



**Universidad
Nacional del
Litoral**

2016

**Tesis
Doctoral
Química**

Alcaráz Mirta Raquel

FIQ

UNL

“Nuevas estrategias analíticas para el análisis de muestras de composición compleja basadas en generación y modelado de datos multidimensionales”

Alcaráz, Mirta Raquel

Licenciada en Biotecnología

*Tesis presentada como parte de los requisitos de la Facultad de Ingeniería Química de la Universidad Nacional del Litoral para la obtención del Grado Académico de **Doctor en Química**, en el campo de la Química Analítica. Trabajo realizado en el Laboratorio de Desarrollo Analítico y Quimiometría de la Facultad de Bioquímica y Ciencias Biológicas de la Universidad Nacional del Litoral, desarrollado en el periodo comprendido entre Abril de 2012 y Octubre de 2016, bajo la dirección del Dr. Héctor C. Goicoechea.*

~2016~



**Universidad Nacional Del Litoral
Facultad de Ingeniería Química**

“How wonderful it is that nobody need wait a single moment *before starting to improve the world*”

The Diary of a Young Girl – Anne Frank

AGRADECIMIENTOS

En una entrevista a un reconocido músico argentino escuché una de esas frases que le quedan a uno dando vueltas en la cabeza. Ante la pregunta de qué significa la música para él, responde: *“La música es, realmente, una expedición [...] Viste que escuchas por ahí: “No, no me gusta esta música”.....No, no es que no te gusta.... Es que no tenés recursos para comprenderla”*. Y entre tanto pensarlo, puedo decir que es algo aplicable a varios aspectos de la vida, pero, principalmente, a la vida que uno forma como profesional. Y el doctorado, en mi opinión, es exactamente eso: *“una expedición”*. Una expedición voluntaria en búsqueda de nuevos conocimientos, en búsqueda de recursos que permitan reconciliarnos con los conceptos adquiridos en nuestra formación anterior y en búsqueda, principalmente, de aquello que nos permitirá disfrutar y degustar todos los días de lo que hacemos.

Comencé la carrera de doctorado por intención y la concreté por elección. Y en esta expedición me encontré con personas que me han acompañado y ayudado a transitar el camino y a quienes no puedo dejar de agradecer. Es por ello que aprovecho este espacio para expresar, humildemente, mi más sincero agradecimiento.

A Héctor Goicoechea, mi director, por haber confiado en mí, por darme la libertad y el apoyo en la construcción de mi formación profesional. Pero, principalmente, por sus consejos y sus charlas en los momentos precisos en que me sentí superada por la búsqueda.

A todos mis compañeros del Laboratorio de Desarrollo Analítico y Quimiometría y de la Cátedra de Química Analítica I por lo compartido, los buenos momentos y, también, por los vaivenes propios de la convivencia. A Gabriel, por sus profundas reflexiones, por su amistad, por los almuerzos colmados de buena energía, por ser partícipe de buenos momentos. A Julia, quien me ha ayudado y me ha brindado su amistad desde el principio y quien me demuestra que, a pesar de que la vorágine de la vida

nos lleve por caminos diferentes, siempre hay un espacio donde encontrarnos. A Romina, por su locura hermosa, por su apoyo y cariño en los momentos compartidos dentro y fuera del laboratorio. A todos con los que he compartido el día a día y a quienes han venido desde otro lugar ofreciendo su compañerismo y amistad. A todos ellos mi eterna gratitud.

A todos los integrantes del Laboratorio de Control de Calidad de Medicamentos, particularmente a Mercedes y Marisi, que siempre están dispuestas a brindar una ayuda, una sonrisa o una charla amena.

A mi incondicional Milagros, compañera de oficina, compañera de vida, cursos, viajes y locuras. Por compartir conversaciones, la mayoría incoherentes o sin sentido, por compartir risas y emociones, por los recuerdos hermosos que quedarán en la memoria, por ese apoyo mutuo, siempre tan necesario.

A Cosima, a quien le estaré agradecida siempre por su compañía invaluable, quien me ha brindado su amistad con apenas conocerme. A Andreas, una persona maravillosa que no solo me brindó su amistad, sino también su desinteresada e incondicional ayuda. A todo el grupo del Institute of Chemical Technologies and Analytics de la Technische Universität Wien.

A Silvana, mi pampeana divina, que siempre estuvo y está presente regalándole una sonrisa al mundo.

A Yamile, persona hermosa, transparente e incondicional que supo escucharme y entenderme. Por todos los momentos en que necesité su ayuda y comprensión. Gracias eternas por ser la persona que necesito en mi vida, con quien puedo hablar de absolutamente todo.

Mi gratitud hacia Germán por sus mensajes a toda hora, por sus comentarios acertados en el momento justo. Simplemente, gracias por nuestra maravillosa y alocada amistad. *“Lo mejor será que bailemos. - ¿Y que nos juzguen de locos, Sr. Conejo? - ¿Usted conoce cuerdos felices? - Tiene usted razón, bailemos.”*

A Yamile, Matías e Iñaki por formar parte de mi vida, por hacer que todos los fines de semana tengan un condimento especial.

A Bernhard por haber sido partícipe de una de las mejores experiencias que he vivido. Por brindarme la oportunidad que, en cierto modo, me hizo cambiar algunas ideologías y me ha hecho crecer casi coercitivamente.

A Victor, quien me ha dado la primera oportunidad y me abrió las puertas de la cátedra de Química Analítica I haciéndome partícipe de su grupo.

A Gabi, (Árbol para los conocidos) mi hermana, la pequeña, por todo lo que me

ha brindado, por hacerme reír todos los días, por su grata y hermosa compañía.

A Beto, mi hermano, y su hermosa familia, Fiorela y Renata, por ser mi compañero de infancia y por regalarme los mejores recuerdos.

A mis padres, Mario e Imelda, que me han brindado y me brindan su amor, quienes me enseñaron y me inculcaron los principios que aún llevo en mi interior.

A la familia que adopte y con quienes me maravillo de la vida todos los días. Gracias Adriana por tu transparencia y tu ingenuidad, gracias Panchito por tu cariño y cuidado paternal.

A Tito, la persona que elegí para compartir mi vida, gracias por sumergirme en un mundo de fantasías lleno de realidad. Por hacerme feliz todos los días, con quien encuentro la calma. Por darme la libertad de ser quien soy, pero, también, por haberme elegido. Por elegirnos para compartir el maravilloso momento que estamos transitando, por proyectar juntos y formar la hermosa familia que tenemos (junto a Gato-Pato).

A todos los docentes que han ofrecido su aporte y han ayudado al crecimiento y formación de profesionales. A la Facultad de Bioquímica y Ciencias Biológicas y a la Facultad de Ingeniería Química, de la Universidad Nacional del Litoral, que logran formar profesionales de prestigio. A la Technische Universität Wien por brindarme el espacio de trabajo en mi estadía en Viena. Al Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) por el apoyo recibido para el desarrollo de la tesis y mi formación.

Gracias a todos los que directa o indirectamente agregaron su aporte en mi camino de crecimiento.

