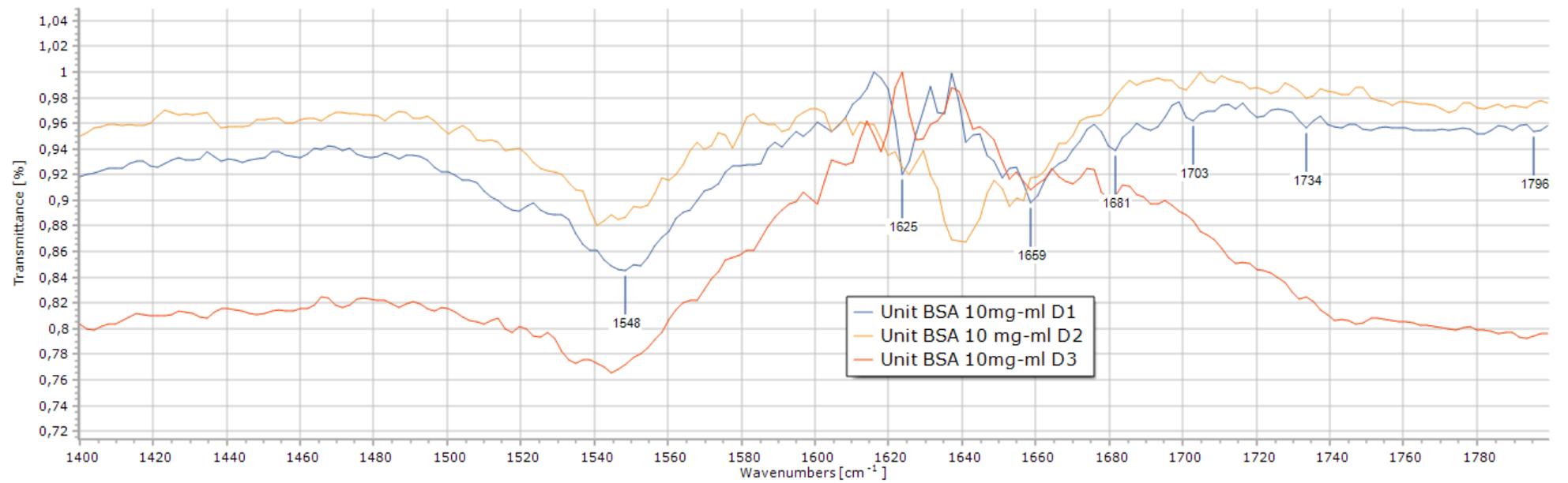


Figura 3.27

Sistema binario BSA 10 mg/ml y Metformina 40 mg/ml

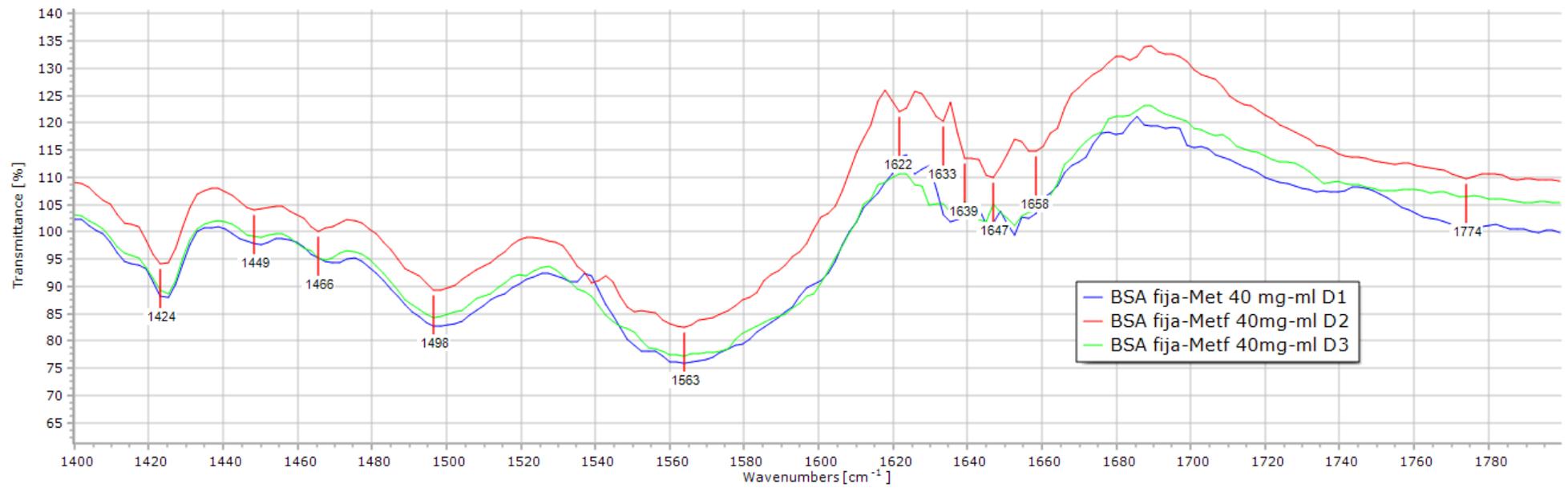


Figura 3.28.

