

**ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL**

**Facultad de Ciencias Naturales y Matemáticas**

Desarrollo de una herramienta espectroscópica para el estudio de  
sistemas ligando-proteína usando 2D-COS FTIR

**PROYECTO INTEGRADOR**

Previo la obtención del Título de:

**Ingeniero Químico**

Presentado por:

Julio Carlos de la Paz Cruz

GUAYAQUIL - ECUADOR

Año: 2023

## **DEDICATORIA**

A mi hermano, la persona que más quiero, y a mis padres, las luces de mi vida.

## **AGRADECIMIENTOS**

Mi más sincero agradecimiento a mi tutor, Dr. Joan Vera, y a mi cotutor, Dr. Dany De Cecchis, por hacer posible este proyecto, y por sus invaluable enseñanzas.

Agradezco de igual manera a la profesora Lorena Quinchuela, por su apoyo, consejo y ayuda durante mi transcurso académico.

A Shaden Intriago Sandoval, su compañía y su cariño hacen todo más sencillo.

## DECLARACIÓN EXPRESA

“Los derechos de titularidad y explotación, me corresponde conforme al reglamento de propiedad intelectual de la institución; *Julio Carlos de la Paz Cruz* y doy mi consentimiento para que la ESPOL realice la comunicación pública de la obra por cualquier medio con el fin de promover la consulta, difusión y uso público de la producción intelectual”



---

Julio Carlos de la Paz Cruz

## **EVALUADORES**

**Dr. Joan Ramón Vera Villalobos**

PROFESOR TUTOR

**M.Sc. María Verónica Ordóñez Pazmiño**

PROFESOR DE LA MATERIA

## RESUMEN

Para el estudio de proteínas en condiciones patológicas se requieren tecnologías avanzadas y costosas como la Resonancia Magnética Nuclear. No obstante, las instituciones de investigación con presupuestos limitados no cuentan con los recursos para adquirir estos dispositivos. Los equipos convencionales, como espectrómetros IR, no proporcionan el detalle necesario para un análisis profundo. En este contexto, se propone la correlación espectroscópica bidimensional como un método alternativo. Este método se basa en aplicar una perturbación medible (ej. química, térmica, etc.) a un sistema de interés, y tomar mediciones espectroscópicas lo largo de la perturbación. Como resultado se tiene un espectro dinámico, a partir del cual se generan espectros de correlación que permiten extraer información relevante sobre el sistema. En el presente trabajo se desarrolló un software computacional utilizando lenguaje Python, que genera gráficos de correlación bidimensional para el análisis de proteínas y sistemas complejos. Se estableció también un caso de estudio sobre la interacción entre Albúmina Sérica Bovina, Glucosa y Metformina. La herramienta desarrollada permitió monitorear cambios estructurales en la proteína y verificar el efecto de la metformina al retrasar el proceso de glucosilación.

**Palabras Clave:** proteínas, espectroscopía, correlación espectroscópica bidimensional, glucosilación.

## **ABSTRACT**

*The study of proteins in pathological conditions requires advanced and expensive technologies such as Nuclear Magnetic Resonance. However, research institutions with limited budgets do not have the resources to acquire these devices. Conventional equipment, such as IR spectrometers, do not provide the detail necessary for in-depth analysis. In this context, Two-Dimensional Correlation Spectroscopy is proposed as an alternative method. This method is based on applying a measurable perturbation (e.g., chemical, thermal, etc.) to a system of interest, and taking spectroscopic measurements along the perturbation. The result is a dynamic spectrum, from which correlation spectra are generated to extract relevant information about the system. In the present work a computational software is developed that generates two-dimensional correlation graphs for the analysis of proteins and complex systems. A case study on the interaction between Bovine Serum Albumin, Glucose and Metformin is also established. The developed tool allowed monitoring structural changes in the protein and verifying the effect of metformin in delaying the glycosylation process.*

*Keywords: proteins, spectroscopy, two-dimensional spectroscopic correlation, glycosylation.*

# ÍNDICE GENERAL

EVALUADORES.....	5
RESUMEN .....	I
<i>ABSTRACT</i> .....	II
ÍNDICE GENERAL .....	III
ABREVIATURAS.....	V
ÍNDICE DE FIGURAS .....	VI
ÍNDICE DE TABLAS.....	VIII
CAPÍTULO 1 .....	1
1.    Introducción .....	1
1.1    Descripción del problema .....	1
1.2    Justificación del problema .....	2
1.3    Objetivos .....	3
1.3.1    Objetivo General .....	3
1.3.2    Objetivos Específicos .....	3
1.4    Marco teórico.....	4
1.4.1    Estructura y función de las proteínas .....	4
1.4.2    Unión de proteínas y ligandos .....	4
1.4.3    Análisis de proteínas por espectroscopía infrarroja .....	5
1.4.4    Correlación espectroscópica bidimensional .....	8
CAPÍTULO 2 .....	11
2.    Metodología .....	11
2.1    Desarrollo del software para 2DCOS .....	11
2.1.1    Librerías y dependencias utilizadas .....	12
2.1.2    Desarrollo de la interfaz gráfica .....	12

2.1.3	Obtención de los espectros de 2DCOS.....	14
2.2	Caso de estudio: interacción de Albúmina Sérica Bovina con Glucosa y Metformina .....	17
2.2.1	Sistemas de control.....	17
2.2.2	Variables de perturbación .....	18
2.2.3	Materiales .....	18
2.2.4	Preparación de muestras y medición .....	18
2.2.5	Sistemas perturbados por concentración .....	19
2.2.6	Sistemas perturbados por tiempo.....	20
2.2.7	Almacenamiento de los datos .....	20
2.2.8	Costos asociados.....	20
CAPÍTULO 3 .....		22
3.	Resultados Y ANÁLISIS .....	22
3.1	Desarrollo de Software Computacional.....	22
3.2	Caso de Estudio: Interacción de BSA con Glucosa y Metformina .....	25
3.2.1	Análisis de Espectros IR .....	25
3.2.2	Resultados de 2DCOS en Sistemas Perturbados por Tiempo .....	32
3.2.3	Resultados de 2DCOS en Sistemas Perturbados por Concentración ...	38
CAPÍTULO 4 .....		43
4.	Conclusiones Y Recomendaciones.....	43
	Conclusiones.....	43
	Recomendaciones.....	44
BIBLIOGRAFÍA .....		45

## **ABREVIATURAS**

ESPOL	Escuela Superior Politécnica del Litoral
FTIR	Espectroscopía Infrarroja por Transformada de Fourier
BSA	Albúmina Sérica Bovina
2DCOS	Correlación Espectroscópica Bidimensional
AGEs	Productos Finales de Glicosilación Avanzada

## ÍNDICE DE FIGURAS

Figura 1-1 Formación de AGEs. Elaborado en ChemSketch .....	5
Figura 1-2 Diagrama de Interferómetro [(Wade & Simek, 2017)] .....	7
Figura 1-3 Ejemplos de espectros síncrono y asíncrono (Xue et al., 2017) .....	10
Figura 2-1 Desarrollo del programa .....	12
Figura 2-2 Diagrama de Flujo de Usuario .....	13
Figura 2-3 Diseño inicial de interfaz.....	14
Figura 2-4 Modelo de almacenamiento de datos en Excel.....	15
Figura 2-5 Obtención de espectro dinámico en Python .....	16
Figura 2-6 Cálculo de espectro síncrono en Python .....	16
Figura 2-7 Obtención de matriz Noda en Python .....	16
Figura 2-8 Cálculo de espectro asíncrono en Python .....	17
Figura 3-1 Captura de Pantalla del Software Desarrollado .....	22
Figura 3-2 Pestaña Para Ingreso de Datos en Excel .....	23
Figura 3-3 Gráficos tridimensionales para los espectros síncrono y asíncrono.....	24
Figura 3-4 Espectro IR de BSA en Estado Basal.....	27
Figura 3-5 Espectro IR de BSA Durante Cinco Días.....	29
Figura 3-6 Espectro IR de Glucosa durante cinco días.....	30
Figura 3-7 Espectro IR de Metformina durante cinco días .....	31
Figura 3-8 2DCOS Metformina (20 mg/ml) por Días .....	33
Figura 3-9 2DCOS Glucosa (20 mg/ml) por Días.....	33
Figura 3-10 2DCOS BSA (10 mg/ml) por Días.....	33
Figura 3-11 2DCOS BSA (10 mg/ml) + Glucosa (40 mg/ml) por Días.....	35
Figura 3-12 2DCOS BSA (10 mg/ml) + Glucosa (20 mg/ml) por Días.....	35
Figura 3-13 2DCOS BSA (10 mg/ml) + Metformina (40 mg/ml) por Días .....	36
Figura 3-14 2DCOS BSA (10 mg/ml) + Glucosa (20 mg/ml) + Metformina (40 mg/ml) por Días.....	37
Figura 3-15 2DCOS BSA (10 mg/ml) + Glucosa (20 mg/ml) + Metformina (20 mg/ml) por Días.....	37
Figura 3-16 2DCOS BSA (10 mg/ml) + Glucosa (20 mg/ml) + Metformina (40 mg/ml) por Días.....	37

Figura 3-17 2DCOS BSA (10 mg/ml) + Glucosa (20 mg/ml) + Metformina (20 mg/ml) por Días.....	37
Figura 3-18 2DCOS Metformina (0 – 40 mg/ml) por Concentración.....	40
Figura 3-19 2DCOS Glucosa (0 – 40 mg/ml) por Concentración .....	40
Figura 3-20 2DCOS BSA (0 – 10 mg/ml) por Concentración .....	40
Figura 3-21 2DCOS BSA (10 mg/ml) + Glucosa (0-40 mg/ml) por Concentración ....	41
Figura 3-22 2DCOS BSA (10 mg/ml) + Metformina (0-40 mg/ml) por Concentración	41
Figura 3-23 2DCOS BSA (10 mg/ml) + Glucosa (20 mg/ml) + Metformina (0-40 mg/ml) por Concentración .....	42
Figura 3-24 2DCOS BSA (10 mg/ml) + Glucosa (20 mg/ml) + Metformina (0-40 mg/ml) por Concentración Imagen Tridimensional.....	42

## ÍNDICE DE TABLAS

Tabla 1-1.1 Asignación de bandas de absorción IR en proteínas	¡Error! Marcador no definido.
Tabla 2-1 Sistemas perturbados por concentración.....	20
Tabla 2-2 Sistemas perturbados por tiempo .....	20
Tabla 3-1 Comparación Con Otras Herramientas De 2DCOS .....	25
Tabla 3-2 Picos Asignados a BSA en Estado Nativo .....	28

# CAPÍTULO 1

## 1. INTRODUCCIÓN

### 1.1 Descripción del problema

El estudio de las modificaciones estructurales o conformacionales de las proteínas en condiciones patológicas es de gran relevancia para el desarrollo de fármacos y estrategias terapéuticas (Tunyasuvunakool et al., 2021). Sin embargo, las instituciones académicas y de investigación con presupuestos limitados enfrentan dificultades para llevar a cabo investigaciones detalladas debido a la falta de acceso a tecnologías avanzadas (Buschiazzo, 2022).

La complejidad estructural de las proteínas y la escala de tiempo en la que ocurren las modificaciones dificultan su estudio. Además, los equipos convencionales, como los espectrofotómetros UV-Vis o de FTIR, no proporcionan la resolución y sensibilidad necesarias para capturar y analizar en detalle las modificaciones estructurales específicas (Varadi et al., 2023).

La falta de acceso a equipos de espectrofluorimetría o resonancia magnética nuclear limita aún más la capacidad de obtener información precisa. Estas tecnologías avanzadas son costosas y no están al alcance de todas las instituciones con presupuestos ajustados (Buschiazzo, 2022). Asimismo, la disponibilidad limitada de muestras biológicas adecuadas y la heterogeneidad de estas dificultan el análisis preciso de las modificaciones estructurales (Tunyasuvunakool et al., 2021).

La falta de recursos para crear un entorno propicio para el desarrollo de métodos no convencionales también afecta la capacidad de las instituciones para realizar estudios experimentales más detallados. Por último, la falta de experiencia y capacitación del personal en el uso de tecnologías avanzadas y métodos no convencionales, así como la escasez de personal especializado en el análisis de las modificaciones estructurales de las proteínas bajo condiciones patológicas, contribuyen a las dificultades en la realización de investigaciones detalladas (Buschiazzo, 2022).

## 1.2 Justificación del problema

Las proteínas son macromoléculas esenciales para la vida. Están involucradas en prácticamente todos los procesos biológicos, y, por tanto, su estudio es de vital importancia. Gracias a los avances en el conocimiento estructural y funcional de las proteínas se ha logrado comprender el mecanismo de numerosos procesos metabólicos, se han desarrollado vacunas y estrategias terapéuticas para combatir múltiples enfermedades, y se han conseguido avances en las industrias química y alimenticia (Nelson & Cox, 2004).

La estructura de las proteínas está estrechamente ligada a su función. Por ende, las modificaciones estructurales de las proteínas pueden alterar su papel en los procesos metabólicos. Una de las vías por las cuales estas sufren cambios estructurales es a través de su unión con ligandos. En condiciones extremas, un ligando puede provocar que una proteína pierda completamente su funcionalidad. Por ejemplo, la glucosa presente en el plasma sanguíneo puede actuar como agente reductor, reaccionar con un grupo amino libre de una proteína, y desencadenar una serie de reacciones cuyo resultado final podrían ser los conocidos productos finales de la glicación avanzada (AGEs). La producción de AGEs se acelera en personas que padecen de Diabetes Mellitus debido a los altos niveles de azúcar en sangre, y esto se ha asociado a la aparición de varias enfermedades neurodegenerativas (Carvajal, 2015). Por tanto, el estudio de la interacción proteína-ligando es muy relevante, sobre todo cuando estas se encuentran en condiciones patológicas.

No obstante, para llevar a cabo dicho estudio, se requiere utilizar técnicas avanzadas de caracterización que puedan proporcionar información detallada sobre las proteínas en su estado patológico. Los equipos de resonancia magnética nuclear (RMN) son una de las tecnologías más utilizadas y eficaces en este campo, pero su adquisición y el mantenimiento son costosos, lo que dificulta su acceso para muchas instituciones y laboratorios de investigación (Buschiazzo, 2022). El costo de un equipo de RMN es de alrededor de “varios cientos de miles de dólares” (Skoog et al., 2018).

Además, si bien existen técnicas convencionales de caracterización menos costosas, como espectrofotómetros UV-Vis (menos de \$20.000 USD (Skoog et al., 2018)) y espectroscopía infrarroja por transformada de Fourier (FTIR) (entre \$10.000 y \$35.000 USD (www.labx.com, 2023)) , estas técnicas ofrecen una visión limitada de las

propiedades estructurales y conformacionales de las proteínas, lo que limita la comprensión de los cambios específicos asociados a las enfermedades (Varadi et al., 2023). Por lo tanto, surge la necesidad de explorar y evaluar métodos alternativos que puedan superar estas limitaciones y brindar una mayor resolución y precisión en el análisis de proteínas en condiciones patológicas. En este contexto, se propone el desarrollo de un método de correlación espectroscópica bidimensional como una alternativa prometedora.

La correlación espectroscópica bidimensional (2DCOS) utiliza dos conjuntos de datos espectrales obtenidos a partir de equipos convencionales, como FTIR, y los correlaciona para producir un espectro bidimensional, que puede proporcionar más información sobre la dinámica y estructura de las proteínas que la observada en un espectro infrarrojo tradicional (Noda, 2022). De esta manera, se consigue un análisis efectivo y económico. En este trabajo, se desarrolla una aplicación basada en lenguaje Python que permite analizar datos espectroscópicos de sistemas complejos constituidos por proteínas y fármacos/ligandos a través de un estudio de 2DCOS. Esta herramienta se aplica en un caso de estudio que analiza las interacciones entre la albúmina sérica bovina (BSA), la glucosa, y la metformina. La hipótesis experimental es que con la aplicación de 2DCOS se podrá evidenciar que la glucosa induce cambios estructurales en la proteína, y que la metformina es capaz de inhibirlos y/o revertirlos. De esta forma, se valida que el uso de 2DCOS es adecuado para el estudio de las interacciones proteína-ligando.

### **1.3 Objetivos**

#### **1.3.1 Objetivo General**

Desarrollar una aplicación de software libre en lenguaje Python para el análisis estructural de sistemas complejos mediante el uso de correlación espectroscópica bidimensional.

#### **1.3.2 Objetivos Específicos**

1. Diseñar una interfaz gráfica de usuario interactiva acoplada a un modelo matemático, para la presentación de los resultados obtenidos.

2. Establecer un caso de estudio sobre la interacción de la albúmina sérica bovina, la glucosa y la metformina, para la validación de la herramienta computacional desarrollada.
3. Interpretar los resultados obtenidos con base a la teoría de interacción proteína-ligando.

## **1.4 Marco teórico**

### **1.4.1 Estructura y función de las proteínas**

Las proteínas son macromoléculas poliméricas compuestas por una o más cadenas de aminoácidos unidos por enlaces peptídicos. Todas las proteínas, desde las más básicas, hasta las más complejas, están formadas por 20 aminoácidos distintos. Su estructura comúnmente se divide en los siguientes niveles organizacionales (Nelson & Cox, 2004):

- Estructura primaria: determinada por la secuencia de aminoácidos en una cadena.
- Estructura secundaria: conformaciones locales en regiones una cadena, como  $\alpha$  hélices y hojas  $\beta$  plegadas.
- Estructura terciaria: disposición tridimensional de la proteína, y responsable principal de su interacción con ligandos.
- Estructura cuaternaria: se conforma por la interacción de varias cadenas peptídicas.

Bajo condiciones determinadas, una proteína tiende a adoptar conformaciones estructurales que disminuyen su energía libre de Gibbs. Cuando una proteína está en una conformación funcional y termodinámicamente estable, se dice que está en su estado nativo (Nelson & Cox, 2004).