*“Hay hombres que de su ciencia
Tienen la cabeza llena;
Hay sabios de todas menas,
Más digo sin ser muy ducho,
Es mejor que aprender mucho
El aprender cosas buenas”*

Martin Fierro

PUBLICACIONES CIENTÍFICAS

External cavity-quantum cascade laser infrared spectroscopy for secondary structure analysis of proteins at low concentration – **Schwaighofer A., Alcaráz M.R., Goicoechea H.C., Lendl B.** *Sci Rep*, 2016, **6**, 33556-33566

EC-QCL mid-IR transmission spectroscopy for monitoring dynamic changes of protein secondary structure in aqueous solution on the example of β -aggregation in alcohol-denaturated α -chymotrypsin – **Alcaráz M.R., Schwaighofer A., Goicoechea H.C., Lendl B.** *Anal Bioanal Chem* 2016, **408**, 3933–3941

Enhanced fluorescence sensitivity by coupling yttrium-analyte complexes and three-way fast high-performance liquid chromatography data modelling – **Alcaráz, M.R., Culzoni M.J., Goicoechea H.C.** *Anal Chim Acta* 2016, **902**, 50–58

External cavity-quantum cascade laser spectroscopy for mid-IR transmission measurements of proteins in aqueous solution – **Alcaráz M.R., Schwaighofer A., Kristament C., Ramer G., Brandstetter M., Goicoechea H., Lendl B.** *Anal Chem* 2015, **87**, 6980–6987

A new modelling strategy for third-order fast high-performance liquid chromatographic data with fluorescence detection. Quantitation of fluoroquinolones in water samples – **Alcaráz, M.R., Bortolato S.A., Goicoechea H.C., Olivieri A.C.** *Anal Bioanal Chem* 2015, **407**, 1999–2011

Ultrafast quantitation of six quinolones in water samples by second-order capillary electrophoresis data modeling with multivariate curve resolution–alternating least squares – **Alcaráz, M.R., Vera-Candioti L., Culzoni M.J., Goicoechea H.C.** *Anal Bioanal Chem* 2014, **406**, 2571–2580

Modelling four and three-way fast high-performance liquid chromatography with fluorescence detection data for quantitation of fluoroquinolones in water samples – Alcaráz, M.R., Siano G.G, Culzoni M.J., Goicoechea H.C. Anal Chim Acta 2014, 809, 37–46

Modeling of second-order spectrophotometric data generated by a pH-gradient flow injection technique for the determination of doxorubicin in human plasma –Alcaráz, M.R., Schenone A.V., Culzoni M.J., Goicoechea H.C. Microchem J 2014, 112, 25–33

*Application of MCR-ALS to reveal intermediate conformations in the thermally induced α - β transition of poly-L-lysine monitored by FT-IR spectroscopy. Alcaráz M.R., Schwaighofer A., Goicoechea H.C., Lendl B. *Enviado para revisión y posterior publicación a Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy**

Chapter 7: Multiway calibration approaches to handle problems linked to the determination of emergent contaminants in waters. En: Current applications of chemometrics. Alcaráz M.R., Brasca R., Cámara M.S., Culzoni M.J., Schenone A.V., Teglia C.M., Vera Candiotti L., Goicoechea H.C. Nova Science Publishers, Inc. 2014, pp.135–154

RESUMEN

La Quimiometría se puede definir como una disciplina química que utiliza métodos matemáticos y estadísticos empleando la lógica formal para diseñar o seleccionar procedimientos de medición óptimos o experimentos, para proveer la máxima información química relevante mediante el análisis de datos y para obtener información acerca del sistema químico bajo estudio. Los procedimientos matemáticos para el procesamiento de datos multidimensionales son algoritmos que permiten el análisis de los datos en base a un determinado modelo matemático.

En este trabajo se evaluaron diferentes estrategias de obtención de datos multidimensionales y se exploraron las potencialidades de los algoritmos quimiométricos para el análisis de muestras complejas de composición variada. Se desarrollaron nuevos métodos analíticos para estudios descriptivos y predictivos de sistemas químicos de diferente naturaleza, combinando diversas técnicas analíticas con modelado quimiométrico de datos multidimensionales.

Se desarrollaron diferentes métodos analíticos para la determinación de contaminantes emergentes a nivel de trazas en muestras agua de recursos naturales, incluyendo un sistema basado en complejos analito-itrio con resolución cromatográfica y detección de fluorescencia de barrido rápido, y un método basado en electroforesis capilar zonal con detección DAD-UV y extracción en fase sólida como etapa de pre-concentración, ambos acoplados a herramientas de calibración multivariada. Las metodologías utilizadas para el incremento de la intensidad de la señal de los analitos evaluados, promovido por la presencia de itrio o por la etapa de pre-concentración, demostraron ser una simple pero eficiente manera de lograr excelentes valores de sensibilidad, permitiendo la detección y cuantificación de analitos en muestras complejas a nivel de trazas.

Además, se propuso un método simple basado en un análisis por inyección en

flujo y detección de absorbancia visible utilizando un detector de arreglo de diodos para la determinación de doxorubicina en plasma humano, obteniéndose datos de segundo orden por generación de gradiente de pH, para la cuantificación del analito y la obtención de información cualitativa acerca sistema en estudio.

La generación de datos de segundo orden demostró ser una excelente estrategia para lograr la resolución y cuantificación de los compuestos de interés en muestras complejas o sistemas multi-componentes, aprovechándose la ventaja de segundo orden que presentan las técnicas quimiométricas de calibración multivariada. El uso de algoritmos quimiométricos, como MCR-ALS, para modelar datos de segundo orden permitió resolver problemas de análisis de mezclas, incluso en presencia de componentes con señales altamente solapadas.

Por otro lado, se propuso el desarrollo de una nueva metodología infrarroja utilizando un láser EC-QCL como fuente lumínica para mediciones de transmisión mid-IR de proteínas en medio acuoso y medio deuterado. Para el tratamiento de señales obtenidas, se desarrolló una estrategia de procesamiento basada en algoritmos quimiométricos. Adicionalmente, se estudió el efecto de la presencia de alcohol sobre una proteína y el efecto de la temperatura sobre un polipéptido modelo, registrándose espectros mid-IR en un determinado periodo de tiempo y temperatura, respectivamente. Con el objetivo de obtener resultados con interpretación física, se seleccionó MCR-ALS como algoritmo quimiométrico para el modelado de los datos multidimensionales obtenidos. Asimismo, se evaluó el efecto de la temperatura utilizando dicroísmo circular y análisis de datos por MCR-ALS.

Finalmente, se analizaron 3 estrategias o metodologías analíticas para la generación de datos de tercer orden, utilizando diferentes arreglos instrumentales, basados en un procedimiento cromatográfico acoplado a la detección de matrices de excitación-emisión de fluorescencia. Se evaluaron los requerimientos instrumentales para cada metodología individual en virtud de la complejidad, los costos y la accesibilidad. Adicionalmente, se realizó un análisis detallado de los datos generados, evaluándose sus propiedades desde el punto de vista de trilinealidad/cuadrilinealidad y complejidad de análisis en el pre-procesamiento y procesamiento de datos.

ÍNDICE GENERAL

CAPÍTULO 1	1.1
1. INTRODUCCIÓN	1.2
1.1. LA QUIMIOMETRÍA COMO HERRAMIENTA DE ANÁLISIS	1.2
1.2. CLASIFICACIÓN DE DATOS INSTRUMENTALES	1.3
<i>Datos de orden cero</i>	1.4
<i>Datos de primer orden</i>	1.5
<i>Datos de segundo orden</i>	1.5
<i>Datos de tercer orden</i>	1.6
1.3. PROPIEDADES DE LOS DATOS MULTIDIMENSIONALES.....	1.6
<i>Bilinealidad</i>	1.6
<i>Trilinealidad</i>	1.7
1.4. GENERACIÓN DE DATOS MULTIDIMENSIONALES.....	1.9
<i>Matrices de excitación-emisión de fluorescencia</i>	1.9
<i>Métodos separativos con detección espectral</i>	1.11
<i>Monitoreo de reacciones con detección espectral</i>	1.12
1.5. ALGORITMOS QUIMIOMÉTRICOS	1.13
<i>Resolución multivariada de curvas-Cuadrados mínimos alternantes (MCR-ALS)</i> ..	1.15
<i>Análisis paralelo de factores (PARAFAC)</i>	1.18
<i>Análisis paralelo de factores de 3-vías aumentado (APARAFAC)</i>	1.21
<i>Otras herramientas quimométricas y de análisis</i>	1.22
1.6. CALIBRACIÓN MULTIVARIADA	1.25
<i>Conceptos básicos</i>	1.25
<i>Cifras de mérito analíticas</i>	1.27
CAPÍTULO 2	1.32
2. OBJETIVOS	2.33
2.1. OBJETIVOS GENERALES	2.33
2.2. OBJETIVOS ESPECÍFICOS.....	2.34
CAPÍTULO 3	2.35
3. OBTENCIÓN Y ANÁLISIS DE DATOS – APLICACIONES	3.36
3.1. MÉTODO CROMATOGRÁFICO PARA LA DETERMINACIÓN DE FLUOROQUINOLONAS EN MEDIO ACUOSO BASADO EN LA FORMACIÓN DE COMPLEJOS METÁLICOS Y MODELADO DE DATOS MULTIDIMENSIONALES	3.36
<i>Consideraciones generales</i>	3.37
<i>Materiales y métodos</i>	3.40
<i>Resultados</i>	3.44