Sistema binario BSA 10 mg/ml y Glucosa 15 mg/ml

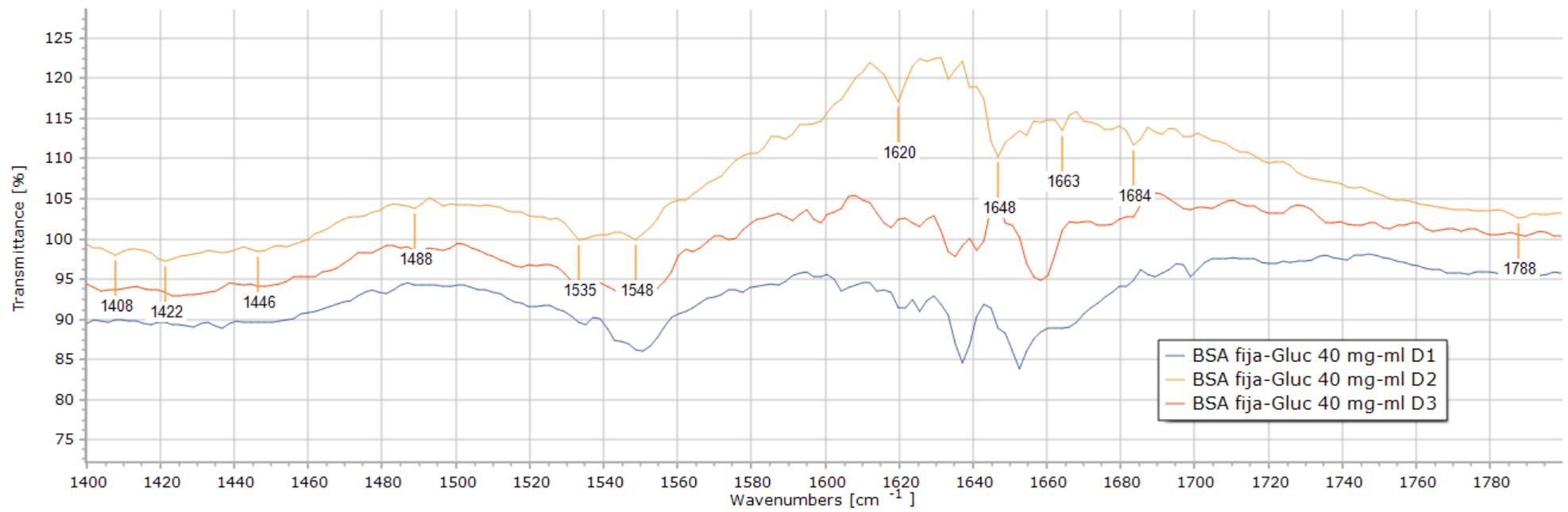


Figura 3.29

Sistema ternario BSA 10 mg/ml con Glucosa 15 mg/ml y Metformina 40 mg/ml

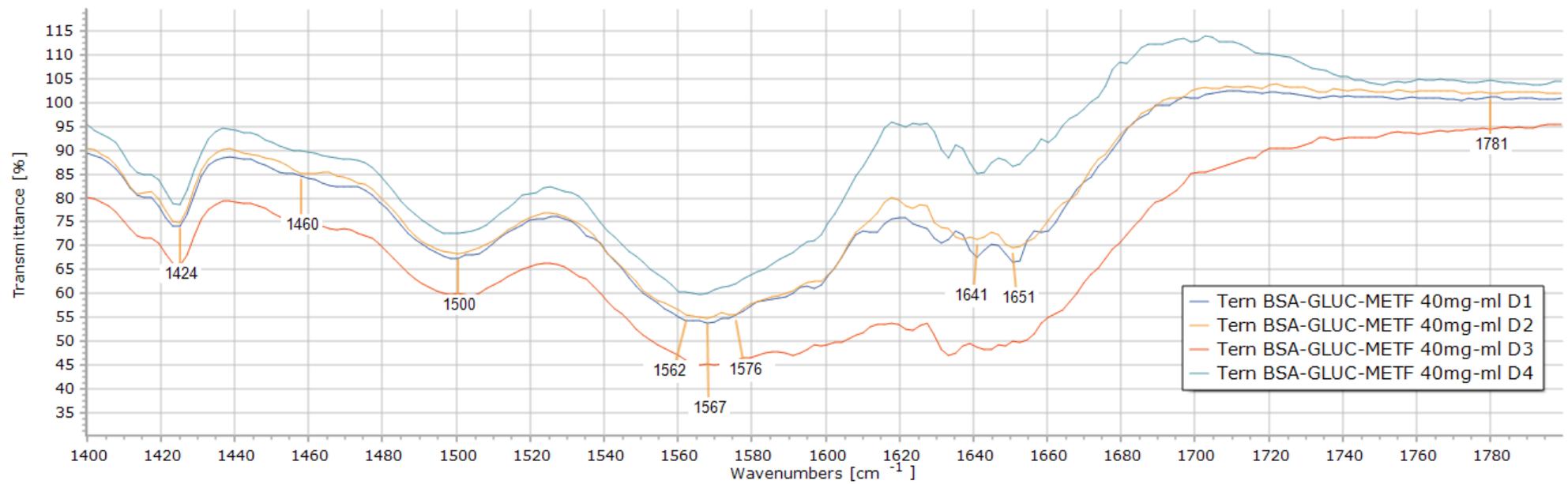


Figura 3.30

PCA de sistema unitario de BSA 10 mg/ml con PEG 8

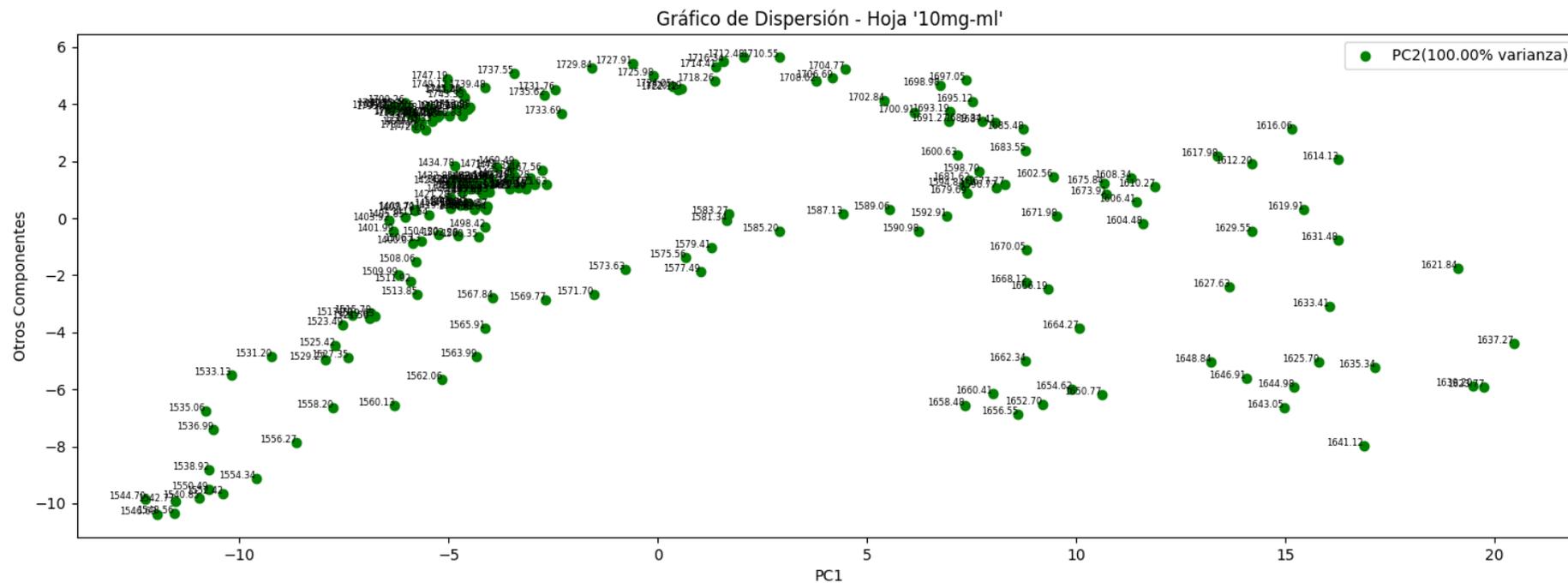


Figura 3.31