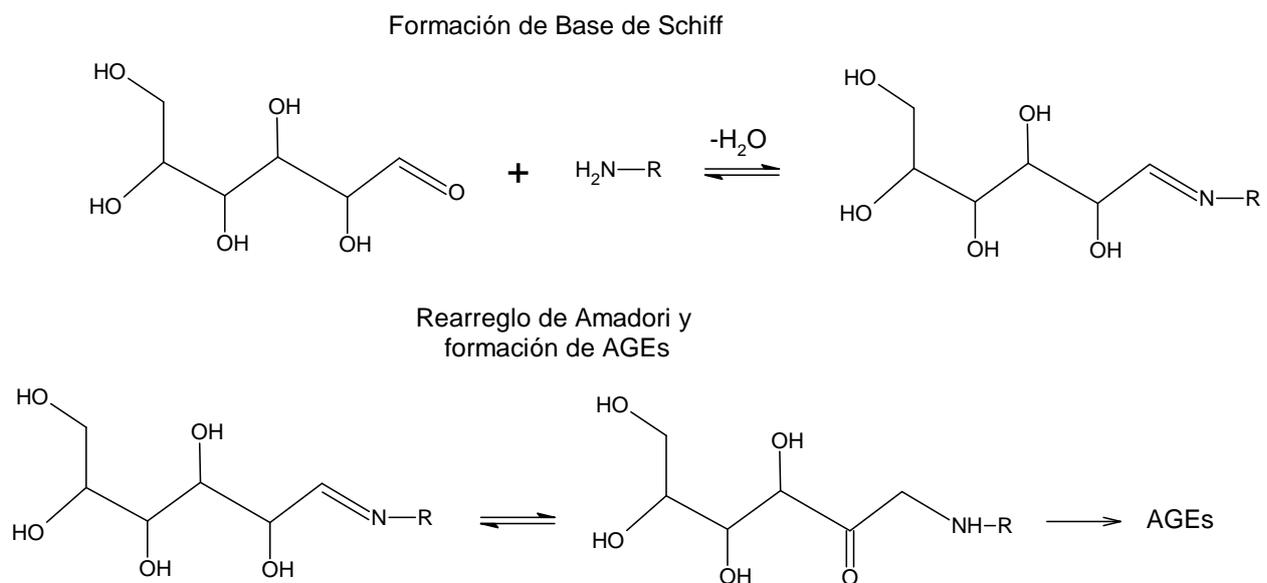
Como consecuencia de su complejidad estructural, las proteínas exhiben una enorme variedad de funciones biológicas, por ejemplo: función hormonal, enzimática, estructural, mecánica, de transporte, señalización, equilibrio ácido base, anticuerpos, etc. (Whitford, 2005).

### **1.4.2 Unión de proteínas y ligandos**

Un ligando es una molécula capaz de unirse a una proteína con alta especificidad y afinidad. La interacción proteína-ligando provoca cambios conformacionales en las

proteínas y en algunos casos es causante de condiciones patológicas. Por tanto, su estudio es de crucial importancia (Du et al., 2016).

Un ejemplo de interacción proteína-ligando es la glicación. La glucosa tiene alta concentración en el plasma sanguíneo y actúa como agente reductor para reaccionar con el grupo amino libre de una proteína. Esta reacción es reversible y forma una base de Schiff, seguido por la formación reversible de un complejo Amadori. Este último producto sufre una serie de reacciones químicas cuyo producto final son los productos finales de la glicación avanzada (AGEs) (Rahman & Khalil, 2022).



**Figura 1-1 Formación de AGEs. Elaborado en ChemSketch**

En los pacientes con Diabetes Mellitus, la formación de AGEs se acelera por la alta concentración de glucosa en sangre. Múltiples trabajos han encontrado relación entre los AGEs y varias patologías, por ejemplo: arterosclerosis, nefropatía, neuropatía, Alzheimer, etc. (Carvajal Carvajal, 2015).

### 1.4.3 Análisis de proteínas por espectroscopía infrarroja

Las técnicas espectroscópicas se basan en medir la cantidad de luz que absorbe un compuesto como función de la longitud de onda. La principal ventaja de estos métodos es que no destruyen las muestras de análisis. Existen varios tipos de técnicas espectroscópicas, las más usuales son: espectroscopía infrarroja, espectroscopía

ultravioleta visible, y resonancia magnética nuclear (Wade & Simek, 2017). En general, las técnicas espectroscópicas son muy útiles para entender diferentes funciones y particularidades estructurales de las proteínas (Hough, 2017). En el presente trabajo, se realiza un caso de estudio con datos de espectroscopía infrarroja.

Un espectro infrarrojo es un gráfico que muestra picos de absorción que resultan de la excitación de modos vibracionales de las moléculas presentes en la muestra. Su principal ventaja es que requiere de una pequeña cantidad de muestra. No obstante, un obstáculo es la dificultad en la interpretación de los espectros causada por la superposición de las señales, y la determinación de correspondencia entre cambios en las bandas de absorción y alteraciones en las propiedades bioquímicas de los componentes de la muestra (Abrosimova et al., 2016).

#### **1.4.3.1 Espectrómetro Infrarrojo**

Un espectrómetro infrarrojo (IR) mide la luz infrarroja (2.5-25  $\mu\text{m}$  de longitud de onda) que absorbe un compuesto a diferentes frecuencias o longitudes de onda. Los espectrómetros infrarrojos modernos son espectrómetros infrarrojos por Transformada de Fourier (FTIR). Su funcionamiento se basa en el uso de un interferómetro (Hough, 2017). El interferómetro tiene un espejo fijo y un espejo móvil. El espejo móvil genera una diferencia variable en el camino óptico de los dos haces de luz, lo cual produce una señal detectada con información espectral. La luz emitida por la fuente se divide utilizando un divisor de haces. Aproximadamente la mitad se refleja hacia el espejo fijo y luego regresa al divisor de haces, pasando alrededor del 50% para llegar al detector. La otra mitad atraviesa el divisor en su primer encuentro, se refleja en el espejo móvil y regresa al divisor, donde alrededor del 50% se refleja hacia el detector. Al recombinarse, los dos haces interfieren, generando una interferencia constructiva o destructiva según la diferencia en el camino óptico. El instrumento registra la intensidad de la luz en relación con la posición del espejo móvil, lo cual se denomina interferograma, que es la transformada de Fourier del espectro. Mediante transformada de Fourier inversa, los datos medidos se convierten nuevamente en un espectro (Wade & Simek, 2017).

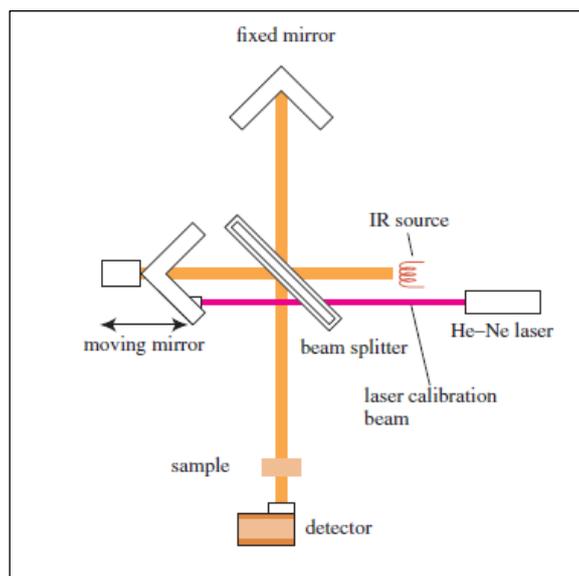


Figura 1-2 Diagrama de Interferómetro [(Wade & Simek, 2017)]

#### 1.4.3.2 Análisis de proteínas por espectroscopía infrarroja

La espectroscopía infrarroja es activamente usada para el análisis de la estructura secundaria de las proteínas (Hough, 2017) y para el estudio de cambios conformacionales, debido a que es una técnica simple, no destructiva y menos costosa que la RMN (Magalhães et al., 2020). Cuando una proteína absorbe luz infrarroja, los enlaces químicos presentes en los grupos funcionales vibran en una frecuencia que depende de la estructura y conformación de la molécula, causando un espectro único para cada muestra analizada. Varios trabajos han asignado los diferentes números de onda que presentan picos de absorción de luz infrarroja con características estructurales de la proteína. En la siguiente tabla se resumen asignaciones importantes:

Característica	Número de onda (1/cm)	Referencia
Amide II modes	1550	(Baiz et al., 2012)
$\nu_{C=C} (D)$	1570	(Baiz et al., 2012)
Aggregates	1620	(Magalhães et al., 2020)
$\nu_{C=C}$	1630	(Baiz et al., 2012)
$\nu_{C=O}$	1640	(Baiz et al., 2012)
$\alpha$ Helices A mode	1650	(Abrosimova et al., 2016)

Random Coils	1655 ± 2	(Abrosimova et al., 2016)
310 helices	Near Amide 1 modes as a single peak	(Abrosimova et al., 2016)
Beta sheets	1670	(Abrosimova et al., 2016)
Característica	Número de onda	Referencia

Tabla 1-1 Asignación de bandas de absorción IR en proteínas

Nótese la asignación en el número de onda 1620 a la formación de agregados. Algunas proteínas, como la albúmina sérica humana (HSA), tienden a auto-agregarse en estructuras poliméricas de fibrilla amiloide. Este proceso tiene dos partes: una nucleación inicial reversible, y un crecimiento irreversible del núcleo de agregación (Maciążek-Jurczyk et al., 2020). La agregación de proteínas está asociada a varias condiciones patológicas, como el Parkinson y el Alzheimer (Magalhães et al., 2020).

#### 1.4.4 Correlación espectroscópica bidimensional

La correlación espectroscópica bidimensional (2DCOS) es un concepto que se origina inicialmente para su aplicación en el análisis de señales de RMN, pero que luego se generaliza para todo tipo de señales espectroscópicas (Noda, 2016). En 2DCOS, se obtiene la intensidad de correlación espectral de una muestra perturbada, como una función de dos variables espectrales independientes, como números de onda de IR. La perturbación que se aplica sobre la muestra puede ser un cambio de temperatura, de concentración, el paso del tiempo, etc. El resultado es un gráfico de contornos bidimensional donde se muestran los picos de correlación. Un análisis de la posición, intensidad y signo de los picos permite extraer información significativa de la muestra (Noda, 2017).

##### 1.4.4.1 Modelo Matemático Resumido de 2DCOS

Sea  $A(v_k, t_i)$  un conjunto discreto de  $m$  espectros vibracionales que se miden en un sistema que sufre una perturbación, la cual provoca cambios en la intensidad espectral. La variable  $v_k$ , donde  $k = 1, 2, \dots, n$ , es la variable espectral. En el caso de 2DCOS-IR, los valores de  $v_k$  son los números de onda. La variable  $t_i$ , donde  $i = 1, 2, \dots, m$  es el valor cuantitativo de la variable de perturbación (p.ej. valores de concentración).

Se define el espectro dinámico  $\tilde{A}(v_k, t_i)$ :

$$\tilde{A}(v_k, t_i) = \begin{cases} A(v_k, t_i) - \bar{A}(v_k, t_i), & 1 \leq i \leq m \\ 0, & \text{en otro caso} \end{cases} \quad (1.1)$$

Donde  $\bar{A}(v_k, t_i)$  es el espectro de referencia, que puede ser el espectro inicial  $A(v_k, t_1)$ , el espectro final  $A(v_k, t_m)$ , el espectro promedio  $\bar{A}(v_k)$ , o el valor cero (0).

Se define el espectro de correlación síncrono  $\Phi(v_1, v_2)$ :

$$\Phi(v_1, v_2) = \frac{1}{m-1} \sum_{i=1}^m \tilde{A}(v_1, t_i) \cdot \tilde{A}(v_2, t_i) \quad (1.2)$$

El espectro de correlación asíncrono  $\Psi(v_1, v_2)$ :

$$\Psi(v_1, v_2) = \frac{1}{m-1} \sum_{i=1}^m \tilde{A}(v_1, t_i) \cdot \sum_{j=1}^m N_{ij} \tilde{A}(v_2, t_j) \quad (1.3)$$

Donde  $N_{ij}$  es un elemento de la matriz de transformación de Hilbert-Noda:

$$N_{ij} = \begin{cases} 0, & i = j \\ \frac{1}{\pi(j-1)}, & \text{en otro caso} \end{cases} \quad (1.4)$$

#### **1.4.4.2 Interpretación de los espectros de 2DCOS**

Un espectro de correlación bidimensional sincrónico es simétrico con respecto a la diagonal principal ( $v_1 = v_2$ ). Los picos de correlación que aparecen en la diagonal principal son autopicos, y representan la susceptibilidad de la señal espectral hacia el cambio inducido por la perturbación. Los picos fuera de la diagonal son picos deslocalizados, y representan cambios simultáneos o coincidentes de dos señales diferentes, lo cual sugiere un posible acoplamiento o relación entre las regiones coordinadas. El signo de un pico deslocalizado es positivo si la intensidad de las señales aumenta o disminuyen en la misma dirección, y negativo en caso contrario (Lasch & Noda, 2019).

El espectro asíncrono consta únicamente de picos deslocalizados. Un pico en un espectro asíncrono se genera solo si dos señales  $v_1$  y  $v_2$  se cambian fuera de fase entre ellas (retrasadas o aceleradas). Esta característica es útil para discriminar bandas sobrepuestas que tienen diferentes orígenes físicos (Noda, 2017).

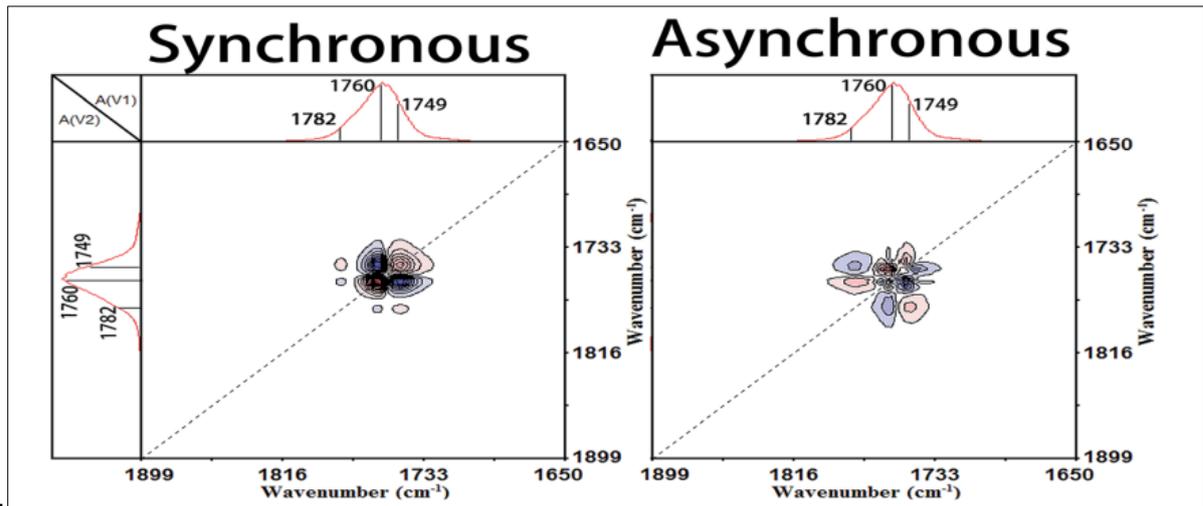


Figura 1-3 Ejemplos de espectros síncrono y asíncrono (Xue et al., 2017)

#### 1.4.4.3 Antecedentes en estudios de proteínas por 2DCOS

La capacidad de 2DCOS para resolver picos superpuestos y distinguir entre respuestas simultáneas o desfasadas lo convierte en una herramienta útil para el análisis de proteínas (Tao et al., 2018). 2DCOS ha sido utilizado para estudiar efectos de agregación polimérica en proteínas sometidas a cambios de concentración de ligandos (Yang et al., 2016). También para analizar cambios conformacionales en las proteínas producto de cambios en temperatura (He et al., 2017), pH (Gomaa et al., 2013), composición (Yang et al., 2016), entre otros. Además, en su trabajo, Rutherford et. al. (2022) exploran la capacidad de 2DCOS para detectar cambios conformacionales en la HSA por concentraciones de drogas tan bajas como 7  $\mu$ M.

# CAPÍTULO 2

## 2. METODOLOGÍA

### 2.1 Desarrollo del software para 2DCOS

Para desarrollar el software computacional orientado a la correlación espectroscópica bidimensional se utilizó lenguaje Python. La elección del lenguaje se basó en que Python cuenta con múltiples bibliotecas y dependencias que ofrecen herramientas robustas y eficientes para el análisis y procesamiento de datos. Además, tiene bibliotecas de visualización que permiten crear imágenes de alta calidad para la generación de los espectros síncrono y asíncrono. Por otro lado, Python tiene varias opciones para desarrollar una interfaz gráfica de usuario con todas las funcionalidades requeridas.

Anteriormente han sido creados varios programas orientados a la generación de gráficos de correlación espectroscópica bidimensional. No obstante, en el contexto actual, estos programas son de pago y/o no cuentan con todas las funcionalidades que el usuario puede requerir para un análisis exhaustivo. El programa 2DShige desarrollado en Python por los profesores Shigeaki Morita y Yukihiro Ozaki (Shigeaki & Yukihiro, 2005) es de acceso libre, pero genera gráficos poco detallados y no provee al usuario de varias opciones para adecuar la apariencia del gráfico. Por otro lado, el programa desarrollado en Matlab “mat2dcorr” (Peter Lasch, 2023) desarrolla gráficos más detallados y elaborados que 2DShige. Sin embargo, no está disponible de forma gratuita, y también presenta limitaciones en cuanto a la personalización de los gráficos. Otro programa enfocado a la correlación espectroscópica bidimensional fue desarrollado por OriginLab (OriginLab Technical Support, 2018) y si bien el programa es gratuito, el software OriginLab sí requiere licencia de pago. Al igual que el programa de Matlab, la herramienta desarrollada en OriginLab no permite al usuario manipular los ajustes del gráfico.