<i>Conclusión</i>	3.57
3. 2. MÉTODO ELECTROFORÉTICO PARA LA DETERMINACIÓN DE QUINOLONAS EN AGUA DE CONSUMO CON MODELADO QUIMIOMÉTRICO DE DATOS MULTIDIMENSIONALES	3.59
<i>Consideraciones generales</i>	3.59
<i>Materiales y métodos</i>	3.62
<i>Resultados</i>	3.66
<i>Conclusión</i>	3.79
3. 3. MÉTODO DE GRADIENTE DE PH POR SISTEMA DE INYECCIÓN EN FLUJO ACOPLADO A MODELADO DE DATOS DE SEGUNDO ORDEN PARA EL ANÁLISIS DE DOXORRUBICINA EN PLASMA ...	3.81
<i>Consideraciones generales</i>	3.81
<i>Materiales y métodos</i>	3.83
<i>Resultados</i>	3.86
<i>Conclusión</i>	3.99
3. 4. DESARROLLO DE UNA MÉTODO PARA EL ANÁLISIS ESTRUCTURAL DE PROTEÍNAS EN MEDIO ACUOSO BASADO EN ESPECTROSCOPIA INFRARROJA LASER COMBINADA A HERRAMIENTAS QUIMIOMÉTRICAS PARA PROCESAMIENTO Y ANÁLISIS DE DATOS.....	3.100
<i>Consideraciones generales</i>	3.100
<i>Materiales y métodos</i>	3.104
<i>Resultados</i>	3.107
<i>Conclusión</i>	3.123
3. 5. MONITOREO Y ANÁLISIS ESTRUCTURAL DE CAMBIOS CONFORMACIONALES INDUCIDOS POR TEMPERATURA DE UN HOMO-POLIPÉPTIDO MEDIANTE DICROÍSMO CIRCULAR, FT-IR Y EC-QCL-IR CON RESOLUCIÓN QUIMIOMÉTRICA DE DATOS.....	3.125
<i>Consideraciones generales</i>	3.125
<i>Materiales y métodos</i>	3.128
<i>Resultados</i>	3.132
<i>Conclusión</i>	3.145
3. 6. OBTENCIÓN DE DATOS DE TERCER ORDEN UTILIZANDO CROMATOGRAFÍA LÍQUIDA DE ALTO RENDIMIENTO ACOPLADA A DETECCIÓN CON MATRICES EXCITACIÓN-EMISIÓN DE FLUORESCENCIA. ANÁLISIS Y RESOLUCIÓN DE DATOS	3.146
<i>Consideraciones generales</i>	3.146
<i>Materiales y métodos</i>	3.148
<i>Resultados</i>	3.152
<i>Conclusión</i>	3.173
CAPÍTULO 4	3.177
4. CONCLUSIONES GENERALES	4.178
CAPÍTULO 5	4.181
5. REFERENCIAS	5.182

ÍNDICE DE FIGURAS

Figura 1.1. Representación esquemática de la clasificación de los datos instrumentales según el “orden” de los datos obtenidos para una muestra única (arriba) y las “vías” de un arreglo multidimensional para múltiples muestras (abajo).....	1.4
Figura 1.2. Matriz de excitación-emisión de fluorescencia obtenida para una solución de sarafloxacin en matriz acuosa. En celeste se muestra el espectro de excitación obtenido utilizando una longitud de onda de emisión de 450 nm y en rosa se muestra el espectro de emisión obtenido utilizando una longitud de onda de excitación de 270 nm.....	1.10
Figura 1.3. Modelo de descomposición MCR-ALS para datos de segundo orden dispuestos en un arreglo de 2-vías aumentado para un sistema de múltiples muestras conteniendo N constituyentes.....	1.18
Figura 1.4. Modelo de descomposición PARAFAC para datos de segundo orden dispuestos en un arreglo de 3-vías para un sistema de múltiples muestras conteniendo N constituyentes.....	1.20
Figura 1.5. Modelo de descomposición APARAFAC para datos de tercer orden dispuestos en un arreglo de 3-vías aumentado para un sistema de múltiples muestras conteniendo N constituyentes.....	1.22
Figura 3.1. Estructura química general de las fluoroquinolonas.....	3.37
Figura 3.2. Modelo de coordinación propuesto de la formación del complejo metálico entre Y^{3+} y CPF.....	3.39
Figura 3.3. Espectros de excitación y emisión de fluorescencia normalizados de SRF en presencia (línea de puntos) y ausencia (línea continua) de Y^{3+} en una solución reguladora de NaAc/HAc pH 4.00.....	3.45
Figura 3.4. Señal cromatográfica obtenida para una solución conteniendo 60.0 ng mL^{-1} de SRF en ausencia (línea de puntos) y presencia (línea continua) de $1 \times 10^{-4} \text{ mol L}^{-1}$ de Y^{3+} . En el recuadro se amplifica la señal obtenida para SRF en ausencia de Y^{3+} . Detección: $\lambda_{\text{ex}} = 280.0 \text{ nm}$; $\lambda_{\text{em}} = 440.0 \text{ nm}$	3.46
Figura 3.5. Espectros de emisión de fluorescencia normalizados de ENO (rojo), CPF (línea azul continua), NRF (línea de rayas y puntos azul), OFL (celeste), ENF (amarillo), SRF (línea continua verde) y DIF (línea de rayas y puntos verde) en presencia de Y^{3+} en solución reguladora de NaAc/HAc pH 4.0.....	3.48

- Figura 3.6.** Cromatograma obtenido para una muestra conteniendo las 7 f-QUI en presencia de Y^{3+} , en el que se muestran las regiones seleccionadas para el modelado quimiométrico. Región 0 (gris): frente de corrida, sin presencia de analitos; Región I (rosa): (1) ENO, (2) NRF, (3) OFL y (4) CPF; Región II (azul): (5) ENF, (6) SRF y (7) DIF3.49
- Figura 3.7.** Perfiles temporales (A) y espectrales (B) de las 7 f-QUI analizadas obtenidos por modelado quimiométrico con MCR-ALS de cada matriz D_{aum} . La Región I contiene los analitos ENO (rojo), CPF/NRF (azul) y OFL (celeste) y la Región II los analitos ENF (amarillo) y SRF/DIF (verde)3.51
- Figura 3.8.** Perfiles temporales obtenidos por modelado con MCR-ALS para una de las muestras de validación analizada. El perfil de CPF y NRF, presente en la Región I, se indica en azul, mientras que el obtenido para SRF y DIF se indica en verde en la Región II. Las flechas rojas señalan el punto de división seleccionado para obtener las contribuciones individuales de los analitos. En líneas de punto gris se muestran los perfiles obtenidos para los demás analitos presentes en la muestra3.52
- Figura 3.9.** Esquema propuesto para los equilibrios acido-base de las piperizinil-QUI (A) y de las QUI ácidas (B).....3.61
- Figura 3.10.** Representación esquemática del diseño experimental combinado Lattice simplex y factorial en 3 niveles. Cada círculo celeste representa los puntos del diseño factorial, mientras que los círculos verdes indican los puntos correspondientes al diseño Lattice simplex3.67
- Figura 3.11.** Electroferograma obtenido para una solución mezcla de 50.0 ng mL^{-1} de (1) ENO, (2) CPF, (3) OFL, (4) (5) ENF y (6) FLU, y 167.0 ng mL^{-1} de CIN, sometida a SPE y analizada en condiciones experimentales óptimas. Detección: 280 nm 3.69
- Figura 3.12.** Estructura química de las QUI evaluadas.....3.70
- Figura 3.13.** Espectros de absorbancia UV normalizados de ENO (amarillo), CPF (línea continua celeste), ENF (línea celeste de rayas y puntos), OFL (verde), FLU (azul) y CIN (rojo) en 0.015 mol L^{-1} de solución $PO_4^{-3}:B_2O_7^{-2}$ pH 8.23.71
- Figura 3.14.** Electroferograma obtenido para una muestra de agua de consumo conteniendo 50.0 ng mL^{-1} de (1) ENO, (2) CPF, (3) OFL, (4) ENF y (5) FLU, y 167.0 ng mL^{-1} de (6) CIN, en presencia de potenciales interferentes, sometida a SPE. Detección: 280 nm3.72
- Figura 3.15.** Perfiles temporales (A) y espectrales (B) obtenidos por modelado quimiométrico con MCR-ALS para una muestra de agua de consumo conteniendo ENO (amarillo), CPF/ENF (celeste), OFL (verde), FLU (azul) y CIN (rojo), en presencia de interferentes (línea de puntos gris)3.74
- Figura 3.16.** Perfiles temporales obtenidos por modelado con MCR-ALS para una muestra de validación conteniendo los 6 compuestos de interés. El perfil de CPF y ENF se indica en azul, mientras que en líneas de punto gris se muestran los perfiles obtenidos para los demás analitos presentes en la muestra. La flecha roja señala el punto de división seleccionado para obtener las contribuciones individuales de CPF y ENF3.75
- Figura 3.17.** Esquema propuesto para los principales equilibrios acido-base de DXR en medio acuoso, con $pK_{a,1} = 8.15$, $pK_{a,2} = 10.16$ y $pK_{a,3} = 13.2$ 3.86
- Figura 3.18.** Perfiles espectrales (A) y de distribución de especies (B) obtenidos por PKFIT para DXR en el rango de pH de entre 6.0 y 14.0. Se distinguen 3 especies: DXR^+ (azul), DXR (rojo) y DXR^- (amarillo).....3.87