PCA de sistema binario con BSA 10 mg/ml y glucosa 15 mg/ml con PEG 8

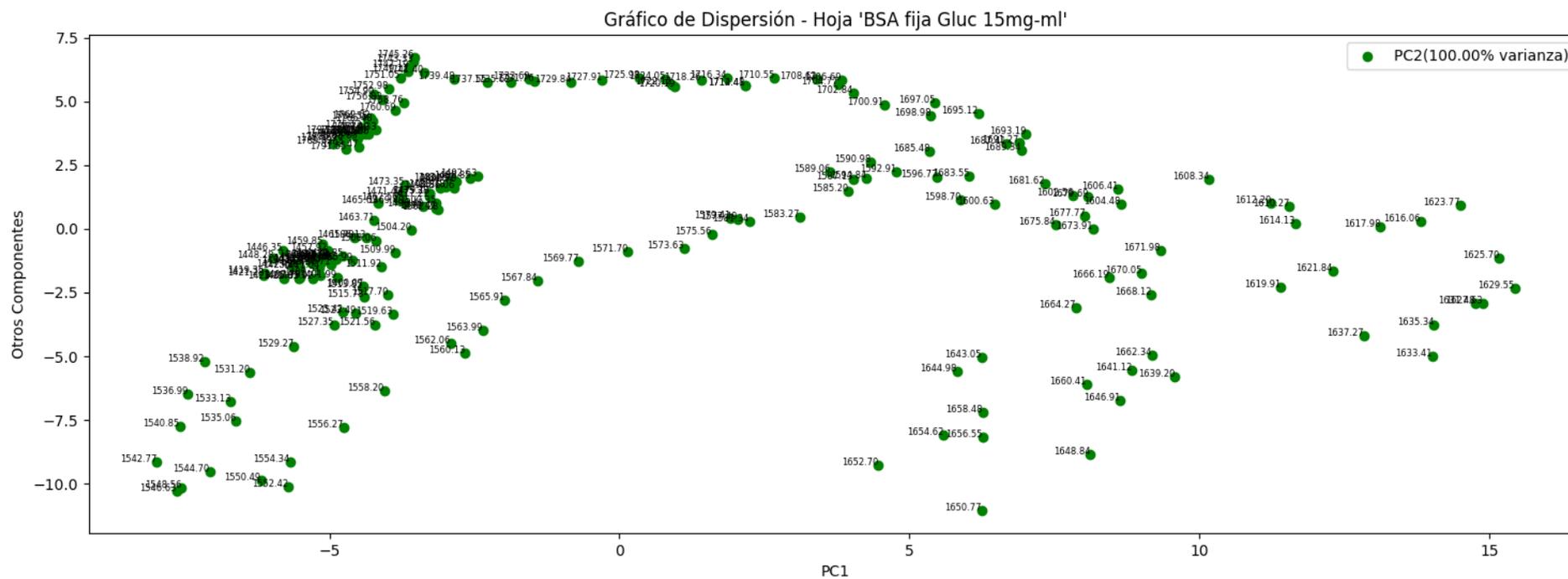


Figura 3.32

PCA de sistema binario de BSA 10 mg/ml y metformina 40 mg/ml con PEG 8

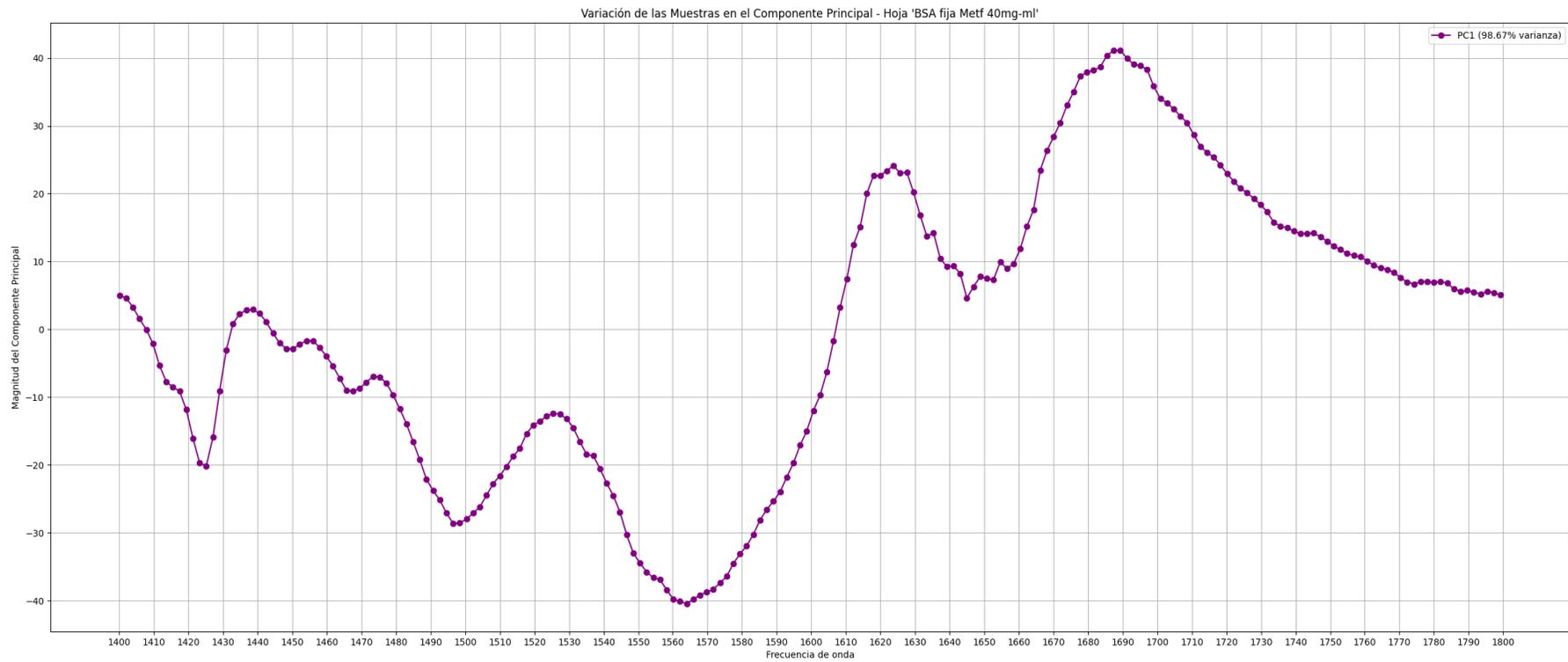


Figura 3.33

PCA de sistema ternario BSA 10 mg/ml, glucosa 15 mg/ml y metformina 40 mg/ml con PEG 8