Tomando en cuenta lo antes expuesto, el desarrollo del software presentado en el siguiente trabajo se centra en los siguientes factores:

- Ser un programa de libre acceso
- Su uso debe ser sencillo/intuitivo
- Generar gráficos de correlación bidimensional precisos y detallados

- Ofrecer al usuario múltiples opciones de configuración y ajuste de los gráficos

Para el desarrollo del software enfocado a correlación espectroscópica bidimensional se diseñó una interfaz gráfica y se programó un modelo matemático de acuerdo con las ecuaciones discretizadas para espectros síncrono y asíncrono. Al acoplar los códigos se obtienen resultados, y con la retroalimentación de un usuario experto se realizan las modificaciones requeridas para mejorar el programa.

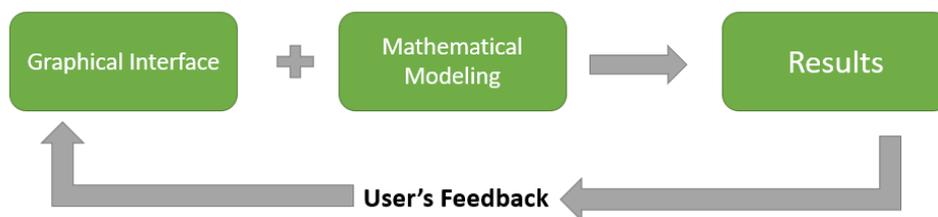


Figura 2-1 Desarrollo del programa

### 2.1.1 Librerías y dependencias utilizadas

Para el desarrollo del software, se utilizaron las siguientes herramientas:

- Para el desarrollo de la interfaz de la aplicación se utiliza la biblioteca gráfica PyQt. La elección de PyQt se basa en su capacidad para crear interfaces gráficas interactivas completas y funcionales.
- Los datos de espectroscopía se almacenan en dataframes, utilizando la librería Pandas. Esto se debe a la flexibilidad para manipular datos que se obtiene en un dataframe, además de las múltiples operaciones de cálculos y reordenamientos que pueden realizarse en este tipo de datos.
- Para cálculos numéricos, se utiliza la librería NumPy. Esta librería ofrece una amplia gama de métodos para realizar cálculos de álgebra lineal. Además, para almacenar los datos usa arreglos optimizados para operaciones numéricas.
- Para mostrar el diagrama de contornos que se obtiene a partir de los cálculos, se emplea la librería Matplotlib. Esto se debe a que ofrece un alto grado de personalización de los gráficos. Además, genera imágenes de alta calidad y con una amplia gama de formatos de salida disponibles.

### 2.1.2 Desarrollo de la interfaz gráfica

Para desarrollar la interfaz gráfica se planteó el siguiente diagrama de flujo de usuario:

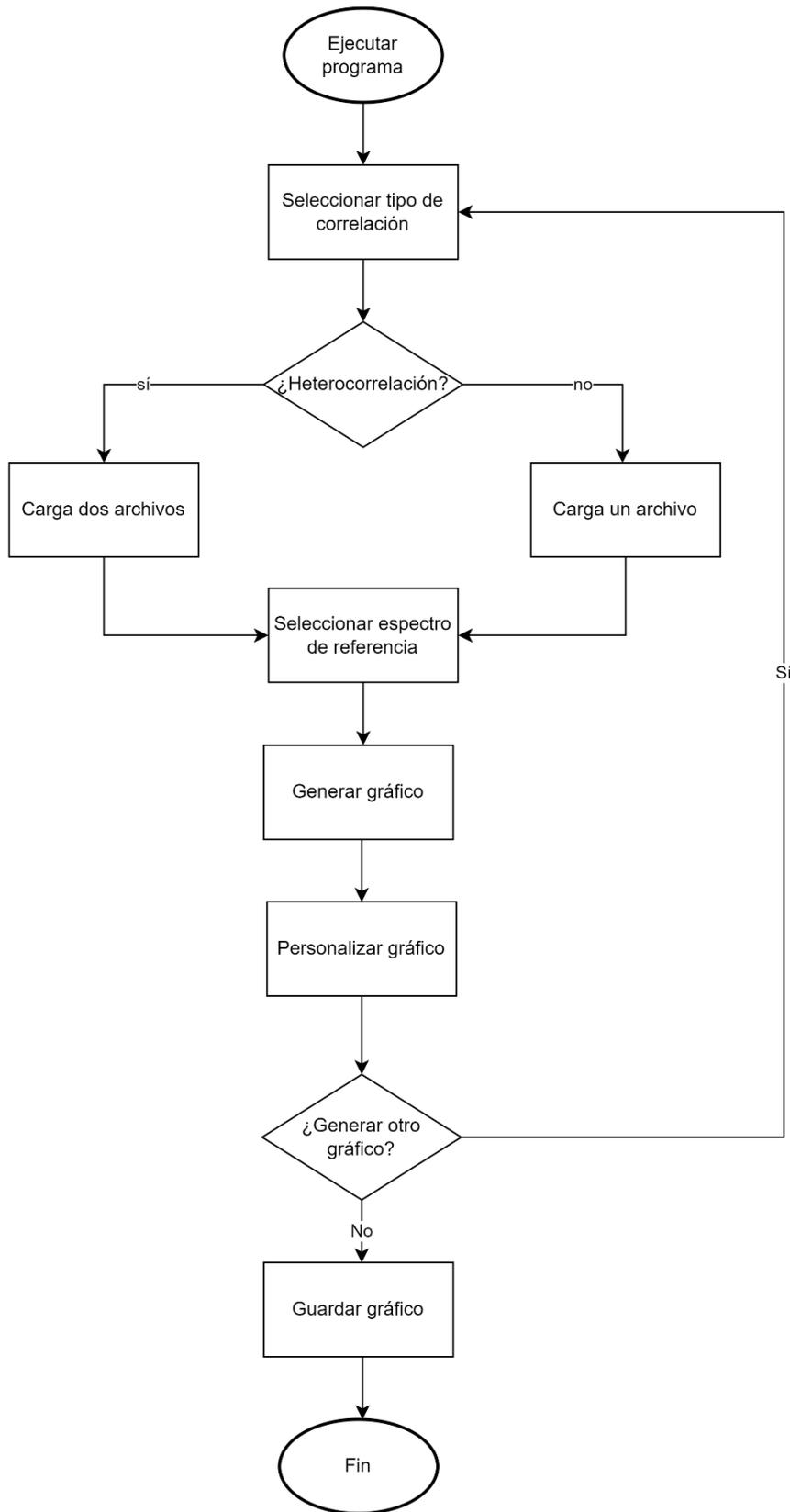


Figura 2-2 Diagrama de Flujo de Usuario

Se diseñó un primer prototipo de baja resolución utilizando Figma. El prototipo se mostró a un investigador con experiencia en el uso de 2DCOS y sus recomendaciones se tomaron en cuenta para el desarrollo de la interfaz gráfica.

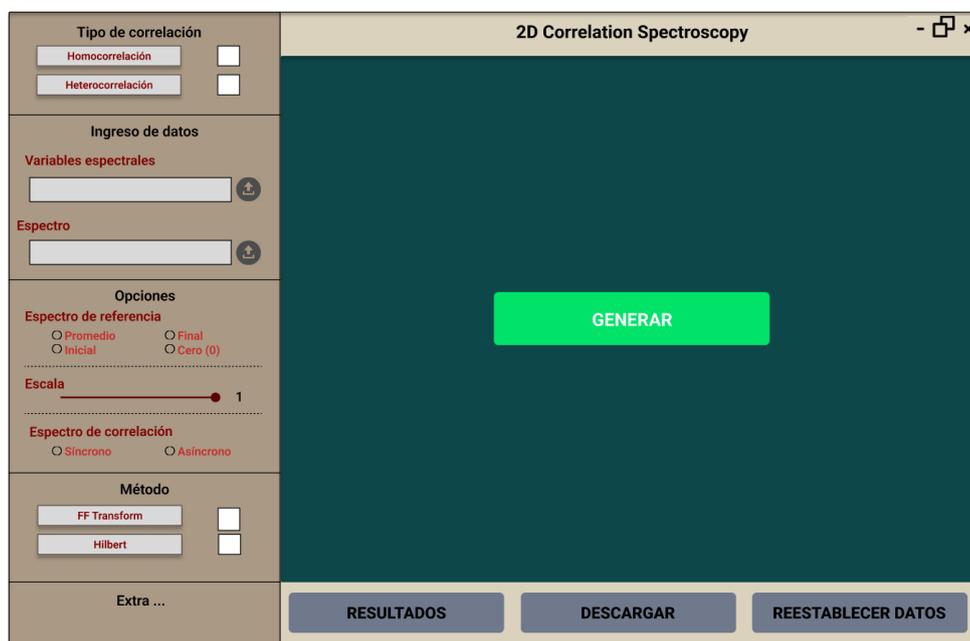


Figura 2-3 Diseño inicial de interfaz

### 2.1.3 Obtención de los espectros de 2DCOS

Una vez que se presiona el botón “Generar”, se muestran los gráficos de correlación bidimensional síncrono y asíncrono. Cuando están generados los gráficos, el usuario puede realizar ajustes que actualizan el gráfico automáticamente.

Los datos deben estar organizados en un archivo de la siguiente manera:

- Una columna inicial de índices que indica el valor de la variable independiente (ej. Longitud de Onda).
- De manera opcional, una fila inicial que indica el valor numérico de la variable de perturbación (ej. Concentración, tiempo, etc.), caso contrario se asume que están igualmente espaciados.
- Los datos de mediciones espectroscópicas que completan la matriz.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1		% Transmitancia												
2		Concentración de BSA (mg/ml)												
3	Longitud de onda [1/cm]	0.5	1	1.5	VARIABLE DE PERTURBACIÓN				7	8	9	10		
4	VARIABLE INDEPENDIENTE													
5	1801.	100.039	100.188	101.247	100.106	100.917	100.087	100.822	100.042	100.313	99.8838	100.626	101.269	100.522
6	1799.	100.223	100.291	101.201	100.07	101.087	100.042	100.685	100.095	100.198	99.7304	100.665	101.275	100.466
7	1797.	100.153	100.197	101.117	99.9893	100.907	100.099	100.508	99.928	100.082	99.5631	100.471	101.306	100.283
8	1795.	99.9885	100.116	101.17	100.093	100.846	100.205	100.604	99.863	100.27	99.7523	100.544	101.428	100.301
9	1793.	100.009	100.307	101.136	100.367	101.136	100.235	100.844	100.093	100.558	100.103	100.903	101.581	100.502
10	1791.	100.087	100.244	101.146	100.146	101.146	99.9844	100.793	100.147	100.403	99.9362	100.864	101.443	100.342
11	1789.	99.9643	100.164	101.146	100.146	101.146	99.8198	100.704	99.875	100.186	99.6511	100.551	101.263	100.131
12	1787.	99.863	100.089	101.319	100.146	101.146	99.8259	100.682	99.6347	100.205	99.8168	100.416	101.165	100.233
13	1785.	99.8258	99.974	101.226	100.04	101.146	100.538	99.7088	100.153	99.9976	100.457	100.99	100.423	
14	1783.	99.9276	99.9501	101.214	99.9429	100.146	100.516	99.9348	100.197	99.9209	100.613	101.012	100.481	
15	1781.	99.9549	99.9405	101.183	99.8732	101.013	100.516	100.011	100.366	99.856	100.658	101.186	100.312	
16	1780.	99.7609	99.982	101.062	99.9596	100.851	99.907	100.144	99.7377	100.453	101.181	100.121		
17	1778.	99.7778	100.152	100.943	99.9626	100.746	99.8337	100.028	99.7123	100.449	101.107	100.062		
18	1776.	99.8834	100.167	100.985	100.026	100.858	99.9177	100.821	99.8234	100.735	101.256	100.272		
19	1774.	100.29	100.236	101.31	100.205	101.323	100.208	101.127	100.39	101.203	101.604	100.664		
20	1772.	100.424	100.281	101.483	100.37	101.446	100.418	101.16	100.372	101.394	101.86	100.704		
21	1770.	99.8802	100.085	101.406	100.241	100.99	100.206	100.805	100.335	100.671	101.138	101.52	100.463	
22	1768.	99.67	100.236	101.201	100.175	100.911	100.115	100.632	100.23	100.55	99.74	100.998	101.298	100.292
23	1766.	99.8344	100.309	101.109	100.146	101.16	100.061	100.619	100.013	100.466	99.7689	100.793	101.401	100.328
24	1764.	100.037	100.2	101.232	100.347	101.367	100.107	100.663	99.9901	100.488	99.9573	100.682	101.415	100.277
25	1762.	99.8988	99.8992	101.096	100.204	101.056	100.052	100.485	99.7714	100.332	99.9423	100.514	101.17	100.038
26	1760.	99.8160	99.6700	100.85	99.7303	100.601	99.884	100.46	99.653	100.16	99.7017	100.306	100.933	99.6501

Figura 2-4 Modelo de almacenamiento de datos en Excel

Los formatos de archivos para cargar datos que acepta el programa son: documentos de texto (.txt), archivos de valores separados por comas (.csv) y documentos Excel (.xlsx).

En los documentos Excel, se distinguen dos casos:

- Caso base: Los datos que se tienen son únicamente numéricos, similar a lo que se tiene en un archivo .csv.
- Caso especial: Hay celdas con letras, y/o las columnas están etiquetadas. En este caso el usuario debe proporcionar información sobre la fila (indicar número de fila) a partir de la cual comienzan los datos de mediciones espectroscópicas, y el rango de columnas (indicar en el formato "A:F").

Con los datos cargados y el espectro de referencia especificado, se crea un objeto a partir de una clase previamente desarrollada, que permite guardar atributos correspondientes a los datos y ajustes que selecciona el usuario, y contiene los métodos de cálculo requeridos. El usuario puede modificar los siguientes ajustes:

- Número de contornos
- Mapa de colores
- Color de las líneas de contorno
- Intensidad de color de gráfico
- Intensidad de color de las líneas de contorno
- Diagonal principal o antidiagonal en cada espectro
- Eje horizontal creciente o decreciente
- Gráfico mostrado (síncrono, asíncrono o ambos)

Dentro del objeto, se obtienen los espectros síncrono y asíncrono de la siguiente manera:

1. Se verifica que los datos subidos son válidos
2. Sea  $v_k$ , donde  $k = 1, 2, 3, \dots, n$ , el valor de la variable independiente (ej. Número de onda), y sea  $t_i$ , donde  $i = 1, 2, 3, \dots, m$  el valor de la variable de perturbación (ej. Concentración, tiempo, etc.)
3. Las matrices  $A_1(v_k, t_i)$ ,  $A_2(v_k, t_i)$  son los conjuntos de datos de mediciones espectroscópicas. Si se trata de homocorrelación, entonces  $A_2 = A_1$ .
4. Según el espectro de referencia seleccionado por el usuario, se obtiene el espectro dinámico de acuerdo con la ecuación 1.1:

```
# creates the dynamic spectrum depending on the chosen reference
if ref == 'zero':
    self.spec1 = spec1
    self.spec2 = spec2
elif ref == 'mean':
    self.spec1 = spec1.sub(self.describe1['mean'], axis=0)
    self.spec2 = spec2.sub(self.describe2['mean'], axis=0)
elif ref == 'min':
    self.spec1 = spec1.sub(self.describe1['min'], axis=0)
    self.spec2 = spec2.sub(self.describe2['min'], axis=0)
elif ref == 'max':
    self.spec1 = spec1.sub(self.describe1['max'], axis=0)
    self.spec2 = spec2.sub(self.describe2['max'], axis=0)
elif ref == 'ini':
    self.spec1 = spec1.sub(spec1[1], axis=0)
    self.spec2 = spec2.sub(spec2[1], axis=0)
elif ref == 'end':
    self.spec1 = spec1.sub(spec1[spec1.columns[-1]], axis=0)
    self.spec2 = spec2.sub(spec2[spec2.columns[-1]], axis=0)
```

Figura 2-5 Obtención de espectro dinámico en Python

5. El espectro síncrono se calcula según la ecuación 1.2:

```
self.syncr = pd.DataFrame(self.spec1.values @ self.spec2.values.T / float(len(self.spec1.columns) - 1),
                          index=self.spec1.index,
                          columns=self.spec2.index)
```

Figura 2-6 Cálculo de espectro síncrono en Python

6. Se define la matriz de Hilbert-Noda:

```
def noda(self):
    nn = len(self.spec1.columns)
    ## The actual values are calculated only once
    fill_pattern = [1. / (np.pi * float(i)) if i != 0 else 0. for i in range(nn)]
    noda_matrix = np.empty((nn, nn), dtype=float)
    for i in range(nn):
        noda_matrix[i, :] = [-fill_pattern[i - j] for j in range(i)] + [fill_pattern[j] for j in range(nn - i)]

    return noda_matrix
```

Figura 2-7 Obtención de matriz Noda en Python

7. El espectro asíncrono se calcula:

```

self.asyncr = pd.DataFrame(
    self.spec1.values @ self.noda() @ self.spec2.values.T / (len(self.spec1.columns) - 1),
    index=self.spec1.index,
    columns=self.spec2.index) # .transpose()

```

Figura 2-8 Cálculo de espectro asíncrono en Python

8. De acuerdo con los ajustes de gráfico seleccionados por el usuario, se muestra el diagrama de contorno para los espectros síncrono y asíncrono.
9. Si el usuario realiza nuevos ajustes en los gráficos, estos se grafican automáticamente, sin necesidad de calcular los espectros síncrono y asíncrono.
10. El usuario puede observar los gráficos tridimensionales correspondientes a los diagramas de contornos.