Figura 3.19. Representación esquemática de un sistema FIA para el análisis de una muestra ácida en un flujo transportador alcalino	3.88
Figura 3.20. Matriz de datos obtenida por FIA para una solución de DXR de C_{DXR} de $7.0 \mu\text{g mL}^{-1}$ indicándose el perfil de dispersión del bolo de muestra generado en el punto isobéptico $\lambda_{pi} = 530 \text{ nm}$	3.90
Figura 3.21. Perfiles espectrales (A) y temporales (B) obtenidos por MCR-ALS para una muestra de DXR de concentración $C_{DXR} = 7.0 \mu\text{g mL}^{-1}$, mostrándose la especie DXR^+ (azul), DXR (rojo) y el perfil de dispersión del bolo de muestra (línea de puntos verde) ante la restricción de condición de balance de masa. En los recuadros se muestran las curvas de distribución para la especies DXR^+ (azul) y DXR (rojo) obtenidas por el gradiente de pH generado en el sistema FIA para los rangos de tiempo comprendidos entre 0.00 min y 0.25 min (b.1) y entre 0.25 min y 1.00 min (b.2).....	3.91
Figura 3.22. Perfiles espectrales (A) y temporales (B) obtenidos para una muestra de plasma conteniendo DXR en presencia de RFP, mostrándose las contribuciones de las especies involucradas en el equilibrio: DXR^+ (azul), DXR (rojo), RFP (línea de puntos naranja) y plasma (línea de puntos gris).....	3.93
Figura 3.23. Curvas de abundancia con respecto a la concentración, obtenidas de la descomposición por MCR-ALS, para DXR^+ (azul), DXR (rojo) y para la suma de contribuciones individuales (verde). En línea de puntos se muestran las curvas ajustadas de regresión.....	3.94
Figura 3.24. (A) Región elíptica de confianza conjunta en el plano pendiente-ordenada al origen, para un nivel de confianza del 95 %, indicándose el punto ideal [1, 0] (estrella violeta) para pendiente y ordenada al origen, respectivamente (B) Gráfico de concentración predicha con respecto a concentración nominal, mostrándose la línea ajustada para los valores experimentalmente (línea de puntos celeste) y la curva de regresión ideal (línea de puntos violeta)	3.98
Figura 3.25. Cromatograma obtenido para la muestra Val _{RFP_4} , mostrándose las señales obtenidas para DXR (1, azul) y RFP (2, naranja), registrado a una longitud de onda de 254 nm	3.98
Figura 3.26. Representación esquemática del funcionamiento de un láser de cascada cuántica, QCL	3.102
Figura 3.27. Representación esquemática del dispositivo utilizado para las mediciones EC-QCL-IR.....	3.106
Figura 3.28. Curvas de transmisión consecutivas obtenidas por EC-QCL-IR para agua ultrapura. En el recuadro se amplifica la región espectral comprendida entre 1605 cm^{-1} y 1600 cm^{-1} , mostrándose las diferencias entre las fluctuaciones presentes de las estructuras finas de 2 líneas de transmisión consecutivas.....	3.108
Figura 3.29. Metodología de análisis desarrollada para el tratamiento de señales obtenidas por EC-QCL-IR.....	3.111
Figura 3.30. Espectros de absorbancia IR obtenidos por EC-QCL-IR sin procesamiento (azul), con corrección por alineación horizontal (celeste) y utilizando la metodología de procesamiento basada en COW (rojo) para una solución de 5.0 mg mL^{-1} de α -CT en PBS pH 7.0	3.111

- Figura 3.31. (Arriba)** Espectros de absorbancia IR obtenidos por EC-QCL-IR de 5 soluciones proteicas conteniendo 10.0 mg mL^{-1} (línea continua azul), 5.0 mg mL^{-1} (línea de rayas y puntos azul) y 2.5 mg mL^{-1} (línea de puntos azul) y **(abajo)** obtenidos por FT-IR (línea continua roja) para soluciones proteicas conteniendo 20.0 mg mL^{-1} . Las líneas de punto grises remarcan la congruencia entre las formas espectrales de los espectros obtenidos con los diferentes instrumentos 3.113
- Figura 3.32. (Arriba)** Espectros de segunda derivada invertidos obtenidos por EC-QCL-IR (líneas azules) y **(abajo)** FT-IR (líneas rojas) para las diferentes soluciones proteicas. Las líneas de punto grises remarcan la congruencia entre las formas espectrales de los espectros obtenidos con los diferentes instrumentos 3.114
- Figura 3.33.** Espectros de absorbancia IR obtenidos por EC-QCL-IR para una solución conteniendo 20.0 mg mL^{-1} de α -CT en TFEPBS (50:50) pH 7.0 a $25 \text{ }^\circ\text{C}$. Los espectros se registraron en un periodo de tiempo comprendido entre 2 min y 240 min. Las líneas azules muestran los espectros de α -CT conformada mayoritariamente por hélices α inducidas, mientras que las líneas rojas muestran los espectros de la proteína luego de la formación gradual de láminas β intermoleculares. El espectro correspondiente a α -CT nativa se muestra en línea de puntos negra. A la derecha se indica la transición conformacional dependiente del tiempo de reacción en escala de colores 3.116
- Figura 3.34.** Perfiles temporales relativos **(A)** y espectrales **(B)** obtenidos por MCR-ALS del monitoreo de la reacción de β -agregación de 20.0 mg mL^{-1} de α -CT en TFEPBS pH 8.2, mostrándose la evolución y los perfiles espectrales de las estructuras secundarias correspondientes a láminas β (rojo) y hélices α (azul). En gris se muestra el perfil obtenido para el ruido instrumental. Los perfiles temporales obtenidos para las láminas β a los diferentes valores de pH analizados **(C)** se muestran en escala de colores de verde a rojo para valores desde pH de 5.8 a 8.2, respectivamente 3.118
- Figura 3.35.** Efecto del pH del medio sobre la velocidad inicial v_0 de la reacción de agregación de 20.0 mg mL^{-1} de α -CT en TFEPBS. El valor de v_0 se calculó utilizando los perfiles temporales obtenidos por MCR-ALS 3.119
- Figura 3.36.** Perfiles temporales relativos **(A)** y espectrales **(B)** obtenidos por MCR-ALS del monitoreo de la reacción de β -agregación de 60.0 mg mL^{-1} de α -CT en TFEPBS pH 6.6, mostrándose la evolución y el perfil espectral de las estructuras secundarias correspondientes a láminas β (rojo) y hélices α (azul). En gris se muestra el perfil obtenido para el ruido instrumental. Los perfiles temporales obtenidos para las láminas β en los diferentes niveles de concentración proteica **(C)** se muestran en escala de colores desde verde a rojo para valores de concentración desde 5.0 mg mL^{-1} a 60.0 mg mL^{-1} , respectivamente 3.121
- Figura 3.37.** Velocidad inicial absoluta v_0 **(A)** y relativa $v_{0,R}$ **(B)** de la reacción de β -agregación para diferentes niveles de concentración de α -CT en TFEPBS pH 6.6, calculadas a partir de los perfiles temporales obtenidos de MCR-ALS 3.122
- Figura 3.38.** Curvas de transmisión consecutivas obtenidas por EC-QCL-IR para D_2O . En el recuadro se amplifica la región espectral comprendida entre 1655 cm^{-1} y 1650 cm^{-1} , mostrándose las fluctuaciones presentes entre las estructuras finas de 2 líneas de transmisión consecutivas 3.132