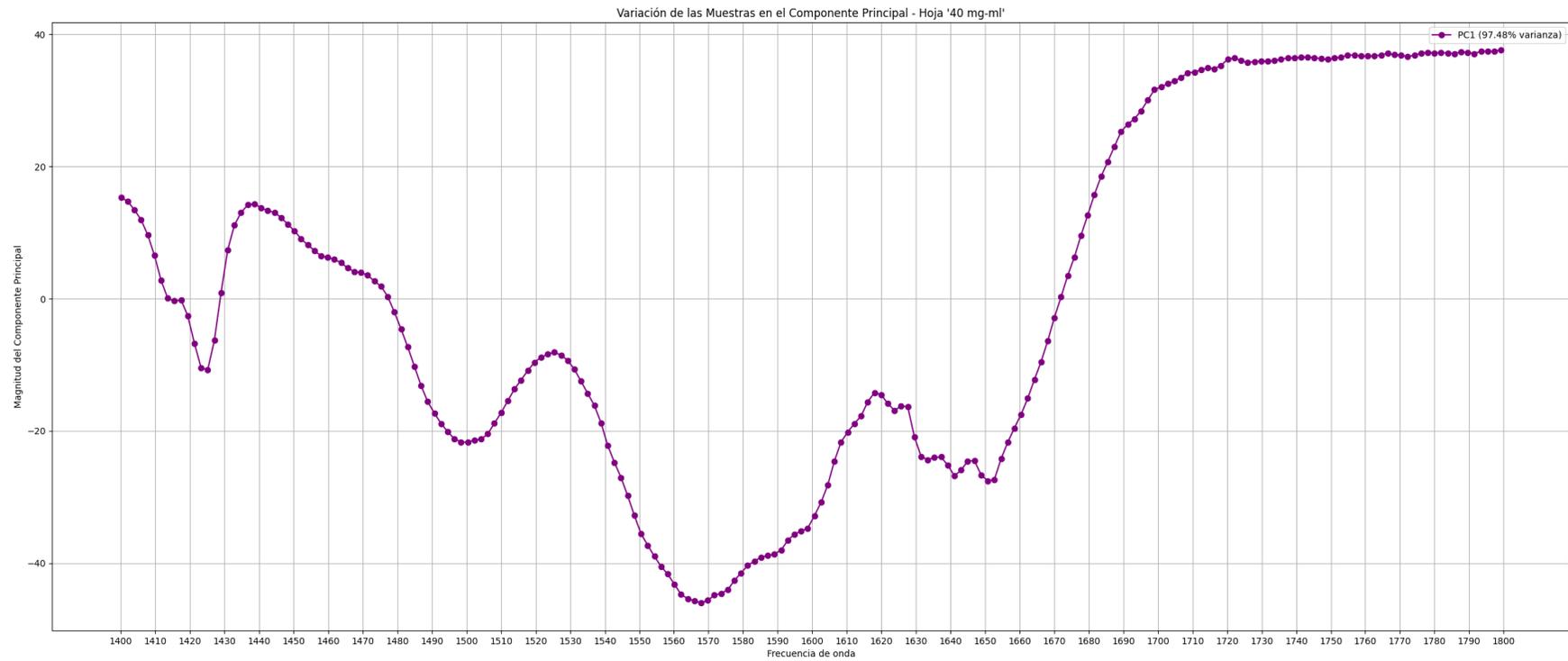


Figura 3.34

2DCOS PEG BSA 10 mg/ml por días

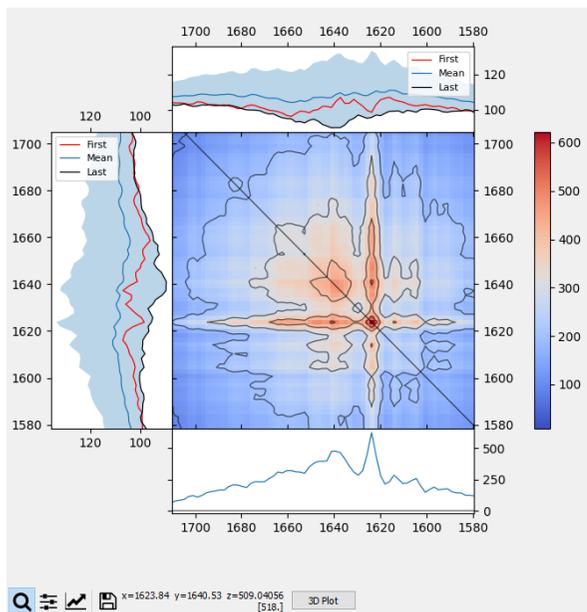


Figura 3.35

2DCOS PEG BSA 10 mg/ml y glucosa 15 mg/ml por días

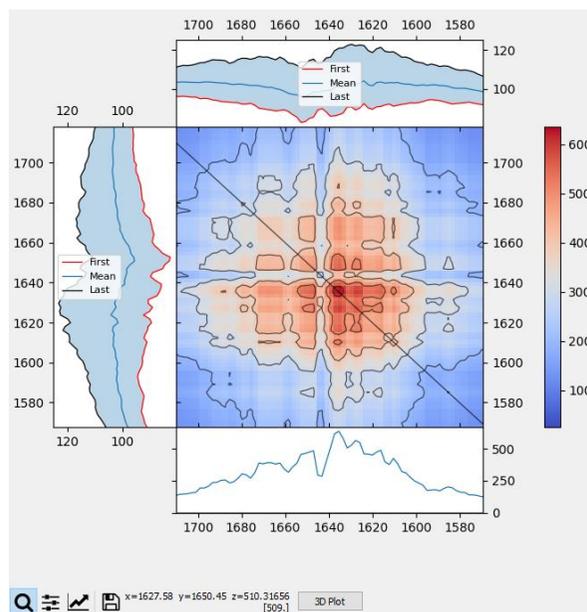


Figura 3.36

2DCOS PEG BSA 10 mg/ml y metformina 40 mg/ml por días

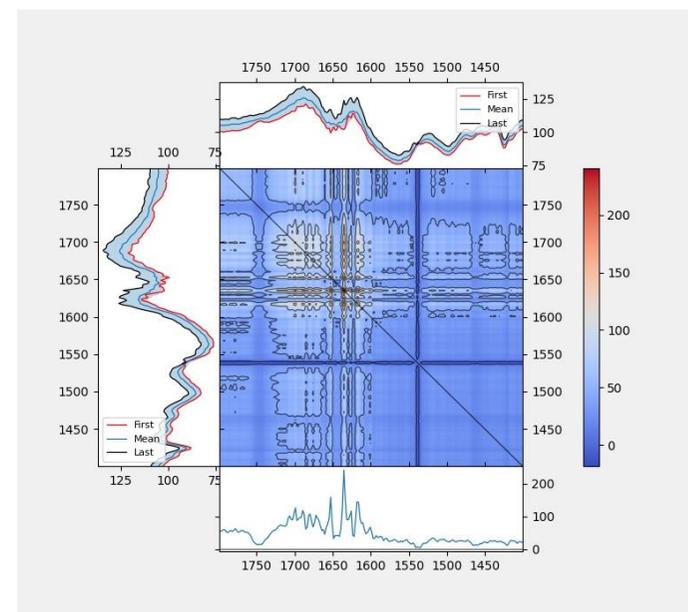
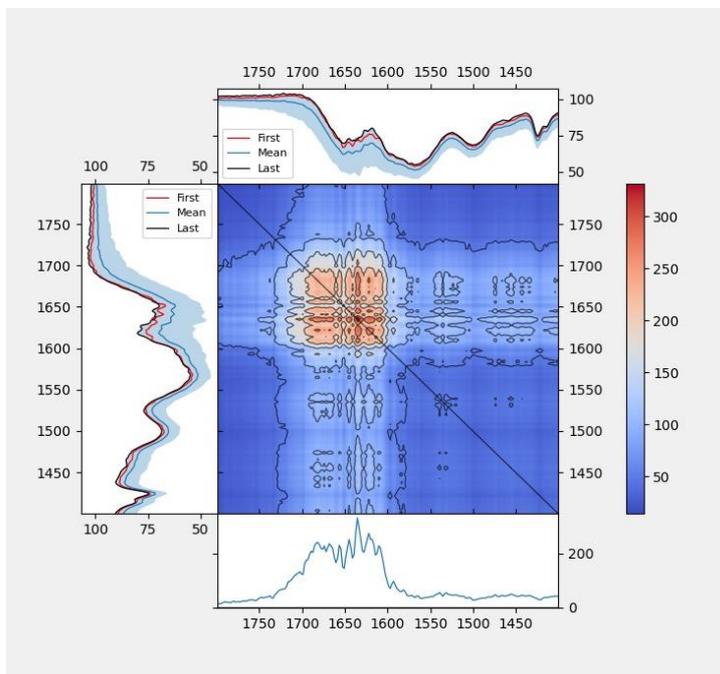


Figura 3.37

2DCOS PEG BSA 10 mg/ml, glucosa 15 mg/ml y metformina 40 mg/ml por días



3.5 Análisis de costos

El desarrollo del proyecto no generó costos adicionales, ya que todos los materiales, equipos y reactivos necesarios estaban disponibles en el laboratorio de la universidad. Los insumos más costosos, como las proteínas BSA y BHp, fueron adquiridos previamente por el departamento de investigación dentro de su presupuesto anual, garantizando acceso a materiales de alta calidad sin requerir inversiones externas.