## 2.2 Caso de estudio: interacción de Albúmina Sérica Bovina con Glucosa y Metformina

Para la aplicación de la herramienta de correlación bidimensional realizada se tomaron mediciones FTIR de sistemas con albúmina sérica bovina (BSA), glucosa y metformina. Se seleccionó la proteína BSA por su similitud con la albúmina sérica humana (HSA) (Du et al., 2016). Se seleccionó la glucosa como ligando porque se conoce que esta reacciona con la proteína y desencadenan el proceso descrito en el capítulo 1. La metformina se seleccionó porque es un medicamento ampliamente utilizado como hipoglucemiante, y además actúa químicamente compitiendo con la proteína por los grupos carbonilo de la glucosa (Khan et al., 2021).

La hipótesis experimental es que con la aplicación de 2DCOS se podrá evidenciar que la glucosa induce cambios estructurales en la proteína, y que la metformina es capaz de inhibirlos y/o revertirlos.

### 2.2.1 Sistemas de control

Se prepararon sistemas de control unitarios, secundarios y terciarios.

- Sistemas unitarios: contienen únicamente BSA, glucosa o metformina
- Sistemas binarios: BSA + glucosa, BSA + metformina
- Sistema ternario: BSA + glucosa + metformina

Los sistemas de control unitarios son necesarios para tener espectros referenciales de cada compuesto sin la acción de una perturbación externa. Con los sistemas binarios se consigue estudiar la interacción de la proteína con cada ligando por separado. El sistema ternario permite conocer si los cambios estructurales y conformacionales de la proteína varían por la presencia de los dos ligandos.

### **2.2.2 Variables de perturbación**

Sobre los sistemas se aplicaron dos variables de perturbación: concentración y tiempo. Al usar la concentración como una variable de perturbación podemos estudiar la interacción entre los componentes de un sistema, debido a los cambios en los modos vibracionales que suceden. En 2DCOS, la concentración se usa como variable de perturbación para determinar las regiones estructurales que varían con más intensidad al variar la concentración de un componente, y cómo estas regiones se acoplan entre sí (Tao et al., 2018).

Por otro lado, el tiempo como variable de perturbación permite analizar la dinámica del sistema. Usando 2DCOS, puede monitorearse los cambios en los modos vibracionales a medida que avanza una reacción química en el tiempo, como la reacción entre la glucosa y la proteína, o bien a medida que uno de los componentes sufre cambios estructurales, como la formación de agregados en la BSA (Tao et al., 2018).

### **2.2.3 Materiales**

- Espectrómetro infrarrojo Thermo Scientific modelo Nicolet iS10, dotado de celda ATR horizontal con cristal de ZnSe a 45°
- Viales de laboratorio plásticos de 1.5 ml
- Micropipeta
- Solución de búfer fosfato-salino (PBS) 20 mM
- Albúmina Sérica Bovina (BSA)
- Metformina aislada y caracterizada en el laboratorio
- Glucosa grado analítico

### **2.2.4 Preparación de muestras y medición**

Para la preparación de las muestras se prepararon las siguientes soluciones madre:

- Glucosa 100 mg/ml
- Metformina 100 mg/ml

- BSA 100 mg/ml

Para preparar las muestras se calcula la cantidad de solución madre requerida para la concentración objetivo. Con ayuda una pipeta electrónica se agrega esta cantidad a un vial. El volumen del vial se completa con la solución búfer PBS.

En el caso de los sistemas perturbados por la concentración, se prepararon varios viales para cada sistema, cada uno con un volumen distinto de PBS y los demás componentes según el tipo de sistema (unitario, binario o ternario). Para los sistemas perturbados por tiempo, se prepararon viales para las concentraciones definidas y se mantuvieron aislados a lo largo de cinco días, tomando una medición diaria.

Para tomar mediciones se siguió el siguiente procedimiento:

- Se toma lectura del búfer como background para eliminar su aporte al espectro de las muestras.
- Se extraen 300  $\mu$ L de cada vial utilizando una micropipeta.
- La mezcla se introduce en el cristal de la celda horizontal esparciéndola uniformemente.
- Se cierra el espectrómetro, se registra la medición en la computadora conectada, y se guarda el espectro.
- Se retira el líquido del compartimento usando la pipeta y se retira la humedad del cristal con franela de microfibra.

### 2.2.5 Sistemas perturbados por concentración

Se tomaron mediciones FTIR por duplicado de cada sistema aumentando las concentraciones de uno de los componentes. Las concentraciones establecidas se resumen en la siguiente tabla:

Clasificación	Sistema	Concentración Fija (mg/ml)	Rango de Concentración Variable (mg/ml)
Unitario	BSA	-	BSA: 0 – 10
	Glucosa	-	Glucosa: 0 – 40
	Metformina	-	Metformina: 0 – 40
Binario	BSA (conc. fija) + Glucosa (conc. variable)	10	Glucosa: 0 – 40
	BSA (conc. variable) + Glucosa (conc. fija)	15	BSA: 0 – 10
	BSA (conc. fija) + Metformina (conc. variable)	10	Metformina: 0 – 40

	BSA (conc. variable) + Metformina (conc. fija)	15	BSA: 0 – 10
Ternario	BSA (conc. fija) + glucosa (conc. fija) + metformina (conc. variable)	BSA: 10 Glucosa: 15	Metformina: 0 – 40

Tabla 2-1 Sistemas perturbados por concentración

### 2.2.6 Sistemas perturbados por tiempo

Se tomaron mediciones FTIR por duplicado de cada sistema diariamente a lo largo de cinco días. Las concentraciones establecidas se resumen en la siguiente tabla:

Clasificación	Sistema	Concentración (mg/ml)
Unitario	BSA	10
	Glucosa	0 – 40
	Metformina	0 – 40
Binario	BSA + glucosa	BSA: 10 Glucosa: 0 – 40
	BSA + metformina	BSA: 10 Metformina: 0 – 40
Ternario	BSA + glucosa + metformina	BSA: 10 Glucosa: 15 Metformina: 0 – 40

Tabla 2-2 Sistemas perturbados por tiempo

### 2.2.7 Almacenamiento de los datos

Los datos medidos de FTIR para cada sistema se guardaron en formato “. SPA”. Utilizando el software SpectraGryph se abrió cada espectro, y se posteriormente se guardó en formato “.csv”. Para ordenar los datos acorde con lo requerido por el software desarrollado, se creó un documento Excel y se organizaron según la figura 2-4.

### 2.2.8 Costos asociados

El software computacional es de acceso libre, y por tanto puede ser utilizado de manera gratuita. Para realizar un caso de estudio similar al que se desarrolla en el presente trabajo, se requiere tomar mediciones espectroscópicas, y por tanto los costos están asociados a la obtención de reactivos y del dispositivo de espectroscopía.

La mayoría de los laboratorios de investigación y universidades constan de equipos convencionales de espectrometría de IR o UV-Vis, motivo por el cual el costo de un caso de estudio de correlación bidimensional es únicamente la obtención de los reactivos.

# CAPÍTULO 3

## 3. RESULTADOS Y ANÁLISIS

### 3.1 Desarrollo de Software Computacional

Se desarrolló un programa en lenguaje Python que genera gráficos síncrono y asíncrono de correlación espectroscópica bidimensional. El programa es aplicable para todo tipo de datos que resulte de mediciones espectroscópicas, por ejemplo: FTIR, UV-Vis, RMN, etc. La interfaz gráfica desarrollada se muestra en la Figura 3-1 Captura de Pantalla del Software Desarrollado. Los espectros generados corresponden al sistema unitario BSA con la concentración como variable de perturbación, en el rango de 0.5 a 10 mg/ml.

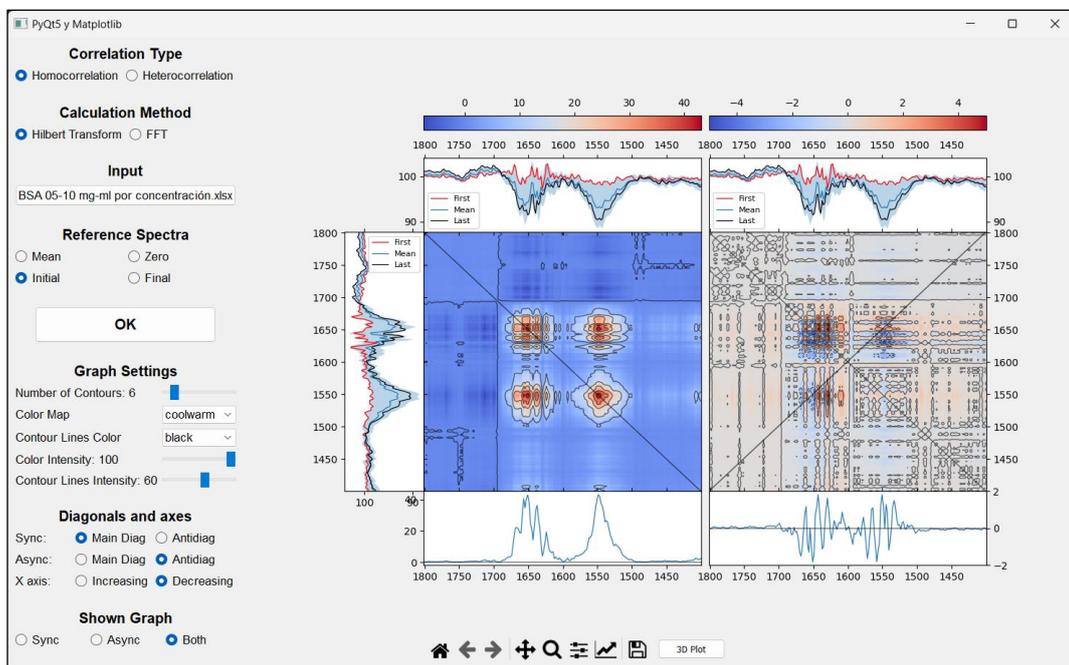


Figura 3-1 Captura de Pantalla del Software Desarrollado

Se programó la carga de datos de manera que se admiten archivos con los formatos: .txt, .csv, .xlsx. Cuando se carga un archivo con formato .xlsx (Excel), se abre una nueva pestaña en la que se diferencia entre los casos “especial” y “base”, que fueron descritos con detalle la sección 2.1.3. En el “caso especial” se indica el rango de columnas en el que se encuentran los datos, y la fila inicial. También se incluyó la opción de indicar si las columnas están etiquetadas, de forma que puede interpolarse dentro de la variable

de perturbación en caso de ser necesario para obtener datos igualmente espaciados. Se trabajó con el modelo de almacenamiento de datos que se muestra en la Figura 2-4 Modelo de almacenamiento de datos en Excel.

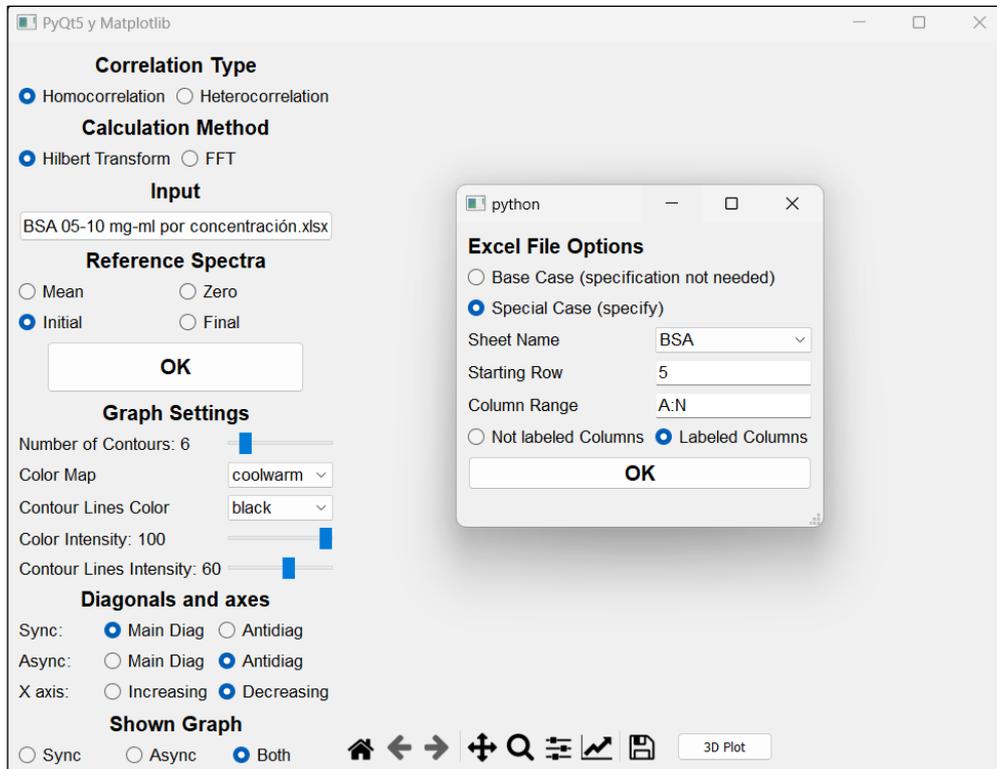


Figura 3-2 Pestaña Para Ingreso de Datos en Excel

Al generar los gráficos pueden modificarse los ajustes y realizar las manipulaciones establecidas en el capítulo anterior. También, puede obtenerse una imagen tridimensional interactiva presionando el botón "3D Plot".

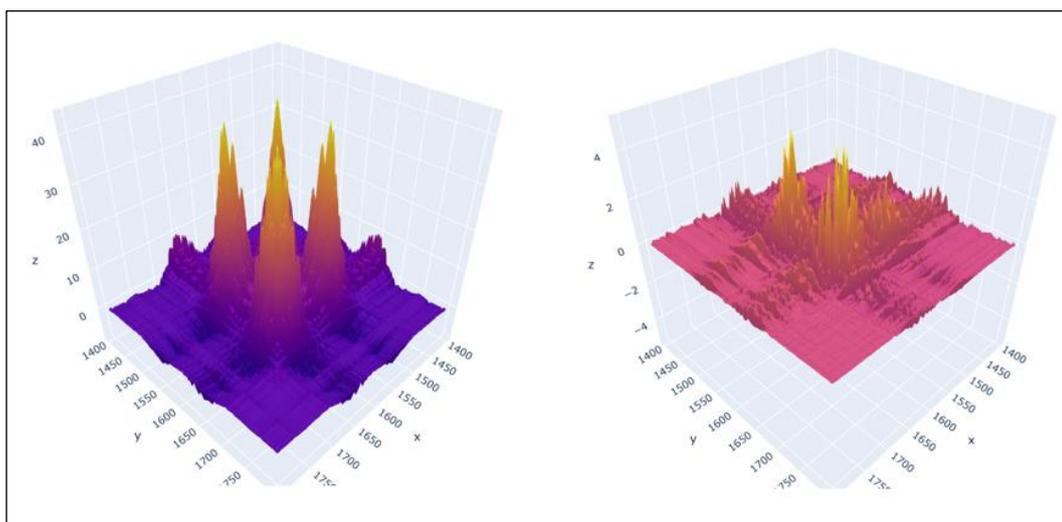


Figura 3-3 Gráficos tridimensionales para los espectros síncrono y asíncrono

En la Tabla 3-1 **Comparación Con Otras Herramientas De 2DCOS** se indican las características y funcionalidades del programa desarrollado y su comparación con los programas mat2dcorr, OriginLab y 2DShige.

Característica	Software Desarrollado	2DShige	Mat2dcorr	OriginLab
Genera gráficos síncrono y asíncrono de 2DCOS	✓	✓	✓	✓
Permite homo y heterocorrelación	✓	✓	✓	✓
Admite archivos .csv, .txt, .xlsx	✓	X	X	X
Permite indicar si las columnas están etiquetadas	✓	X	X	X
Permite cambiar el espectro de referencia	✓	✓	✓	✓
Permite personalizar la apariencia del gráfico (número de contornos, mapa de colores, intensidad de colores, intensidad de líneas)	✓	X	✓	X
Permite cambiar las diagonales mostradas y tendencia de eje X	✓	X	X	X
Muestra los gráficos IR	✓	X	X	X

Muestra el valor en la diagonal	✓	X	X	X
Muestra valores en las coordenadas al navegar con el cursor	✓	✓	✓	X
Permite realizar zoom a una región del gráfico	✓	X	X	✓
Permite guardar el gráfico como imagen	✓	X	✓	✓
Genera gráficos tridimensionales	✓	X	X	X
Libre Acceso	✓	✓	X	X

Tabla 3-1 Comparación Con Otras Herramientas De 2DCOS

### 3.2 Caso de Estudio: Interacción de BSA con Glucosa y Metformina

Se desarrolló un caso de estudio de la interacción entre BSA, Glucosa y Metformina. Se tomaron las mediciones de IR de acuerdo con los sistemas establecidos en las secciones 2.2.5 y 2.2.6. Se extrajeron los datos de IR de la región comprendida entre 1400 y 1800  $\text{cm}^{-1}$ , donde se contiene la información dinámica y estructural de la proteína (Baiz et al., 2012).

#### 3.2.1 Análisis de Espectros IR

Para la proteína BSA, se obtuvo el espectro IR mostrado en la Figura 3-4 Espectro IR de BSA en Estado Basal para una concentración de 10 mg/ml, y en estado nativo (sin acción de perturbación externa). En el espectro se distinguen dos regiones de alta absorbancia aproximadamente entre 1470 y 1580  $\text{cm}^{-1}$ , y 1615 y 1680  $\text{cm}^{-1}$ , que constituyen las regiones de Amida II y Amida I, respectivamente. La región Amida II corresponde en su mayoría a flexiones de enlaces C-N y estiramiento en plano de N-H, y la región de Amida I a vibraciones de enlaces C=O (López-Lorente & Mizaikoff, 2016). La región Amida I es sensible a las particularidades de la estructura secundaria de la proteína, y por este motivo se asignan picos a modos vibracionales de conformaciones estructurales. En el caso de la BSA, se trata de una proteína con estructura secundaria conformada por una presencia mayoritaria de  $\alpha$ -hélices (Abrosimova et al., 2016). No obstante, en el espectro

obtenido se observa un pico en  $1674\text{ cm}^{-1}$  característico de las estructuras  $\beta$ -plegadas. Este comportamiento fue descrito por Grdadolnik (2010), indicando que en condiciones de hidratación existen modificaciones en la estructura secundaria, con la posibilidad de intercambios reversibles entre  $\alpha$ -hélices y hojas  $\beta$ .