- Figura 3.39. (Arriba)** Espectros de absorbancia IR obtenidos por EC-QCL-IR de soluciones proteicas conteniendo 0.25 mg mL^{-1} (línea de puntos azul), 0.50 mg mL^{-1} (línea de rayas y puntos azul), 1.00 mg mL^{-1} (línea de rayas azul) y 2.50 mg mL^{-1} (línea continua azul) y **(abajo)** obtenidos por FT-IR (línea continua roja) para soluciones proteicas conteniendo 5.00 mg mL^{-1} . Las líneas de punto grises remarcan la congruencia entre las formas espectrales de los espectros obtenidos con los diferentes instrumentos 3.133
- Figura 3.40.** Espectros CD obtenidos en el rango UV lejano para una solución de 0.25 mg mL^{-1} de PLL en PBS $\text{pD } 12.4$, registrados en el rango de temperaturas de entre $20 \text{ }^{\circ}\text{C}$ y $50 \text{ }^{\circ}\text{C}$, cada $2 \text{ }^{\circ}\text{C}$. A la derecha se indica en escala de colores la transición conformacional α (azul) \rightarrow β (rojo) inducida por temperatura. En el recuadro se muestra la progresión de la elipticidad registrada a 222 nm en dependencia de la temperatura. La línea de puntos muestra la curva sigmoidea ajustada usando una función de Boltzmann. 3.135
- Figura 3.41.** Espectros IR obtenidos para una solución conteniendo 0.50 mg mL^{-1} de PLL en PBS $\text{pD } 12.4$, registrados en función de la temperatura por espectroscopia EC-QCL-IR. A la derecha se indica en escala de colores la transición conformacional α (azul) \rightarrow β (rojo) inducida por temperatura 3.136
- Figura 3.42.** Progresión del cambio de absorbancia IR relativa obtenida por evaluación de la altura de la banda localizada a 1611 cm^{-1} para distintos niveles de concentración. Para soluciones conteniendo 0.50 mg mL^{-1} (círculos anaranjados) y 3.00 mg mL^{-1} (círculos verde oscuros) se analizaron los espectros obtenidos por EC-QCL-IR, mientras que para las soluciones conteniendo 6.50 mg mL^{-1} (círculos verde claros) y 9.00 mg mL^{-1} (círculos celestes) se utilizaron los espectros obtenidos por FT-IR. Las líneas de puntos representan la curva sigmoidea ajustada a una función de Boltzmann 3.137
- Figura 3.43.** Temperaturas de transición correspondientes al cambio conformacional $\alpha \rightarrow \beta$ inducido por temperatura determinados por espectroscopia CD (círculos celestes), EC-QCL-IR (triángulos azules) y FT-IR (cuadrados rojos). Las bandas coloreadas representan la región de concentración cubierta por CD (celeste), EC-QCL-IR (azul) y FT-IR (rojo) 3.138
- Figura 3.44.** Perfiles espectrales **(A)** y de progresión conformacional dependientes de la temperatura **(B)** correspondientes a la estructura secundaria de helices α (azul), láminas β (rojo) y estructuras desordenadas (verde) 3.140
- Figura 3.45.** Espectros IR obtenidos para una solución conteniendo 18.00 mg mL^{-1} de PLL en PBS $\text{pD } 12.4$ registrados en función de la temperatura por espectroscopia FT-IR, mostrándose la banda amida I' y amida II'. A la derecha se indica en escala de colores la transición conformacional α (azul) \rightarrow β (rojo) inducida por temperatura 3.141
- Figura 3.46.** Perfiles espectrales **(A)** y de progresión conformacional dependiente de la temperatura **(B)** correspondientes a la estructura secundaria de helices α (azul), estructuras desordenadas (verde), láminas β_1 (marrón) y láminas β_2 (rojo) 3.143
- Figura 3.47.** Matrices de excitación-emisión, vistas desde el modo de excitación, registradas para las fracciones recolectadas en pocillos de ELISA de muestras conteniendo los analitos puros de OFL (azul), CPF (verde) y DNF (rojo). En línea de puntos gris se representa el perfil cromatográfico obtenido a partir de los máximos de intensidad registrados en cada pocillo para una muestra conteniendo los tres analitos. En el recuadro se muestra una matriz de excitación-emisión registrada para una de las fracciones recolectadas de CPF, indicándose el espectro de emisión al máximo de excitación correspondiente 3.157

- Figura 3.48.** Cromatogramas correspondientes a $\lambda_{em} = 350$ nm extraídos de las matrices de tiempo de retención-espectro de emisión para cada longitud de onda de excitación, mostrándose los espectros de excitación obtenidos para OFL a un tiempo de retención $t_R = 0.840$ min (azul claro) y $t_R = 0.852$ min (azul oscuro) y para CPF a un tiempo de retención $t_R = 0.930$ min (verde oscuro) y $t_R = 0.942$ min (verde oscuro). 3.158
- Figura 3.49.** Matrices de excitación-emisión consecutivas, vistas desde el modo de excitación, registradas para una muestra conteniendo CPF (verde) (A). En línea de puntos gris se representa el perfil cromatográfico obtenido a partir de los máximos de intensidad registrados para cada matriz. Se muestran las matrices de tiempo de retención-espectro de excitación (B.1) y tiempo de retención-espectro de emisión (B.2) normalizadas, reconstruidas a partir de los espectros de excitación y emisión correspondientes a los máximos de emisión y excitación, respectivamente. 3.160
- Figura 3.50.** Resultados obtenidos por PARAFAC para las 3 metodologías propuestas, mostrándose los perfiles temporales (A), espectrales de excitación (B) y espectrales de emisión (C). Para la Metodología I se muestran los perfiles de CPF (línea continua verde), OFL (línea continua azul) y DNF (línea continua rojo); para la Metodología II se indican los perfiles de CPF (línea continua verde) y OFL (línea continua azul) y un perfil correspondiente a un factor desconocido (línea de puntos y rayas azul verdosa); para la Metodología III se muestran los perfiles obtenidos para CPF (verde), observándose 2 factores para el mismo componente 3.162
- Figura 3.51.** Perfiles temporales (A) y espectrales (B) obtenidos por modelado quimiométrico con MCR-ALS de una muestra conteniendo 3 f-QUI. Los perfiles espectrales (B) se corresponden a las matrices de excitación-emisión desdobladas de OFL (azul), CPF (verde) y DNF (rojo). Los perfiles espectrales de excitación (C.1) y emisión (C.2) se obtuvieron a partir de las matrices de excitación-emisión reconstruidas de cada componente..... 3.167
- Figura 3.52.** Perfiles temporales (A) y espectrales de excitación (B.1) y emisión (B.2) obtenidos por modelado quimiométrico con APARAFAC de una muestra conteniendo OFL (azul), CPF (verde) y DNF (rojo), en presencia de 2 interferentes (línea de puntos gris). 3.169