Además, se desarrolló una aplicación web específica para realizar análisis PCA utilizando herramientas de Python de libre acceso, eliminando la necesidad de aplicaciones comerciales de pago. Esto optimizó los recursos tecnológicos disponibles, permitiendo una solución económica y eficiente para el procesamiento e interpretación de datos espectroscópicos.

Capítulo 4

4. Conclusiones y recomendaciones

4.1 Conclusiones

- Se desarrolló una estrategia eficaz para identificar las frecuencias de onda más relevantes en cada sistema estudiado, facilitando una interpretación más rápida y precisa de los datos obtenidos mediante PCA. Esta estrategia integró diversas etapas clave, como la selección y adaptación del formato de los datos FTIR, la delimitación de rangos de frecuencia relevantes, y la determinación de las perturbaciones más significativas a partir de los pesos de PCA. Además, se incorporó un enfoque iterativo para verificar la correspondencia entre las frecuencias identificadas y los valores reportados en la literatura científica, lo que garantizó la validez de los resultados. Finalmente, la aplicación del análisis 2DCOS permitió establecer correlaciones entre las frecuencias perturbadas, mejorando la comprensión de los fenómenos moleculares subyacentes y optimizando el análisis espectral en sistemas con y sin agentes de congestiónamiento molecular. Esta estrategia no solo permitió confirmar estos hallazgos, sino que también optimizó el tiempo de análisis al centrarse únicamente en las perturbaciones más significativas, proporcionando un enfoque más eficiente y dirigido para el estudio de sistemas biológicos complejos.
- Se estudiaron los sistemas proteína-ligando en condiciones donde ocurre la glucosilación no enzimática y el congestiónamiento molecular utilizando PEG-8 como agente congestionante, usando espectroscopía FTIR para las proteínas BSA y hemoglobina bovina. Los resultados mostraron que la glucosa causa cambios importantes en la estructura de las proteínas, promoviendo la glicación no enzimática. Sin embargo, cuando se añadió metformina, las proteínas mantuvieron su estructura original, lo que evitó la formación de productos avanzados de glicación. Este estudio permitió identificar claramente las diferencias estructurales causadas

por estas condiciones y resaltó el papel protector de la metformina en estos procesos.

- La caracterización de los sistemas con y sin congestiónamiento molecular, empleando Py2DCos y métodos estadísticos avanzados, permitió establecer criterios robustos para la interpretación de señales espectrales. En particular, se observó que en los sistemas con PEG-8000, las proteínas presentaron un comportamiento diferente frente a la unión de ligandos como metformina y glucosa. El análisis combinado de FTIR, 2DCOS y PCA reveló que el PEG, al actuar como un agente de congestiónamiento molecular, reduce el espacio disponible alrededor de la proteína debido a su gran tamaño molecular. Esta reducción de espacio genera un entorno más restringido que dificulta la interacción de los ligandos con la proteína, limitando su acceso a sitios de unión. Como resultado, se identificó que el sistema con PEG ejerce una propiedad protectora para la proteína, al impedir de manera más efectiva la interacción con ligandos en comparación con el sistema sin congestiónamiento molecular.
- Se diseñó una herramienta eficaz que combina el análisis estadístico y bases de interpretación de señales FTIR mediante el uso de librerías de acceso gratuito que permite la implementación de la técnica PCA y brinda reportes completos y detallados que, en conjunto con el análisis Py2DCos, optimiza la interpretación de los espectros obtenidos en FTIR clásico. La herramienta desarrollada permite al analista cargar un archivo .xlsx, seleccionar las hojas que se desea analizar y obtener resultados instantáneos y a su vez tener el acceso a descargar archivos en formato .xlsx para las tablas y .pdf o .png para gráficos de dispersión.

4.2 Recomendaciones

- Se sugiere trabajar en la integración del código de la aplicación web para crear un programa único que permita utilizar los datos del PCA y el análisis Py2DCos en una

misma interfaz. Esto facilitará un análisis más dinámico e interactivo para el usuario, permitiendo visualizar los resultados de ambas herramientas de forma simultánea.

- Se recomienda ampliar las condiciones de perturbación del sistema, no solo variando la concentración y el tiempo, sino también incluyendo la temperatura y pH. Esto permitirá simular los cambios térmicos que experimenta el cuerpo humano y evaluar con mayor precisión la capacidad de la metformina para proteger las proteínas frente a la glicación no enzimática.
- Aprovechando la agilidad obtenida para el análisis de estos sistemas complejos se debe realizar los mismos estudios usando sangre humana para incluir las diferentes proteínas y demás sustancias presentes en este tipo de muestra y así obtener resultados que destaquen las diferencias con los obtenidos en esta investigación.