Figura 3-4 Espectro IR de BSA en Estado Basal

En la Tabla 3-2 se resumen los picos observados en el espectro IR mostrado en la figura 3.4 y su asignación en base a la tabla 1-1.1.

<b>BSA (10 mg/ml)</b>		
<b>Pico</b>	<b>Referencia (cm-1)</b>	<b>Observado (cm-1)</b>
Amide II modes	1550	1545
$\nu_{C=C} (D)$	1570	1569
$\nu_{C=C}$	1630	1633
$\nu_{C=O}$	1640	1643
$\alpha$ Helices A mode	1650	1649
Random Coils	1655	1657
310 helices	Near Amide 1 modes as a single peak	1664
Presence of Beta sheets	1670	1674

Tabla 3-2 Picos Asignados a BSA en Estado Nativo

En la Figura 3-5 se muestran los espectros obtenidos a lo largo de cinco días para el sistema unitario BSA a concentración de 10 mg/ml. Los picos observados en la región amida 1 (1600-1700 cm<sup>-1</sup>) disminuyen su intensidad y se desplazan a la izquierda a lo largo del tiempo. Tomando en consideración que dicha región es altamente sensible a los modos vibracionales de la estructura secundaria de la proteína, entonces de acuerdo con la Ley de Beer-Lamber, misma que relaciona la absorbancia de la luz con la concentración del absorbente, el comportamiento sugiere modificaciones conformacionales de la proteína, ya sea por pérdida de estructura secundaria, y/o formación de agregados proteicos. Por otra parte, la región de amida II comprendida entre 1500 y 1570 cm<sup>-1</sup> no presenta una tendencia clara. Según Chatterley et. al. (2022): “Los modos de amida II no permiten relacionarlos con precisión a particularidades de la estructura secundaria”.

Se observa también un crecimiento de pico cercano a 1620 cm<sup>-1</sup> que podría estar asociado a la presencia de proteínas agregadas, y que es característico de hojas  $\beta$  intermoleculares (Wilkosz et al., 2020).

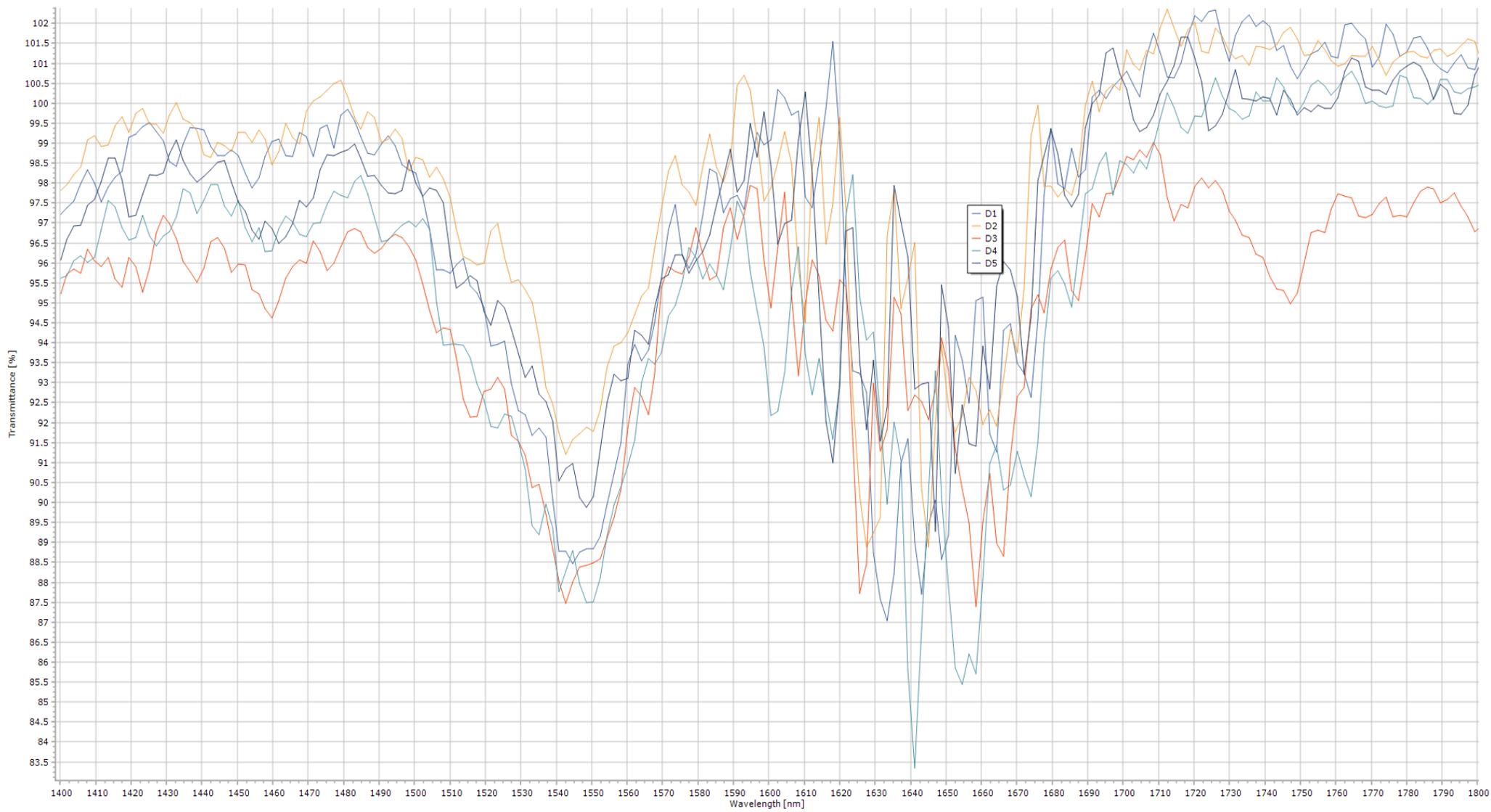


Figura 3-5 Espectro IR de BSA Durante Cinco Días

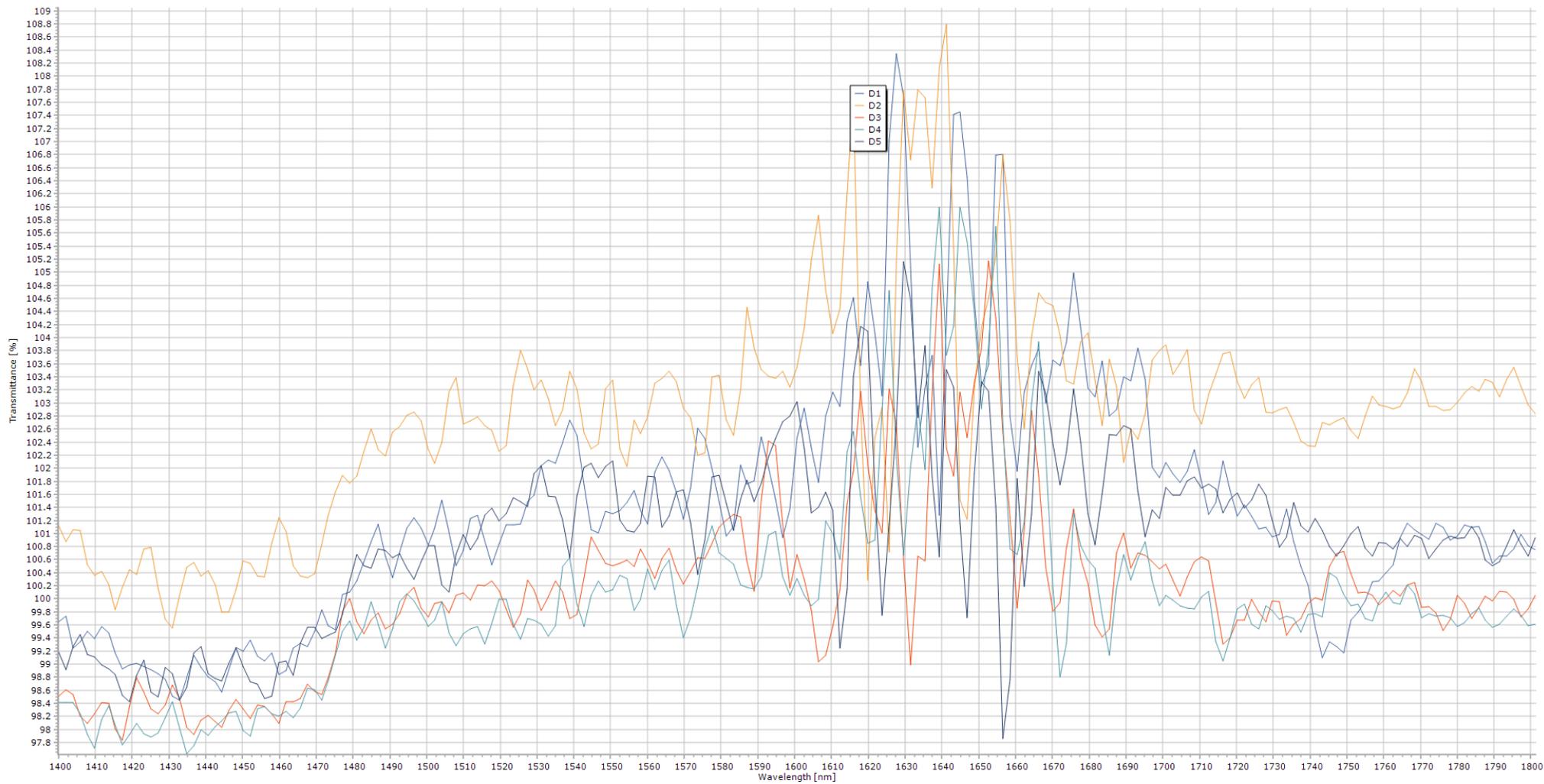


Figura 3-6 Espectro IR de Glucosa durante cinco días

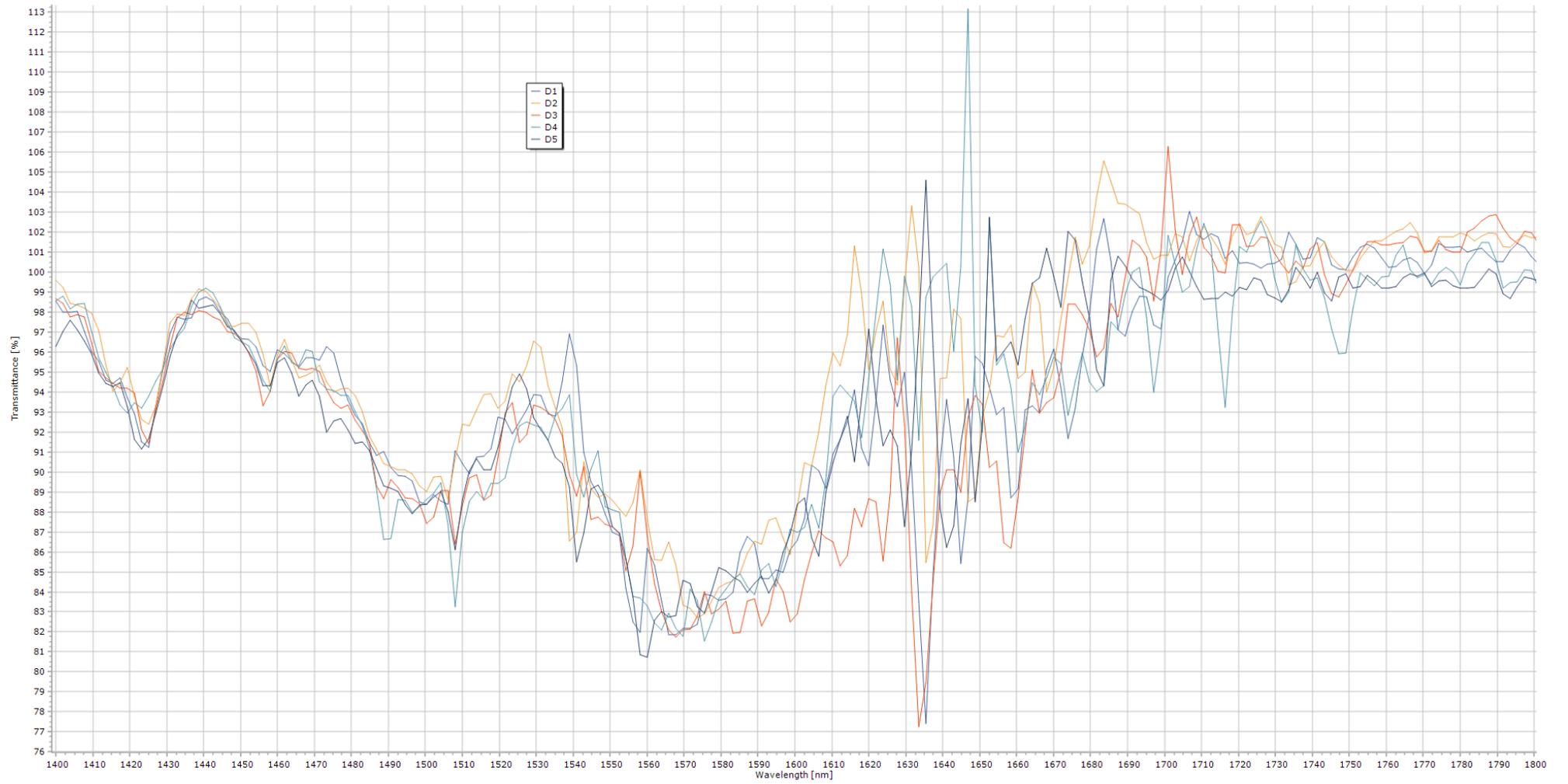


Figura 3-7 Espectro IR de Metformina durante cinco días

Los espectros IR obtenidos de glucosa y metformina a lo largo de cinco días a una concentración de 12 mg/ml se muestran en las figuras Figura 3-6 y Figura 3-7. En ninguno de los casos se observa una tendencia que permita identificar un cambio estructural en los componentes a lo largo del tiempo.

Para el caso de la glucosa, se observa que el porcentaje de transmitancia es cercano al 100% a lo largo de la región de análisis. La escala en el eje vertical verifica que los picos generados son despreciables en comparación a la absorción que se genera en los espectros de BSA y de Metformina.

En el caso de la metformina, sí se tiene una actividad marcada en la región de 1400 a 1800  $\text{cm}^{-1}$  que está caracterizada por el estiramiento C=N alrededor de 1660  $\text{cm}^{-1}$  y flexión C-H en la región de 1558  $\text{cm}^{-1}$  (Sabbagh et al., 2022). Tales zonas de absorción causan superposiciones al analizar un espectro IR de un sistema que contenga metformina y BSA.

### **3.2.2 Resultados de 2DCOS en Sistemas Perturbados por Tiempo**

Se generaron los espectros síncronos correspondientes a los sistemas perturbados por tiempo. Los espectros unitarios de correlación bidimensional obtenidos para BSA a 10 mg/ml, glucosa a 20 mg/ml y metformina a 20 mg/ml se muestran en las figuras Figura 3-10, Figura 3-9 y Figura 3-8, respectivamente.

El espectro bidimensional de la proteína muestra un pico de autocorrelación en 1620  $\text{cm}^{-1}$ , lo que permite verificar la tendencia que se observó en los espectros IR. Este pico está asociado a la presencia de agregados de proteína. Se observan picos de correlación deslocalizados acoplados con el pico de 1620  $\text{cm}^{-1}$ , con la particularidad de que la región entre 1615 hasta 1650  $\text{cm}^{-1}$  es negativa, y la región cercana a 1655  $\text{cm}^{-1}$  es positiva. Este comportamiento sugiere que la aparición de agregados ocurre a medida que se pierde la estructura secundaria que causa actividad infrarroja en la región amida I, y también ocurre a medida que aumenta la presencia de espirales aleatorias (1655  $\text{cm}^{-1}$ ).

Los espectros de 2DCOS de la metformina y la glucosa verifican lo observado en los espectros IR. No presentan una tendencia clara, lo cual provoca imágenes ruidosas y con pequeños picos de autocorrelación asociados a fluctuaciones en los diferentes espectros IR.