ÍNDICE DE TABLAS

Tabla 1.1. <i>Parámetros utilizados en la expresión de sensibilidad de diferentes métodos de calibración</i>	1.30
Tabla 3.1. <i>Áreas obtenidas de las señales cromatográficas correspondientes para las f-QUI analizadas en ausencia y presencia de Y^{2+}</i>	3.46
Tabla 3.2. <i>Valores de concentración predicha en muestras de validación para las 7 f-QUI determinadas simultáneamente^a</i>	3.53
Tabla 3.3. <i>Valores de concentración encontrados de f-QUI en muestras acuosas ambientales^a</i>	3.54
Tabla 3.4. <i>Valores de límite de detección y cuantificación estimados para las f-QUI analizadas siguiendo las tres estrategias de cálculo^a</i>	3.56
Tabla 3.5. <i>Constantes de disociación y masa molar de las 6 QUI analizadas</i>	3.69
Tabla 3.6. <i>Valores de concentración predicha en muestras de validación para las 6 QUI determinadas simultáneamente^a</i>	3.76
Tabla 3.7. <i>Valores de concentración encontrados de QUI en muestras de agua de consumo^a</i>	3.77
Tabla 3.8. <i>Valores de límite de detección y cuantificación obtenidos para las QUI analizadas siguiendo tres estrategias de estimación^a</i>	3.78
Tabla 3.9. <i>Concentración de DXR predicha en muestras de validación y muestras de prueba en presencia de RFP^a</i>	3.95
Tabla 3.10. <i>Resultados obtenidos para el análisis de repetibilidad y precisión intermedia^a</i> ...	3.97
Tabla 3.11. <i>Valores de concentración predicha para para las 3 f-QUI en muestras de validación determinados por modelado con MCR-ALS y APARAFAC^a</i>	3.170
Tabla 3.12. <i>Valores de concentración encontrados de OFL, CPF y DNF en muestras de agua de consumo fortificadas aplicando el método desarrollado y utilizando modelado con MCR-ALS y APARAFAC^a</i>	3.171
Tabla 3.13. <i>Cifras de mérito analíticas obtenidas para el método desarrollado utilizando diferentes algoritmos quimiométricos para el modelado de datos de tercer orden^a</i> ..	3.172
Tabla 3.14. <i>Cifras de mérito obtenidas para OFL utilizando MCR-ALS y PARAFAC/APARAFAC para el modelado de datos de datos de segundo y tercer orden</i>	3.172

A stylized logo consisting of the letters 'E1' in white, bold, sans-serif font, enclosed within a blue circle. The circle is partially obscured by a white, curved, 3D-style shape that appears to be a stylized letter 'E' or a similar form. The background features a large, dark blue triangle on the right side and a vertical dark blue line on the left. At the bottom left, there are two small squares, one red and one grey, on a dark blue horizontal bar.

E1

Introducción

1. INTRODUCCIÓN

“The important thing in science is not so much to obtain new facts as to discover new ways of thinking about them”

Bragg, W.L.

1. 1. LA QUIMIOMETRÍA COMO HERRAMIENTA DE ANÁLISIS

La Quimiometría se puede definir como una disciplina química que utiliza métodos matemáticos y estadísticos empleando la lógica formal para diseñar o seleccionar procedimientos de medición óptimos o experimentos, para proveer la máxima información química relevante mediante el análisis de datos y para obtener información acerca del sistema químico bajo estudio (Massart *et al.*, 1998; Massart *et al.*, 1988).

Las herramientas quimiométricas son utilizadas, generalmente, para resolver problemas químicos descriptivos y predictivos. Las aplicaciones descriptivas se basan en modelar las propiedades de los sistemas químicos a fin de comprender el compor-

tamiento del sistema. Por otro lado, las aplicaciones predictivas se focalizan en el modelado de las propiedades químicas del sistema a fin de poder predecirlas en sistemas nuevos de composición desconocida. Por ello, es importante comprender que el análisis de datos multidimensionales no solo beneficia al área de la química analítica, sino que es posible encontrar aplicaciones quimiométricas en diferentes áreas científicas, convirtiéndola en una disciplina versátil y muy útil para la obtención de información, tanto en el ámbito académico como en el campo de la industria.

1. 2. CLASIFICACIÓN DE DATOS INSTRUMENTALES

Clasificación, según el diccionario de la Real Academia Española, es un concepto vinculado al ordenamiento o la disposición en clases de algo determinado. En Quimiometría, la clasificación de datos instrumentales resulta en una herramienta que ayuda a la comprensión del sistema bajo estudio, pero, sobre todo, a la toma de decisiones y selección de un correcto modelo de análisis.

Los datos instrumentales pueden ser categorizados o jerarquizados según el número de modos o dimensiones instrumentales – orden o vía – que presenten los datos que, a su vez, dependerá de la manera en que son colectados y organizados. Las dimensiones del dato dependerán, entonces, del diseño de los experimentos, de los instrumentos utilizados y de la forma en que los datos son organizados. A su vez, es importante remarcar que los datos generados para una sola muestra se clasifican según su orden, mientras que los datos generados para un grupo de muestras se clasifican según el número de vías en los que se disponen (Escandar *et al.*, 2007; Kiers, 2000; Olivieri *et al.*, 2014b). En la Figura 1.1 se muestra esquemáticamente la clasificación o jerarquización de los datos instrumentales según el orden de los datos obtenidos para una única muestra y las vías de un arreglo de datos para múltiples muestras.



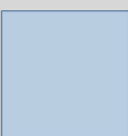
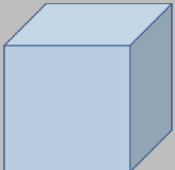


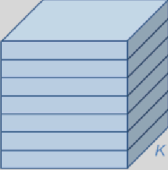
	Orden 0	1 ^{er} Orden	2 ^{do} Orden	3 ^{er} Orden
Única muestra	Escalar  1×1	Vector  $J \times 1$	Matriz  $J \times K$	Arreglo tridimensional  $J \times K \times L$
Múltiples muestras	1-vía  $I \times 1$ \mathbb{R}^I	2-vías  $I \times J$ $\mathbb{R}^{J \times I}$	3-vías  $I \times J \times K$ $\mathbb{R}^{J \times K \times I}$	4-vías No es posible graficar en el espacio tridimensional $I \times J \times K \times L$ $\mathbb{R}^{J \times K \times L \times I}$

Figura 1.1. Representación esquemática de la clasificación de los datos instrumentales según el “orden” de los datos obtenidos para una muestra única (**arriba**) y las “vías” de un arreglo multidimensional para múltiples muestras (**abajo**)

DATOS DE ORDEN CERO

Los datos de orden cero corresponden a los datos obtenidos por instrumentos que generan una única respuesta por muestra. Ejemplos representativos de este tipo de datos son las señales de absorción molecular registradas a una única longitud de onda, las lecturas de los electrodos ion selectivos, las señales obtenidas por absorción atómica o espectrofotometría de llama, entre otros. En términos algebraicos, los objetos de orden cero, es decir, las señales obtenidas para una muestra, se identifican como escalares.

Por otro lado, cuando se dispone de un conjunto de datos para un grupo de muestras, es posible construir un arreglo de datos de N -vías, obteniéndose un modo o dimensión adicional, es decir, un modo correspondiente al número de muestras. En este caso, los datos de orden cero pueden ser agrupados en un arreglo vectorial ($I \times 1$) o un arreglo de 1-vía, que en términos matemáticos significa obtener un tensor en el espacio \mathbb{R}^I , para un conjunto de I muestras (Kiers, 2000; Olivieri *et al.*, 2014b).

DATOS DE PRIMER ORDEN

Los datos de primer orden se refieren a aquellos datos generados por instrumentos que permiten la obtención de un conjunto de señales dispuestas en un arreglo vectorial por muestra. En otras palabras, la señal correspondiente a una única muestra es un vector de dimensiones $J \times 1$ para J entidades o sensores instrumentales. Ejemplos de este tipo de datos son los espectros de absorción Ultra Violeta-Visible (*UV-Vis*), los espectros de emisión de fluorescencia, los espectros de absorbancia infrarroja (*IR*) o las señales electroquímicas, como voltamperogramas o cronoamperogramas (Escandar *et al.*, 2007). Por otro lado, con un conjunto de datos de primer orden, obtenido para un grupo de muestras, es posible construir un arreglo matricial o un arreglo de 2-vías con dimensiones $I \times J$, es decir, un tensor en $\mathbb{R}^{J \times I}$, para J sensores instrumentales e I muestras.