Referencias

- Alhazmi, H. A. (2019). FT-IR spectroscopy for the identification of binding sites and measurements of the binding interactions of important metal ions with bovine serum albumin. *Scientia Pharmaceutica*, 87(1). <https://doi.org/10.3390/scipharm87010005>
- Andraka, N., Gonzalez-Velasco, J., Celeiro, J., Arrondo, J. L. R., & Bilbao, P. (2012). An infrared microspectroscopy 2DCOS study of the effect of radiation on normal and cancer cells. *Vibrational Spectroscopy*, 60, 189–192. <https://doi.org/10.1016/j.vibspec.2011.11.008>
- Barth, A. (2007). Infrared spectroscopy of proteins. In *Biochimica et Biophysica Acta - Bioenergetics* (Vol. 1767, Issue 9, pp. 1073–1101). <https://doi.org/10.1016/j.bbabi.2007.06.004>
- Brixner Riça, L., Sari Cassol, O., Rieger, A., & Corbellini, V. A. (2019). Discrimination of healthy and colorectal cancer patients using FTIR and PLS-DA. *Revista Jovens Pesquisadores*, 9(2), 115–130. <https://doi.org/10.17058/rjp.v9i2.13372>
- Bro, R., & Smilde, A. K. (2014). Principal component analysis. In *Analytical Methods* (Vol. 6, Issue 9, pp. 2812–2831). Royal Society of Chemistry. <https://doi.org/10.1039/c3ay41907j>
- Caccamo, M. T., & Magazù, S. (2022). Multiscale Spectral Analysis on Lysozyme Aqueous Solutions in the Presence of PolyEthyleneGlycol. *Molecules*, 27(24). <https://doi.org/10.3390/molecules27248760>
- Chen, R., & Spiro, T. G. (2002). Monitoring the allosteric transition and CO rebinding in hemoglobin with time-resolved FTIR spectroscopy. *Journal of Physical Chemistry A*, 106(14), 3413–3419. <https://doi.org/10.1021/jp012491v>

- De La Arada, I., Andraka, N., Pacios, M. G., Luís, J., & Arrondo, R. (2011). A Conventional and 2DCOS Infrared Approach to the Kinetics of Protein Misfolding. In *Current Protein and Peptide Science* (Vol. 12).
- De la Paz, J. (2023). *Desarrollo de una herramienta espectroscópica para el estudio de sistemas ligando-proteína usando 2D-COS FTIR*.
- Dong, J. E., Zhang, S., Li, T., & Wang, Y. Z. (2022). 2DCOS combined with CNN and blockchain to trace the species of boletes. *Microchemical Journal*, 177. <https://doi.org/10.1016/j.microc.2022.107260>
- Foist, R. B., Schulze, H. G., Ivanov, A., & Turner, R. F. B. (2011). Noise reduction methods applied to two-dimensional correlation spectroscopy (2D-COS) reveal complementary benefits of pre- and post-treatment. *Applied Spectroscopy*, 65(5), 474–487. <https://doi.org/10.1366/10-06107>
- Gholami, A., & Shirmohammadi, S. (2024). Advanced in Protein-Ligand Interaction Studies Optical-Based Spectroscopic Insights and Simulation Perspectives. *Journal of Biomedical Research & Environmental Sciences*, 5(6), 614–629. <https://doi.org/10.37871/jbres1934>
- Hniopek, J., Schmitt, M., Popp, J., & Bocklitz, T. (2020). PC 2D-COS: A Principal Component Base Approach to Two-Dimensional Correlation Spectroscopy. *Applied Spectroscopy*, 74(4), 460–472. <https://doi.org/10.1177/0003702819891194>
- Ibrahim, M., Ramadan, E., Elsadek, N. E., Emam, S. E., Shimizu, T., Ando, H., Ishima, Y., Helmy, O., Sarhan, H. A., Hussein, A. K., & Ishida, T. (2022). *Polyethylene glycol (PEG): The nature, immunogenicity, and role in the hypersensitivity of PEGylated products 2 3*.
- Ioannou, A., & Varotsis, C. (2019). Reaction of Hemoglobin With the Schiff Base Intermediate of the Glucose/Asparagine Reaction: Formation of a Hemichrome. In *Polyphenols in Plants* (pp. 317–325). Elsevier. <https://doi.org/10.1016/b978-0-12-813768-0.00017-7>

- Jankovic, A., Chaudhary, G., & Goia, F. (2021). Designing the design of experiments (DOE) – An investigation on the influence of different factorial designs on the characterization of complex systems. *Energy and Buildings*, 250. <https://doi.org/10.1016/j.enbuild.2021.111298>
- Jiang, E. Y., & Rieppo, J. (2006). Enhancing FTIR imaging capabilities with two-dimensional correlation spectroscopy (2DCOS): A study of concentration gradients of collagen and proteoglycans in human patellar cartilage. *Journal of Molecular Structure*, 799(1–3), 196–203. <https://doi.org/10.1016/j.molstruc.2006.03.069>
- Jung, Y. M. (2003). Principal Component Analysis Based Two-Dimensional (PCA-2D) Correlation Spectroscopy: PCA Denoising for 2D Correlation Spectroscopy. In *Correlation Spectroscopy Bull. Korean Chem. Soc* (Vol. 24, Issue 9).
- Jung, Y. M., Czarnik-Matusiewicz, B., Wu, Y., Zhang, L., & Noda, I. (2015). Application of Two-dimensional Correlation Spectroscopy in Protein Research. In *Encyclopedia of Analytical Chemistry* (pp. 1–27). Wiley. <https://doi.org/10.1002/9780470027318.a9513>
- Khan, M. W. A., Otaibi, A. Al, Al-Zahrani, S. A., Alshammari, E. M., Haque, A., Alouffi, S., Khan, W. A., & Khan, S. N. (2021). Experimental and theoretical insight into resistance to glycation of bovine serum albumin. *Journal of Molecular Structure*, 1230. <https://doi.org/10.1016/j.molstruc.2020.129645>
- Larios, G., Ribeiro, M., Arruda, C., Oliveira, S. L., Canassa, T., Baker, M. J., Marangoni, B., Ramos, C., & Cena, C. (2021). A new strategy for canine visceral leishmaniasis diagnosis based on FTIR spectroscopy and machine learning. *Journal of Biophotonics*, 14(11). <https://doi.org/10.1002/jbio.202100141>