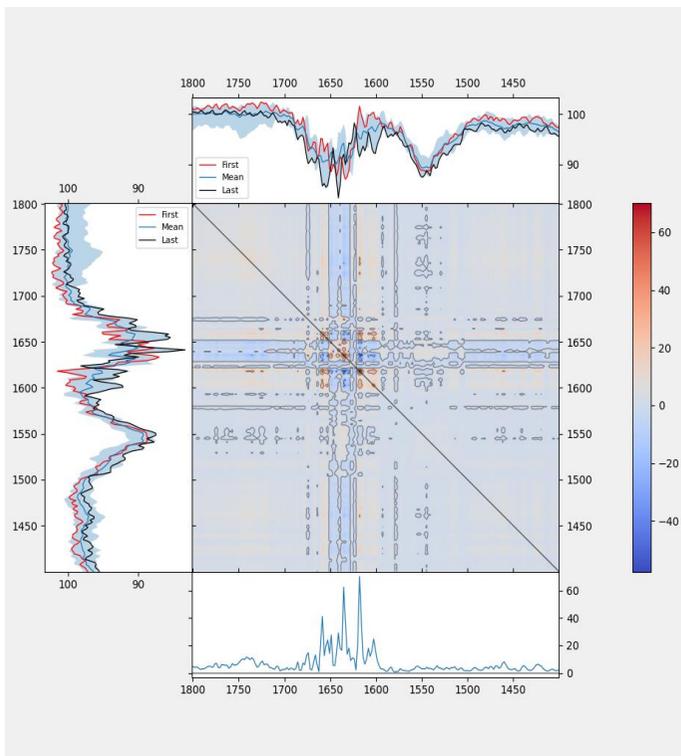


Figura 3-10 2DCOS BSA (10 mg/ml) por Días

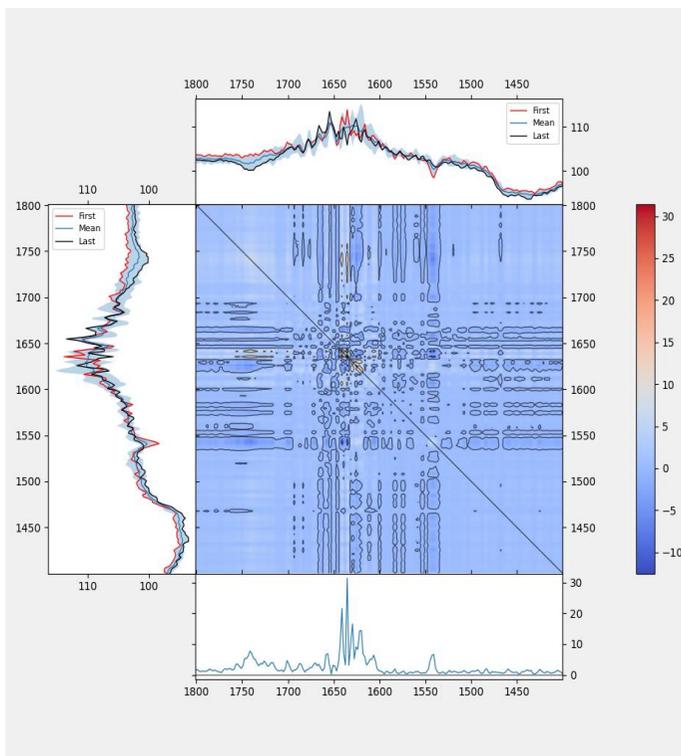


Figura 3-9 2DCOS Glucosa (20 mg/ml) por Días

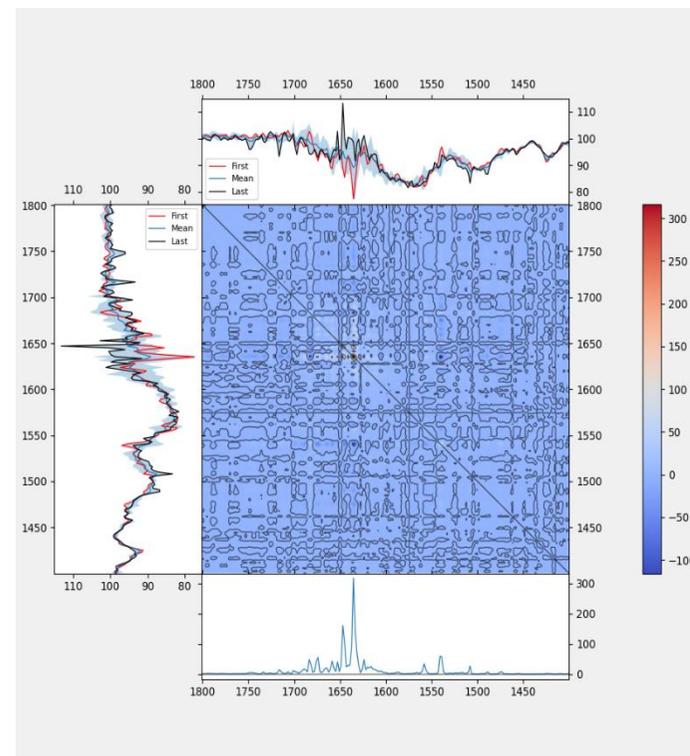


Figura 3-8 2DCOS Metformina (20 mg/ml) por Días

En las figuras Figura 3-11 y Figura 3-12 se muestran los espectros de 2DCOS generados para los sistemas binarios BSA (10 mg/ml) + glucosa (20 mg/ml) y BSA (10 mg/ml) + glucosa (40 mg/ml) a lo largo de cinco días. Se observaron cambios en los espectros perturbados por tiempo a diferentes concentraciones de glucosa.

En el espectro bidimensional de BSA (10 mg/ml) + glucosa (20 mg/ml) se observa un pico de autocorrelación en  $1650\text{ cm}^{-1}$ , asociado a la región de  $\alpha$ -hélices, con un corrimiento a un pico de deslocalización negativo hacia  $1660\text{ cm}^{-1}$ . Este desplazamiento del pico asociado a  $\alpha$ -hélices ha sido descrito como un efecto inicial de la glicación (Bohlooli et al., 2016) y como resultado de la presencia de bases de Schiff (Ioannou, 2016).

En el espectro bidimensional de BSA (10 mg/ml) + glucosa (40 mg/ml) se observan autopicos y picos deslocalizados en la región cercana a  $1620\text{ cm}^{-1}$ , lo cual está asociado a la presencia de estructuras  $\beta$  en forma de giros y hojas plegadas. También se tiene un pico que acopla las regiones  $1620$  y  $1658\text{ cm}^{-1}$ . En comparación al espectro bidimensional obtenido para el sistema unitario BSA, se tiene que con glucosa a 40 mg/ml como ligando hay mayor actividad en la zona de estructuras  $\beta$ . Según Huang et. al. (2016), la albúmina glicosilada presenta una transición conformacional desde su forma nativa de  $\alpha$ -hélices hacia hojas  $\beta$ , pasando por espirales aleatorias, y esto puede ser atribuido a un incremento de conformaciones tipo  $\beta$  causado por la formación de AGEs. Esta tendencia también fue descrita por Mou et. al. (Mou et al., 2022) para la albúmina en presencia de ribosa, que también es un azúcar con carácter reductor.

En cuanto al espectro bidimensional de BSA (10 mg/ml) + metformina (40 mg/ml) por cinco días mostrado en la figura 3-13, se obtuvo una imagen con regiones menos distinguibles, lo cual provoca ruido espectral. Se observan picos de autocorrelación y picos deslocalizados en las mismas regiones que fueron observadas para el sistema unitario BSA (10 mg/ml), pero con diferentes intensidades. Esta variación de intensidades de correlación puede atribuirse a la presencia del ligando, cuya interacción con la proteína induce cambios en los espectros IR tomados. En los gráficos IR de los paneles superior e izquierdo se identifica que los espectros inicial, promedio, y final no tienen diferencias que permitan distinguir una tendencia.

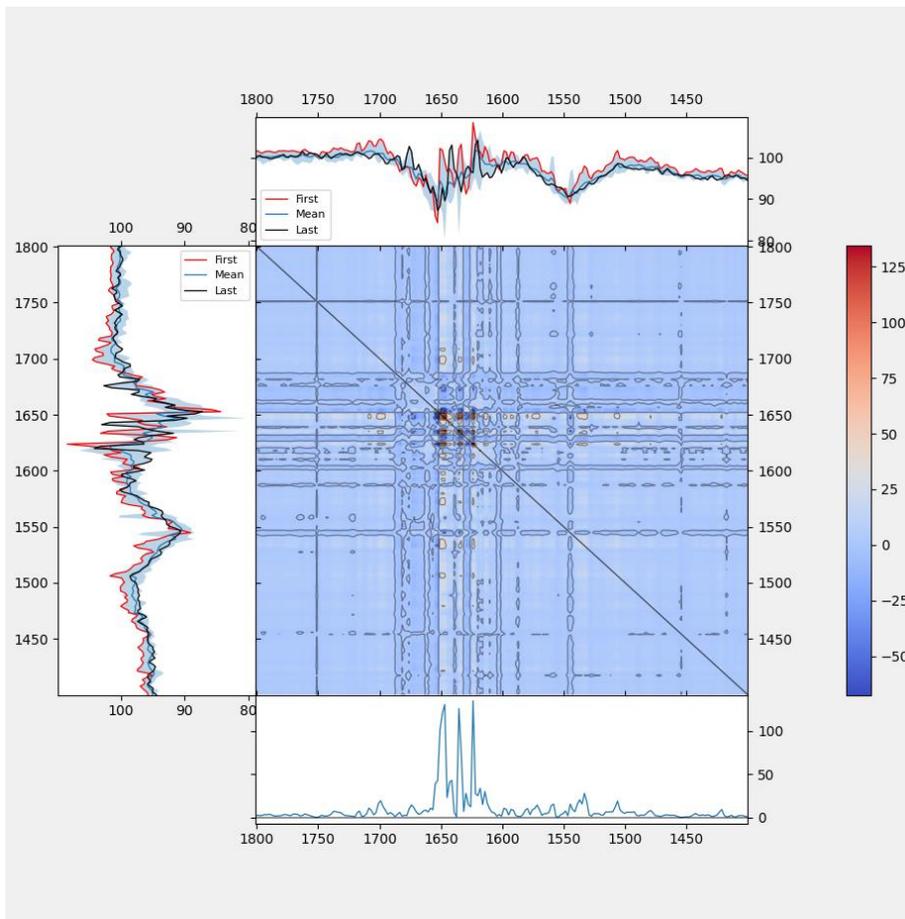


Figura 3-12 2DCOS BSA (10 mg/ml) + Glucosa (20 mg/ml) por Días

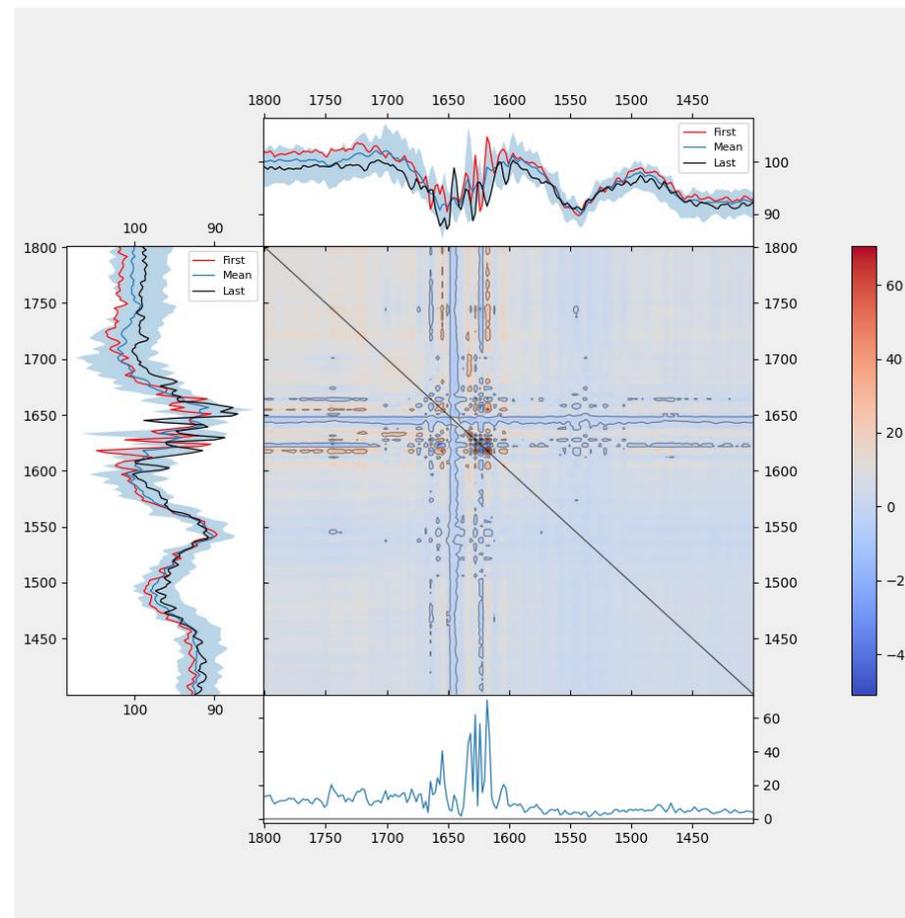


Figura 3-11 2DCOS BSA (10 mg/ml) + Glucosa (40 mg/ml) por Días

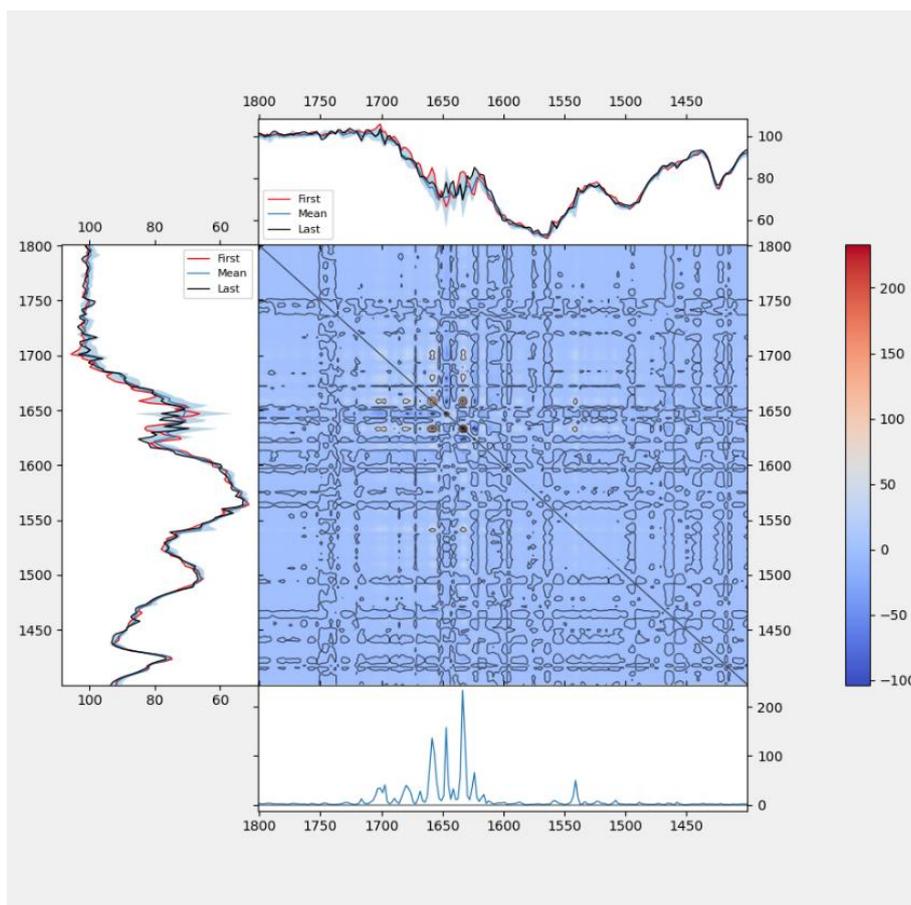


Figura 3-13 2DCOS BSA (10 mg/ml) + Metformina (40 mg/ml) por Días

Finalmente, se obtuvieron espectros de 2DCOS para los sistemas ternarios con BSA (10 mg/ml), glucosa (20 mg/ml) y metformina a las concentraciones establecidas, a lo largo de cinco días. En las figuras 3-14 y 1-15 se muestran los gráficos bidimensionales obtenidos para los sistemas ternarios con una concentración de metformina de 20 mg/ml y 40 mg/ml, respectivamente.

El espectro bidimensional obtenido para el sistema ternario con metformina a una concentración de 20 mg/ml muestra un fuerte pico en  $1620\text{ cm}^{-1}$  y un pico cercano a  $1650\text{ cm}^{-1}$ . A pesar de la posible presencia de agregados que muestra el autopico en  $1620\text{ cm}^{-1}$ , no se presenta el corrimiento del pico asociado a estructuras  $\alpha$ -hélices ni el resto de los picos de correlación asociados a estructuras  $\beta$ , que sí se observaron en el sistema binario BSA (10 mg/ml) + glucosa (20 mg/ml). Esto sugiere que el comportamiento del sistema ternario con metformina y glucosa a 20 mg/ml se asemeja más al sistema unitario BSA (10 mg/ml) que al sistema binario BSA (10 mg/ml) + glucosa (20 mg/ml).