DATOS DE SEGUNDO ORDEN

Los datos de segundo orden son datos generados para una muestra obtenidos como un arreglo matricial, es decir, la señal correspondiente a una única muestra es una matriz de datos de dimensiones $J \times K$, para J y K sensores instrumentales en la primera y segunda dimensión, respectivamente. Estos datos pueden ser obtenidos utilizando un único instrumento o mediante diferentes técnicas analíticas acopladas o dispuestas en tándem. Entre los ejemplos se destacan las matrices de excitación-emisión de fluorescencia (*EEM: Excitation-Emission Matrix*), los datos obtenidos de una reacción cinética o de gradiente de pH con detección espectral y los obtenidos por cromatografía acoplada a un detector de arreglo de diodos o a espectrometría de masas, entre otros.

En este caso, un conjunto de datos obtenidos para un grupo de muestras se pueden disponer en un arreglo de 3-vías $I \times J \times K$, para J y K sensores instrumentales en la primera y segunda dimensión, respectivamente, e I muestras, es decir, un tensor en $\mathbb{R}^{J \times K \times I}$. Sin embargo, es posible construir, también, un arreglo de datos de 2-vías aumentado cuando las matrices son ordenadas consecutivamente en sentido de filas o columnas, es decir, un arreglo de 2-vías aumentado en filas o columnas, respectivamente (de Juan *et al.*, 2009; Muñoz de la Peña *et al.*, 2015) y un arreglo de 2-vías, utilizando los vectores obtenidos de matrices desdobladas, de dimensiones $1 \times JK$, dispuestos en un arreglo aumentado en columna ($I \times JK$).

DATOS DE TERCER ORDEN

Sobre la base de lo explicado anteriormente se puede inferir que la introducción de una nueva dimensión instrumental para una muestra lleva consigo un incremento en el orden de los datos obtenidos.

Los datos de tercer orden pueden ser obtenidos, por ejemplo, por monitoreo de una reacción cinética o de gradiente de pH con espectroscopia de excitación-emisión de fluorescencia u obtenidos con un cromatógrafo acoplado a detección de EEM de fluorescencia.

Con un conjunto de datos de tercer orden, de manera análoga a lo mencionado para datos de orden inferior, se pueden construir arreglos de datos diferentes: **1-** Arreglo de 4-vías, construido a partir de la disposición de hiper-matrices $J \times K \times L$ generando un arreglo con dimensiones $I \times J \times K \times L$, es decir, un tensor en el espacio $\mathbb{R}^{J \times K \times L \times I}$, para J , K y L sensores instrumentales en la primer, segunda y tercera dimensión, respectivamente, e I muestras; **2-** Arreglo de 3-vías aumentado, construido con las hiper-matrices $J \times K \times L$ ordenadas consecutivamente en el sentido de una de las J , K , L dimensiones, obteniéndose una hiper-matriz aumentada de dimensiones $JL \times K \times L$, por ejemplo. **3-** Arreglo de 2-vías aumentado, construido con matrices desdobladas $J \times KL$, por ejemplo, y ordenadas de manera consecutivas en sentido de las filas o las columnas, generándose un arreglo de 2-vías aumentado de dimensiones $JL \times KL$, siguiendo el mismo ejemplo, **4-** Arreglo de 2-vías, construido con vectores desdoblados con dimensiones $JKL \times 1$ y ordenados de manera consecutiva en el sentido de columnas, generándose un arreglo de 2-vías de dimensiones $I \times JKL$.

1. 3. PROPIEDADES DE LOS DATOS MULTIDIMENSIONALES

BILINEALIDAD

Para datos de segundo orden, la señal instrumental obtenida para un compuesto puro, en ausencia de ruido, puede expresarse como el producto de dos vectores correspondientes a cada modo instrumental, según

$$\mathbf{X} = \mathbf{b} \mathbf{c}^T \quad 1.1$$

siendo \mathbf{X} una matriz de datos ($J \times K$), \mathbf{b} un vector columna con J filas ($J \times 1$) y \mathbf{c} un vector columna con K filas ($K \times 1$). El superíndice "T" implica la operación transpuesta de una matriz, por lo que \mathbf{c}^T será un vector fila con K columnas ($1 \times K$).

En una muestra conteniendo N componentes, la matriz de datos \mathbf{X} se expresa como la sumatoria de todas las contribuciones individuales de cada componente puro, es decir:

$$\mathbf{X} = \sum_{n=1}^N \mathbf{b}_{jn} \mathbf{c}_{kn}^T = \mathbf{B} \mathbf{C}^T \quad 1.2$$

siendo \mathbf{B} una matriz ($J \times N$) conteniendo los perfiles de uno de los modos instrumentales de cada N componente y \mathbf{C}^T una matriz ($N \times K$) conteniendo los N perfiles del modo instrumental restante.

Debido a que \mathbf{X} es lineal en \mathbf{B} (o \mathbf{b}) y lineal en \mathbf{C} (o \mathbf{c}), para \mathbf{C} y \mathbf{B} fijas, respectivamente, la ecuación 1.2 define a una matriz doblemente lineal, es decir, a una matriz de datos bilineal. En otras palabras, una matriz bilineal \mathbf{X} puede ser descompuesta en el producto de dos matrices individuales o sub-matrices.

Es importante remarcar que desde el punto de vista químico el fenómeno de bilinealidad ocurre cuando los modos instrumentales, es decir, los perfiles contenidos en \mathbf{B} y \mathbf{C} , muestran independencia entre sí. Adicionalmente, se debe cumplir que una matriz bilineal \mathbf{X} es el producto son dos sub-matrices \mathbf{B} y \mathbf{C} que contiene un número de columnas igual al número de constituyentes químicos presentes en el sistema, es decir, contiene N columnas. Esto conlleva a la introducción del concepto de *rango* que se define como el número de términos bilineales necesarios para reproducir una matriz de datos. En consecuencia, cuando el *rango* es igual al número de constituyentes químicos presentes en el sistema, la matriz es bilineal; de otra manera, la matriz \mathbf{X} es una matriz no-bilineal de datos (Olivieri *et al.*, 2014b).

TRILINEALIDAD

El concepto de trilinealidad puede ser explicado para un arreglo de 3-vías cons-

truido con datos de segundo orden obtenidos para un grupo de muestras, o bien, para un dato de tercer orden obtenido para una única muestra.

De manera análoga a lo explicado para datos bilineales, para un modelo trilineal se tiene que,

$$x_{ijk} = \sum_{n=1}^N a_{in} b_{jn} c_{kn} + e_{ijk} \quad 1.3$$

donde x_{ijk} representa la respuesta para la muestra i a un sensor determinado del modo j y k ; a_{in} , b_{jn} y c_{kn} son los componentes de cada modo instrumental y e_{ijk} colecta los errores del modelo. También es posible disponer estos valores en arreglos de matrices donde \mathbf{B} es una matriz ($J \times N$, *conteniendo todos los b_{jn}*) que contiene a los N perfiles correspondientes a uno de los modos instrumentales y \mathbf{C} es una matriz ($K \times N$, *conteniendo todos los c_{kn}*) conteniendo los N perfiles del segundo modo instrumental. Aquí, \mathbf{A} ($I \times N$, *conteniendo todos los a_{in}*) puede tener identidades diferentes dependiendo de si se trata de un arreglo de 3-vías o un dato de tercer orden: **1- Arreglo de 3-vías:** \mathbf{A} *representa el modo correspondiente a las diferentes muestras;* **2- Dato de tercer orden:** \mathbf{A} *contiene el perfil de cada N componente correspondiente al tercer modo instrumental* (Bro, 1997; Olivieri *et al.*, 2014b).

Una de las condiciones que se deben cumplir para la trilinealidad de un arreglo de 3-vías, \mathbf{X} , para un sistema conteniendo un único componente, es que las matrices individuales obtenidas para cada muestra independiente deben ser bilineales. Además, los elementos contenidos en \mathbf{B} deben ser independientes de la muestra, es decir, debe existir un vector \mathbf{b} que describa el comportamiento del constituyente n en el mismo modo instrumental en todas las muestras. Lo mismo debe ocurrir para los elementos contenidos en \mathbf{c} . En conclusión, la trilinealidad de un arreglo de 3-vías compuesto por matrices de datos para múltiples muestra requiere que cada matriz de datos sea bilineal y que ambos perfiles, \mathbf{b} y \mathbf{c} para cada constituyente, sean el mismo en todas las muestras. Por lo tanto, la falta de reproducibilidad entre muestras en uno de los perfiles conlleva a la obtención de un arreglo \mathbf{X} no trilineal (Olivieri *et al.*, 2014b).