- Li, X., Li, H., Lai, K., & Miao, J. (2023). The Effect of Glucose on the Interaction of Bisphenol A and Bovine Hemoglobin Characterized by Spectroscopic and Molecular Docking Techniques. *International Journal of Molecular Sciences*, 24(19).
<https://doi.org/10.3390/ijms241914708>
- Noda, I. (2016). Quadrature two-dimensional correlation spectroscopy (Q-2DCOS). *Journal of Molecular Structure*, 1124, 42–52.
<https://doi.org/10.1016/j.molstruc.2016.01.090>
- Olabinjo, Oyebola, & Oduyano. (2024). *Response Surface Techniques as an Inevitable Tool in Optimization Process*. www.intechopen.com
- Outman, A., Deracinois, B., Flahaut, C., Diab, M. A., Dhaouefi, J., Gressier, B., Eto, B., & Nedjar, N. (2023). Comparison of the Bioactive Properties of Human and Bovine Hemoglobin Hydrolysates Obtained by Enzymatic Hydrolysis: Antimicrobial and Antioxidant Potential of the Active Peptide α 137-141. *International Journal of Molecular Sciences*, 24(17).
<https://doi.org/10.3390/ijms241713055>
- Pataridis, S., Štastná, Z., Sedláková, P., & Mikšík, I. (2013). Monotopic modifications derived from in vitro glycation of albumin with ribose. *Electrophoresis*, 34(12), 1757–1763.
<https://doi.org/10.1002/elps.201300014>
- Rahbar, S., & Figarola, J. L. (2003). Novel inhibitors of advanced glycation endproducts. In *Archives of Biochemistry and Biophysics* (Vol. 419, Issue 1, pp. 63–79). Academic Press Inc.
<https://doi.org/10.1016/j.abb.2003.08.009>
- Sánchez-Brito, M., Luna-Rosas, F. J., Mendoza-González, R., Mata-Miranda, M. M., Martínez-Romo, J. C., & Vázquez-Zapién, G. J. (2021). A machine-learning strategy to evaluate the use of FTIR spectra of saliva for a good

control of type 2 diabetes. *Talanta*, 221.

<https://doi.org/10.1016/j.talanta.2020.121650>

Seichter, F., Vogt, J., Radermacher, P., & Mizaikoff, B. (2017). *Response-surface fits and calibration transfer for the correction of the oxygen effect in the quantification of carbon dioxide via FTIR spectroscopy.*

Venkatesan, P., Dharuman, C., & Gunasekaran, S. (2011). A comparative study of principal component regression and partial least squares regression with application to FTIR diabetes data. *Journal of Science and Technology*, 4.

Vijayakumar, S., Schwaighofer, A., Ramer, G., & Lendl, B. (2024). Multivariate curve resolution -alternating least squares augmented with partial least squares baseline correction applied to mid-IR laser spectra resolves protein denaturation by reducing rotational ambiguity. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, 315. <https://doi.org/10.1016/j.saa.2024.124228>

Wander L, Vianello A, Vollertsen J, Braun U, & Paul A. (2019). Multivariate Analysis of Large μ -FTIR Data Sets in Search of Microplastics. *16th International Conference on Environmental Science and Technology*, 4.

Wang, Y., & Annunziata, O. (2007). Comparison between Protein-polyethylene Glycol (PEG) interactions and the effect of PEG on protein-protein interactions using the liquid-liquid phase transition. *Journal of Physical Chemistry B*, 111(5), 1222–1230. <https://doi.org/10.1021/jp065608u>

Younus, H., & Anwar, S. (2016). Prevention of non-enzymatic glycosylation (glycation): Implication in the treatment of diabetic complication. In *International Journal of Health Sciences* (Vol. 10, Issue 2).

Yue, J. Q., Huang, H. Y., & Wang, Y. Z. (2021). A practical method superior to traditional spectral identification: Two-dimensional correlation spectroscopy combined with deep learning to identify Paris species.

<https://doi.org/10.1016/j.microc.2020.105731>

Zhang, L., Zhang, L., Yan, L., Liu, B., Wang, X., & Wang, J. (2005). *Application and improvement of partial-least-squares in Fourier transform infrared spectroscopy.*

Zhang, Q., Ames, J. M., Smith, R. D., Baynes, J. W., & Metz, T. O. (2009). A Perspective on the Maillard Reaction and the Analysis of Protein Glycation by Mass Spectrometry: Probing the Pathogenesis of Chronic Disease. *Journal of Proteome Research*, 8(2), 754–769. <https://doi.org/10.1021/pr800858h>

Zhang, Y., Zhang, C., Wang, M., Du, K., & Peng, X. (2023). Effect of macromolecular crowding on the inhibition of non-enzymatic glycation of bovine serum albumin by cichoric acid. *Food Bioscience*, 53. <https://doi.org/10.1016/j.fbio.2023.102801>

Zhou, H. X. (2004). Protein folding and binding in confined spaces and in crowded solutions. *Journal of Molecular Recognition*, 17(5), 368–375. <https://doi.org/10.1002/jmr.711>

Anexos

Figura A

Validación con Statgraphics XIX para el PCA. a) Tablas de pesos obtenida con Python. b)

Tabla de pesos obtenidos con Statgraphics XIX

a)		b)	
Tabla de pesos de los componentes		Tabla de Pesos de los Componentes	
Columnas	PC1		Componente
BSA fija-Gluc 5 mg-ml	0,34484		1
BSA fija-Gluc 10 mg-ml	0,35016	Col_1	0,344858
BSA fija-Gluc 15 mg-ml	0,35669	Col_2	0,350158
BSA fija-Gluc 20 mg-ml	0,35657	Col_3	0,35668
BSA fija-Gluc 25 mg-ml	0,36455	Col_4	0,356573
BSA fija-Gluc 30 mg-ml	0,35749	Col_5	0,364556
BSA fija-Gluc 35 mg-ml	0,35586	Col_6	0,357479
BSA fija-Gluc 40 mg-ml	0,34172	Col_7	0,35585
		Col_8	0,341726