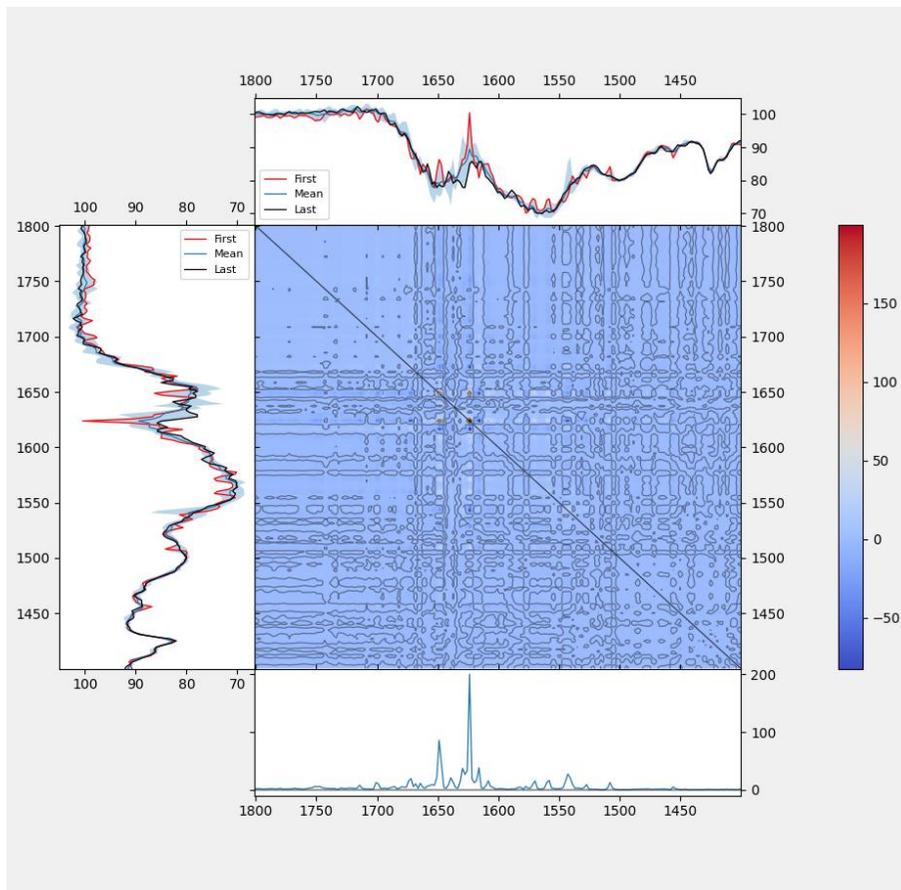


Figura 3-15 2DCOS BSA (10 mg/ml) + Glucosa (20 mg/ml) + Metformina (20 mg/ml) por Días

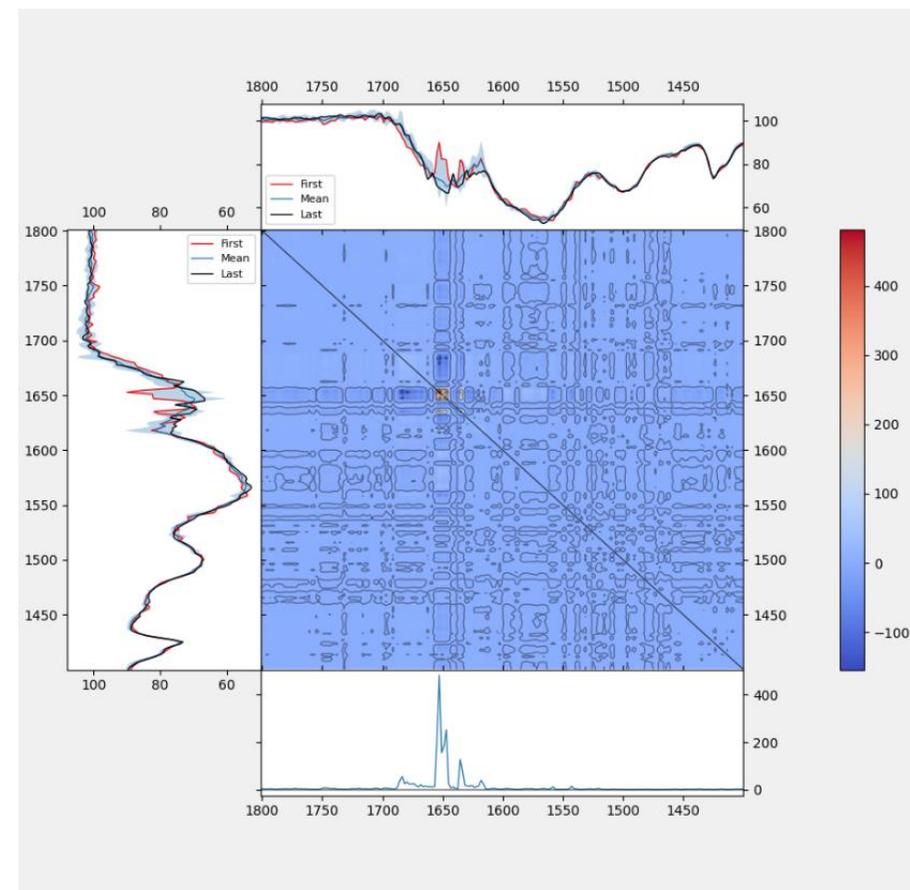


Figura 3-14 2DCOS BSA (10 mg/ml) + Glucosa (20 mg/ml) + Metformina (40 mg/ml) por Días

Por otro lado, el sistema ternario con metformina a una concentración de 40 mg/ml muestra una única región de autocorrelación cercana a  $1650\text{ cm}^{-1}$ , y no hay actividad clara en la región de  $1620\text{ cm}^{-1}$ . Este comportamiento sugiere que la presencia de metformina a 40 mg/ml en un sistema con BSA (10 mg/ml) y glucosa (20 mg/ml) retrasa el proceso de glicosilación y formación de agregados.

### **3.2.3 Resultados de 2DCOS en Sistemas Perturbados por Concentración**

Se generaron los espectros síncronos correspondientes a los sistemas perturbados por concentración. A diferencia de los sistemas perturbados por tiempo, los sistemas perturbados por concentración permiten estudiar características estructurales de las mezclas analizadas según el componente con concentración variable, mas no la dinámica del sistema.

En las figuras 3-16, 3-17 y 3-18 se muestran los sistemas unitarios de BSA (0 – 10 mg/ml), glucosa (0-40 mg/ml) y metformina (0-40 mg/ml) perturbados por concentración. El sistema unitario BSA (0 – 10 mg/ml) presenta regiones de autocorrelación correspondientes a las bandas amida I y amida II típicas de la proteína, con la particularidad de que están acopladas por picos deslocalizados positivos.

El sistema unitario glucosa (0-40 mg/ml) muestra picos de baja intensidad que se pueden considerar como pequeñas fluctuaciones causadas por el incremento de concentración. Esto se verifica en los paneles superior e izquierdo, donde la transmitancia está cercana a 100% en toda la región de observación. Por otro lado, el sistema unitario de metformina (0-40 mg/ml) sí muestra zonas de correlación en las regiones de absorción que habían sido observadas previamente en los espectros IR.

Los sistemas binarios BSA (10 mg/ml) + glucosa (0-40 mg/ml), y BSA (10 mg/ml) + Metformina (0-40 mg/ml) se muestran en las figuras 3-19 y 3-20, respectivamente. El sistema que contiene glucosa a concentración variable muestra dos picos de autocorrelación en  $1620$  y  $1630\text{ cm}^{-1}$  y varias regiones de correlación deslocalizadas. Si bien las intensidades de los picos son bajas y las muestras no se analizaron a lo largo de varios días, la autocorrelación entre  $1620\text{ cm}^{-1}$  y  $1630\text{ cm}^{-1}$  sugieren que la presencia de glucosa acelera el proceso de formación de agregados y de transición a estructuras  $\beta$ . Por otro lado, el sistema con metformina a concentración variable presenta los picos de autocorrelación y picos deslocalizados que fueron observados en el sistema unitario para la metformina. No se tiene actividad relevante en las regiones de amida I ni amida

II, lo cual indica que el aumento de concentración de metformina no tiene un impacto directo sobre las características conformacionales de la proteína.

El espectro de 2DCOS obtenido para sistema ternario con BSA (10 mg/ml), glucosa (15 mg/ml) y metformina a concentración variable de 0 a 40 mg/ml se muestra en la figura 3-21. En esta imagen se observan los picos de autocorrelación y deslocalizados característicos de la metformina, y no se tiene una correlación marcada en la región cercana a  $1620\text{ cm}^{-1}$ . Esto puede verificarse en la imagen tridimensional de la figura 3-22, donde se evidencia una depresión de los picos de correlación en las coordenadas del plano  $(1620, 1620)$ . No obstante, sí existe actividad en la región cercana a la banda de amida I ( $1650\text{ cm}^{-1}$ ). Tal comportamiento sugiere que en presencia de glucosa y metformina la proteína BSA sufre cambios estructurales, pero sin llegar a la formación de agregados ni estructuras  $\beta$ .

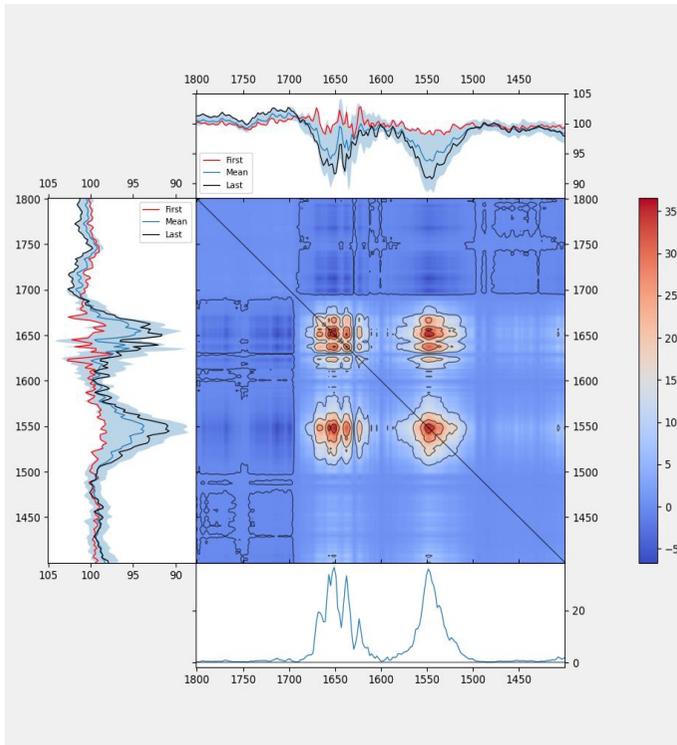


Figura 3-20 2DCOS BSA (0 – 10 mg/ml) por Concentración

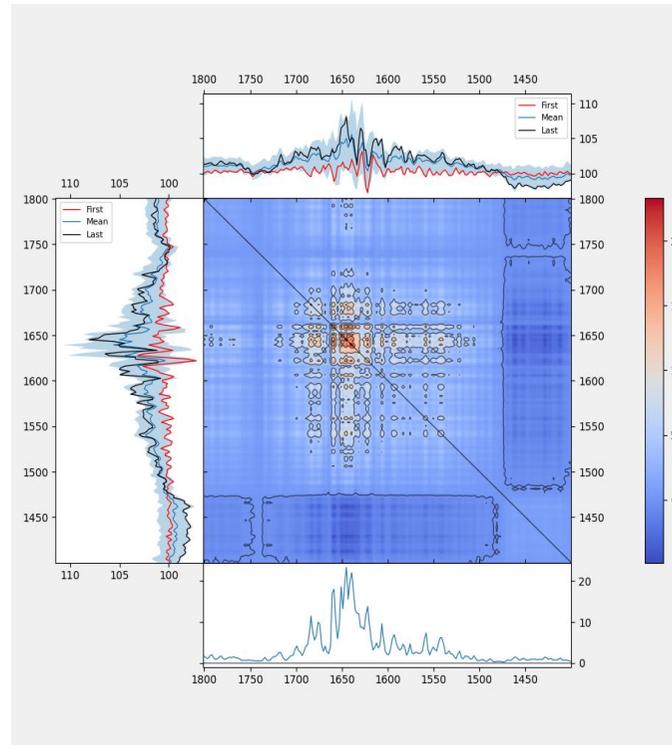


Figura 3-19 2DCOS Glucosa (0 – 40 mg/ml) por Concentración

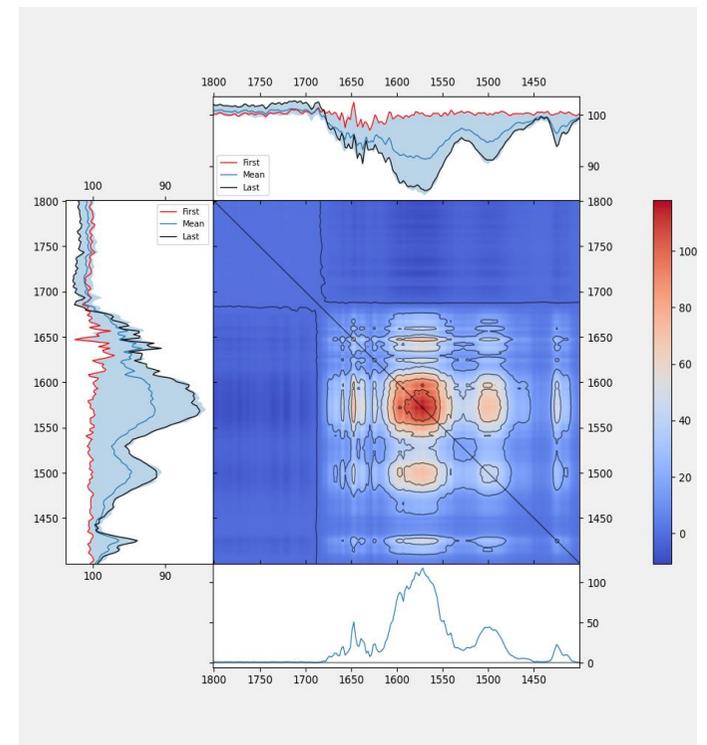


Figura 3-18 2DCOS Metformina (0 – 40 mg/ml) por Concentración

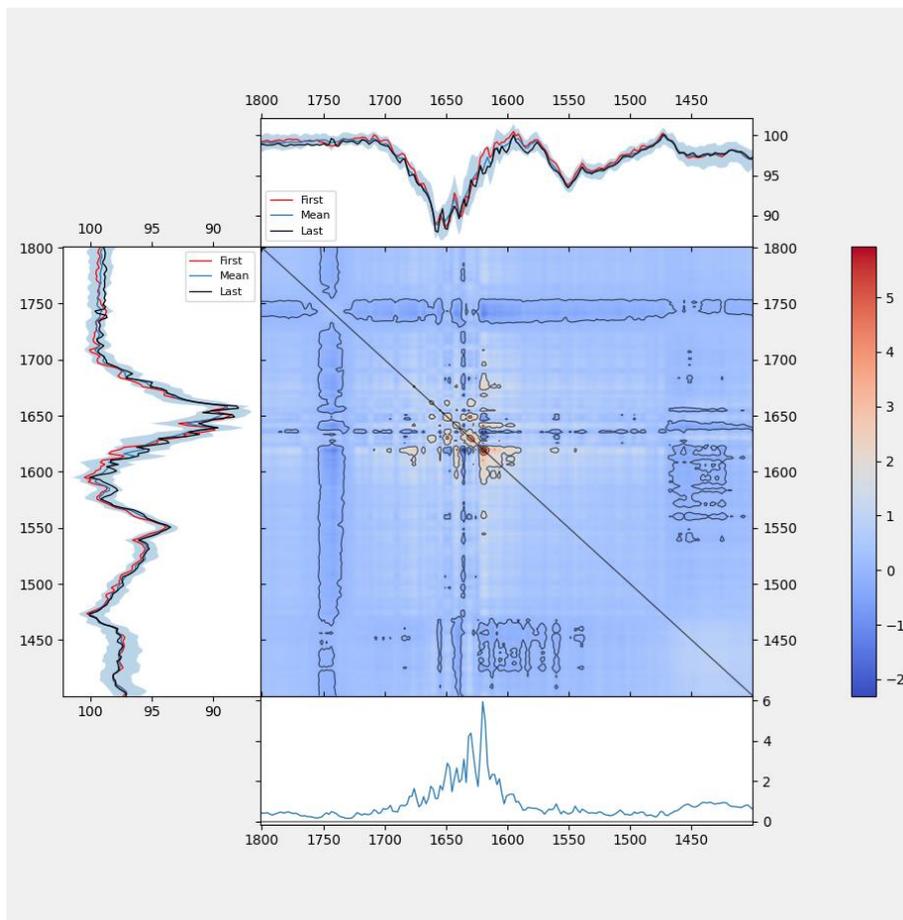


Figura 3-21 2DCOS BSA (10 mg/ml) + Glucosa (0-40 mg/ml) por Concentración

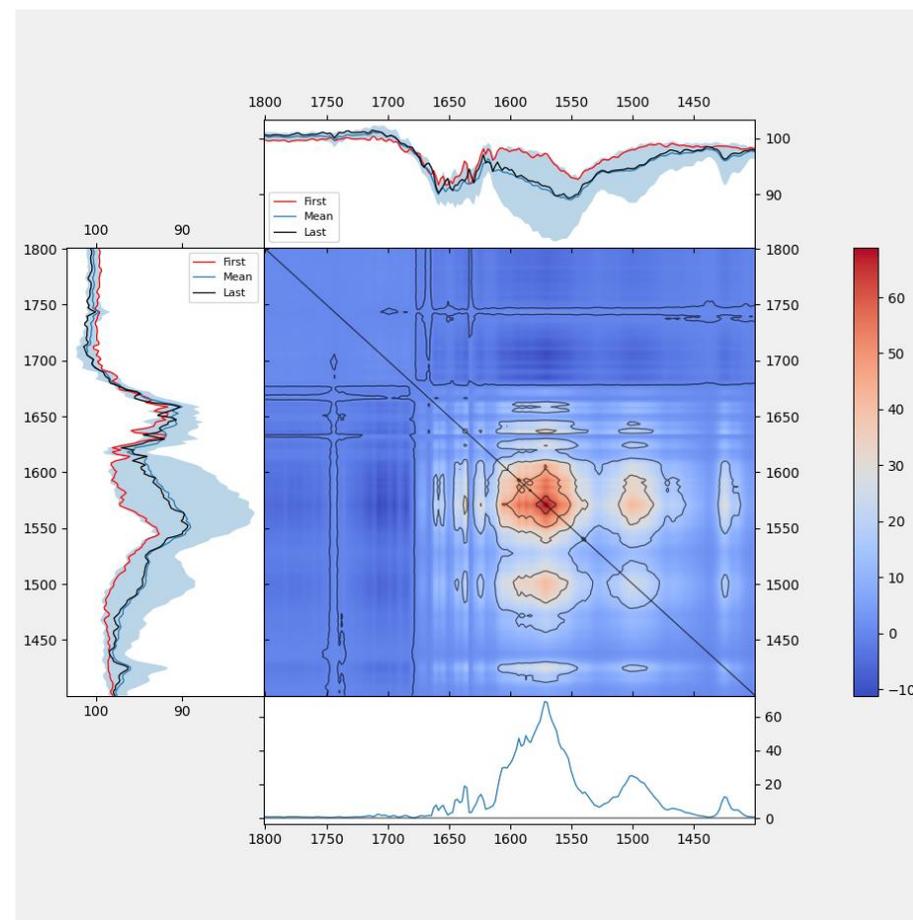


Figura 3-22 2DCOS BSA (10 mg/ml) + Metformina (0-40 mg/ml) por Concentración

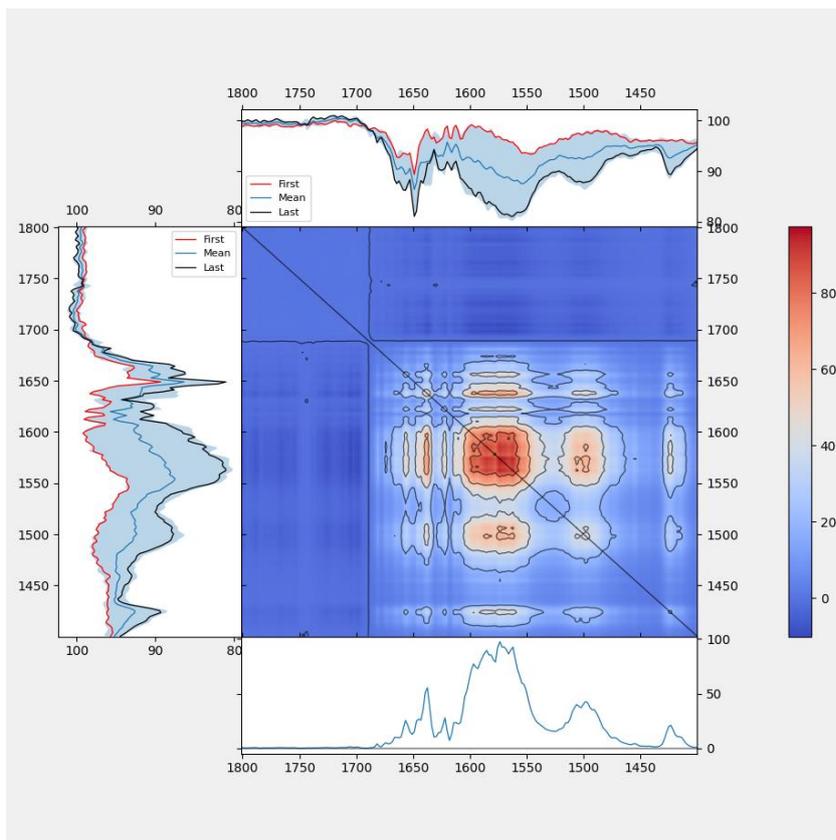


Figura 3-23 2DCOS BSA (10 mg/ml) + Glucosa (20 mg/ml) + Metformina (0-40 mg/ml) por Concentración

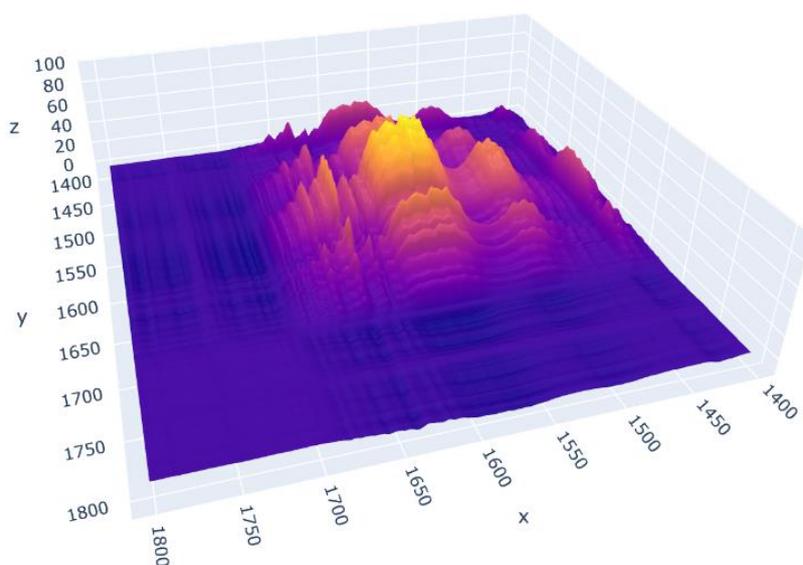


Figura 3-24 2DCOS BSA (10 mg/ml) + Glucosa (20 mg/ml) + Metformina (0-40 mg/ml) por Concentración Imagen Tridimensional

# CAPÍTULO 4

## 4. CONCLUSIONES Y RECOMENDACIONES

### Conclusiones

- Se desarrolló una aplicación de software libre en lenguaje Python que permite el análisis estructural de sistemas complejos en base a la correlación espectroscópica bidimensional. El programa genera espectros síncronos y asíncronos, y contiene funcionalidades que permiten al usuario interpretar resultados de mediciones espectroscópicas modificadas por una perturbación establecida. La aplicación desarrollada constituye un método alternativo al uso de costosos equipos avanzados de caracterización.
- Se desarrolló una interfaz interactiva acoplada a un modelo matemático que genera los gráficos síncronos y asíncronos de correlación espectroscópica bidimensional. La interfaz provee al usuario de la posibilidad de modificar múltiples ajustes de personalización, así como de manipular los espectros obtenidos para facilitar el análisis. Entre las opciones y funcionalidades disponibles para el usuario, destacan: los diferentes formatos de carga de archivo admitidos, la generación de gráficos IR adjuntos con los espectros de correlación, y los gráficos tridimensionales; opciones que no están disponibles en otros programas de correlación espectroscópica bidimensional desarrollados y mencionados en el presente proyecto.
- Se estableció un caso de estudio sobre la interacción entre albúmina sérica bovina, la glucosa y la metformina, que permitió validar la herramienta computacional desarrollada. Se tomaron mediciones de espectroscopía infrarroja para sistemas unitarios, binarios, y ternario, considerando el paso del tiempo y la variación de concentración como variables de perturbación. El tiempo permite estudiar la dinámica de cada sistema, y la variación de concentración posibilita el análisis de las características estructurales. Con las mediciones espectroscópicas obtenidas, se utiliza la aplicación desarrollada para generar espectros de correlación espectroscópica bidimensional.

- Se generaron espectros síncronos y asíncronos de correlación espectroscópica bidimensional para los sistemas establecidos en el caso de estudio sobre la interacción entre albúmina sérica bovina, glucosa y metformina. Los espectros correspondientes a las mediciones perturbadas por tiempo evidenciaron que la proteína sin ligandos sufrió cambios estructurales a lo largo de cinco días, caracterizados principalmente por la formación de agregados y pérdida de estructura secundaria propia de su estado nativo. Se observó que, en presencia de glucosa, además de los cambios descritos para el sistema unitario de proteína, hay formación de estructuras tipo  $\beta$ , que son posibles resultados de los productos finales de la glicosilación avanzada. En presencia de metformina y glucosa, los espectros obtenidos sugieren que la proteína sufrió menos cambios en su estructura. Por otra parte, los espectros que corresponden a las mediciones perturbadas por concentración permitieron verificar las particularidades estructurales de los compuestos analizados.

### **Recomendaciones**

- Analizar el sistema de albúmina sérica bovina, glucosa y metformina bajo el efecto de otras variables de perturbación (ej.: pH, temperatura) o en diferentes rangos de las variables estudiadas en este trabajo.
- Validar el uso de la correlación espectroscópica bidimensional y de la herramienta computacional desarrollada en el análisis de otros tipos de sistemas complejos, como plásticos o sistemas biológicos.
- Desarrollar casos de estudio con diferentes técnicas espectroscópicas, como UV-Vis.
- Aportar al desarrollo de la herramienta informática elaborada con la inclusión de nuevas funcionalidades, como la limpieza previa de datos.

# BIBLIOGRAFÍA

- Abrosimova, K. V., Shulenina, O. V., & Paston, S. V. (2016). FTIR study of secondary structure of bovine serum albumin and ovalbumin. *Journal of Physics: Conference Series*, 769(1). <https://doi.org/10.1088/1742-6596/769/1/012016>
- Baiz, C. R., Reppert, M., & Tokmakoff, A. (2012). *Introduction to Protein 2D IR Spectroscopy*.
- Bohlooli, M., Ghaffari-Moghaddam, M., Khajeh, M., Shahraki-Fallah, G., Haghghi-Kekhaiye, B., & Sheibani, N. (2016). The role of acetoacetate in Amadori product formation of human serum albumin. *Journal of Photochemistry and Photobiology B: Biology*, 163, 345–351. <https://doi.org/10.1016/j.jphotobiol.2016.09.004>
- Buschiazzo, A. (2022). *Latin America and Caribbean Structural Biology Landscape Analysis Report Providing the Tools Researchers Need: The Structural Biology Case Contributors and Affiliations*.
- Carvajal Carvajal, C. (2015). REVISIÓN BIBLIOGRÁFICA PRODUCTOS FINALES DE GLICACIÓN (AGEs) Y LA NEFROPATÍA DIABÉTICA. 32(1).
- Chatterley, A. S., Laity, P., Holland, C., Weidner, T., Woutersen, S., & Giubertoni, G. (2022). Broadband Multidimensional Spectroscopy Identifies the Amide II Vibrations in Silkworm Films. *Molecules*, 27(19), 6275. <https://doi.org/10.3390/molecules27196275>
- Du, X., Li, Y., Xia, Y. L., Ai, S. M., Liang, J., Sang, P., Ji, X. L., & Liu, S. Q. (2016). Insights into protein–ligand interactions: Mechanisms, models, and methods. In *International Journal of Molecular Sciences* (Vol. 17, Issue 2). MDPI AG. <https://doi.org/10.3390/ijms17020144>
- Gomaa, A. I., Sedman, J., & Ismail, A. A. (2013). An investigation of the effect of microwave treatment on the structure and unfolding pathways of  $\beta$ -lactoglobulin using FTIR spectroscopy with the application of two-dimensional correlation spectroscopy (2D-COS). *Vibrational Spectroscopy*, 65, 101–109. <https://doi.org/10.1016/j.vibspec.2012.11.019>
- Grdadolnik, J., & Chal, Y. M. (2000). *Bovine Serum Albumin Observed by Infrared Spectrometry. I. Methodology, Structural Investigation, and Water Uptake*.

- He, Z., Zhao, T., Zhou, X., Liu, Z., & Huang, H. (2017). Sequential Order of the Secondary Structure Transitions of Proteins under External Perturbations: Regenerated Silk Fibroin under Thermal Treatment. *Analytical Chemistry*, 89(10), 5534–5541. <https://doi.org/10.1021/acs.analchem.7b00592>
- Hough, M. A. (2017). Choosing the optimal spectroscopic toolkit to understand protein function. In *Bioscience Reports* (Vol. 37, Issue 3). Portland Press Ltd. <https://doi.org/10.1042/BSR20160378>
- Huang, Y. T., Liao, H. F., Wang, S. L., & Lin, S. Y. (2016). Glycation and secondary conformational changes of human serum albumin: Study of the FTIR spectroscopic curve-fitting technique. *AIMS Biophysics*, 3(2), 247–260. <https://doi.org/10.3934/biophy.2016.2.247>
- Ioannou, A. (2016). Real Time Monitoring the Maillard Reaction Intermediates by HPLC-FTIR. *Journal of Physical Chemistry & Biophysics*, 6(2). <https://doi.org/10.4172/2161-0398.1000210>
- 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>
- Lasch, P., & Noda, I. (2019). Two-Dimensional Correlation Spectroscopy (2D-COS) for Analysis of Spatially Resolved Vibrational Spectra. In *Applied Spectroscopy* (Vol. 73, Issue 4, pp. 359–379). SAGE Publications Inc. <https://doi.org/10.1177/0003702818819880>
- López-Lorente, Á. I., & Mizaikoff, B. (2016). Mid-infrared spectroscopy for protein analysis: Potential and challenges. In *Analytical and Bioanalytical Chemistry* (Vol. 408, Issue 11, pp. 2875–2889). Springer Verlag. <https://doi.org/10.1007/s00216-016-9375-5>
- Maciążek-Jurczyk, M., Janas, K., Pożycka, J., Szkudlarek, A., Rogóż, W., Owczarzy, A., & Kulig, K. (2020). Human serum albumin aggregation/fibrillation and its abilities to drugs binding. *Molecules*, 25(3). <https://doi.org/10.3390/molecules25030618>
- Magalhães, S., Trindade, D., Martins, T., Martins Rosa, I., Delgadillo, I., Goodfellow, B. J., da Cruz e Silva, O. A. B., Henriques, A. G., & Nunes, A. (2020). Monitoring plasma protein aggregation during aging using conformation-specific antibodies and FTIR

- spectroscopy. *Clinica Chimica Acta*, 502, 25–33.  
<https://doi.org/10.1016/j.cca.2019.11.025>
- Mou, L., Hu, P., Cao, X., Chen, Y., Xu, Y., He, T., Wei, Y., & He, R. (2022). Comparison of bovine serum albumin glycation by ribose and fructose in vitro and in vivo. *Biochimica et Biophysica Acta - Molecular Basis of Disease*, 1868(1).  
<https://doi.org/10.1016/j.bbadis.2021.166283>
- Nelson, D. L., & Cox, M. M. (2004). *Lehninger Principles of Biochemistry* (Freeman, Ed.; 4th Edition).
- Noda, I. (2016). 2DCOS and I. Three decades of two-dimensional correlation spectroscopy. *Journal of Molecular Structure*, 1124, 3–7.  
<https://doi.org/10.1016/j.molstruc.2016.01.035>
- Noda, I. (2017). Vibrational two-dimensional correlation spectroscopy (2DCOS) study of proteins. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, 187, 119–129. <https://doi.org/10.1016/j.saa.2017.06.034>
- Noda, I. (2022). Two-trace two-dimensional (2T2D) correlation applied to a number of spectra beyond a simple pair. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, 277. <https://doi.org/10.1016/j.saa.2022.121258>
- OriginLab Technical Support. (2018). *2D Correlation Spectroscopy Analysis* (1.22).
- Peter Lasch. (2023). *mat2dcorr - 2D correlation spectroscopy toolbox with Matlab* (1.04).
- Rahman, N., & Khalil, N. (2022). Effect of glycation of bovine serum albumin on the interaction with xanthine oxidase inhibitor allopurinol: Spectroscopic studies and molecular modeling. *Journal of Molecular Liquids*, 367.  
<https://doi.org/10.1016/j.molliq.2022.120396>
- Rutherford, S. H., Greetham, G. M., Towrie, M., Parker, A. W., Kharratian, S., Krauss, T. F., Nordon, A., Baker, M. J., & Hunt, N. T. (2022). Detection of paracetamol binding to albumin in blood serum using 2D-IR spectroscopy. *Analyst*, 147(15), 3464–3469.  
<https://doi.org/10.1039/d2an00978a>
- Sabbagh, B. A., Kumar, P. V., Chew, Y. L., Chin, J. H., & Akowuah, G. A. (2022). Determination of metformin in fixed-dose combination tablets by ATR-FTIR spectroscopy. *Chemical Data Collections*, 39, 100868.  
<https://doi.org/10.1016/j.cdc.2022.100868>
- Shigeaki, M., & Yukihiro, O. (2005). *2DShige* (1.3). Kwansai-Gakuin University, Japan.

- Skoog, D., Holler, F. J., & Crouch, S. (2018). *Principles of Instrumental Analysis* (7th ed.). Cengage Learning.
- Tao, Y., Wu, Y., & Zhang, L. (2018). Advancements of two dimensional correlation spectroscopy in protein researches. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 197, 185–193. <https://doi.org/10.1016/j.saa.2018.01.039>
- Tunyasuvunakool, K., Adler, J., Wu, Z., Green, T., Zielinski, M., Židek, A., Bridgland, A., Cowie, A., Meyer, C., Laydon, A., Velankar, S., Kleywegt, G. J., Bateman, A., Evans, R., Pritzel, A., Figurnov, M., Ronneberger, O., Bates, R., Kohl, S. A. A., ... Hassabis, D. (2021). Highly accurate protein structure prediction for the human proteome. *Nature*, 596(7873), 590–596. <https://doi.org/10.1038/s41586-021-03828-1>
- Varadi, M., Bordin, N., Orengo, C., & Velankar, S. (2023). The opportunities and challenges posed by the new generation of deep learning-based protein structure predictors. In *Current Opinion in Structural Biology* (Vol. 79). Elsevier Ltd. <https://doi.org/10.1016/j.sbi.2023.102543>
- Wade, L., & Simek, J. (2017). *Organic Chemistry* (9th ed.). Pearson.
- Whitford, D. (2005). *Proteins: Structure and Function* (1st ed.). John Wiley & Sons.
- Wilkosz, N., Czaja, M., Seweryn, S., Skirlinska-Nosek, K., Szymonski, M., Lipiec, E., & Sofinska, K. (2020). Molecular spectroscopic markers of abnormal protein aggregation. In *Molecules* (Vol. 25, Issue 11). MDPI AG. <https://doi.org/10.3390/molecules25112498>
- www.labx.com. (2023, June 17). *FT-IR Listings*. LabX.
- Xue, B., Xie, L., & Zhang, J. (2017). Detailed molecular movements during poly(l-lactic acid) cold-crystallization investigated by FTIR spectroscopy combined with two-dimensional correlation analysis. *RSC Adv.*, 7(74), 47017–47028. <https://doi.org/10.1039/C7RA08921J>
- Yang, H., Xu, S., Shen, L., Liu, W., & Li, G. (2016). Changes in aggregation behavior of collagen molecules in solution with varying concentrations of acetic acid. *International Journal of Biological Macromolecules*, 92, 581–586. <https://doi.org/10.1016/j.ijbiomac.2016.07.080>