Sin embargo, ante la presencia de un arreglo de 3-vías no trilineal, la construcción de una matriz aumentada resulta en una excelente alternativa para la descomposición de los datos. De esta manera, se construye una matriz bilineal aumentada \mathbf{X}_{aum} , que sigue la expresión de la ecuación 1.2. Es importante remarcar que para obtener

una matriz aumentada bilineal, se debe realizar la aumentación en la dirección correspondiente al modo que rompe la trilinealidad, que puede ser en filas o en columnas. Así, la descomposición bilineal para una matriz \mathbf{X}_{aum} aumentada en columnas, construida con dos matrices \mathbf{X} individuales, se representa como

$$\mathbf{X}_{\text{aum}} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix} = \mathbf{B}_{\text{aum}} \mathbf{C}^T = \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{bmatrix} \mathbf{C}^T \quad 1.4$$

donde las columnas de \mathbf{B}_{aum} contienen los perfiles aumentados en la dirección de columnas de los N constituyentes de las muestras y \mathbf{C} contiene los perfiles de cada N constituyente del modo instrumental restante.

1. 4. GENERACIÓN DE DATOS MULTIDIMENSIONALES

Los datos de orden superior pueden ser generados, como se explicó más arriba, utilizando un único instrumento o mediante el acoplamiento de distintas técnicas analíticas. Ante los posibles escenarios que se pueden tener para este fin, es importante, en primer lugar, conocer los detalles de los datos generados, ya sea desde el punto de vista experimental, como desde una visión matemática. Aquí, se explicarán algunas de las metodologías más utilizadas para la generación de datos multidimensionales.

MATRICES DE EXCITACIÓN-EMISIÓN DE FLUORESCENCIA

Una de las metodologías más populares para la generación de datos de segundo orden es la basada en la generación de EEM. Esto se debe, entre otras cosas, a la simplicidad de la obtención de los datos, al requerimiento de un único instrumento, a que las señales de emisión de fluorescencia presentan alta la sensibilidad en comparación con espectroscopia UV y a que se requiere de instrumentos de costos relativamente bajos.

Una de sus principales características es que, al no ser una técnica separativa, para una muestra multi-componente se obtendrán señales altamente solapadas requi-

riéndose de las ventajas provistas por las herramientas quimiométricas para la resolución de las contribuciones individuales de cada componente puro (Mbogning Feudjio *et al.*, 2014; Qing *et al.*, 2013; Sajjadi *et al.*, 2016; Zhang *et al.*, 2016).

La obtención de una EEM de fluorescencia se logra registrándose las señales de emisión de fluorescencia en una determinada región espectral a diferentes longitudes de onda de excitación. Las regiones espectrales de emisión y excitación son seleccionadas de acuerdo al sistema en estudio y a los componentes que constituyen las muestras.

Debido a que los espectros de excitación y emisión presentan independencia mutua, las EEM son matrices bilineales y, a causa de que los espectros solo dependen de la composición del medio y de la naturaleza del analito, los datos obtenidos para un grupo de muestras pueden disponerse en un arreglo trilineal de 3-vías, $\underline{\mathbf{X}}$. Sin embargo, en presencia de efectos de filtro interno es posible observar ruptura de trilinealidad en el modo correspondiente a las muestras y, entonces, se deberán analizar las posibles alternativas para el ordenamiento de los datos.

En la Figura 1.2 se expone una EEM obtenida para una muestra conteniendo un analito, mostrándose las proyecciones de los espectros de excitación y emisión correspondientes.

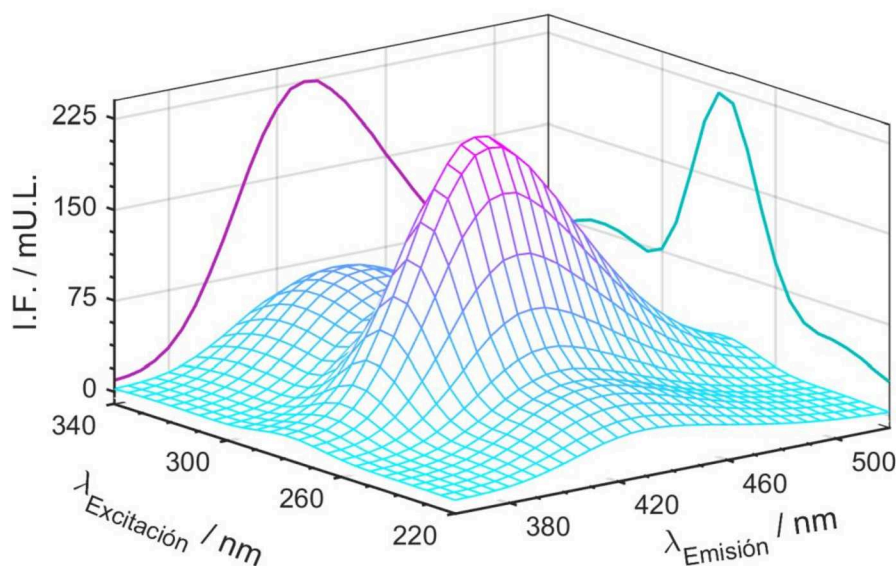


Figura 1.2. Matriz de excitación-emisión de fluorescencia obtenida para una solución de sarafloxacin en matriz acuosa. En celeste se muestra el espectro de excitación obtenido utilizando una longitud de onda de emisión de 450 nm y en rosa se muestra el espectro de emisión obtenido utilizando una longitud de onda de excitación de 270 nm

MÉTODOS SEPARATIVOS CON DETECCIÓN ESPECTRAL

Las técnicas cromatográficas acopladas a detección espectral UV, de emisión de fluorescencia o a espectroscopia de masas han sido ampliamente utilizadas para la generación de datos de segundo orden basados en métodos separativos (Giuliani *et al.*, 2016; López-Cobo *et al.*, 2016; Tajabadi *et al.*, 2016), si bien, los métodos electroforéticos han empezado a ocupar su lugar en el área de datos multidimensionales (de Oliveira Mendes *et al.*, 2016; Pasquini *et al.*, 2016).

Tradicionalmente, los métodos separativos han sido utilizados para el análisis de muestras multi-componentes bajo la premisa de la obtención de señales individuales puras de cada constituyente, es decir, lográndose la resolución completa entre señales cromatográficas o electroforéticas de los distintos constituyentes de la muestra. Sin embargo, la resolución completa de las señales no siempre se logra y es posible obtener señales solapadas y también, efectos de línea de base.

La matriz de datos generada para una muestra única utilizando un método separativo con detección espectral contiene los tiempos de retención/elución en una de las dimensiones y la información espectral en la dimensión restante. Las matrices se obtienen registrándose las señales espectrales en un rango espectral determinado para cada tiempo de retención/elución en un rango de tiempo establecido. Los instrumentos que permiten obtener este tipo de datos deben contar de detectores que permitan la obtención de espectros en un tiempo muy corto, a fin de no introducir efectos de distorsión en virtud de la movilidad del flujo o la fase móvil. Dentro de los detectores más comunes encontrados en los instrumentos analíticos están los detectores de arreglo de diodos UV-Vis (*DAD: Diode Array Detector*), los detectores de fluorescencia de barrido rápido (*FSFD: Fast-Scanning Fluorescence Detector*) y los detectores de masas (*MSD: Mass Detector*).

Si bien el modo correspondiente a los tiempos de retención/elución es independiente al modo espectral, es decir, las matrices generadas son bilineales, los arreglos de 3-vías \mathbf{X} , generalmente, no resultan en una opción conveniente, debido a la pérdida de trilinealidad en virtud de las variaciones en los tiempos de retención/elución de los constituyentes entre diferentes muestras. Por lo tanto, para el análisis de múltiples muestras es conveniente utilizar arreglos de 2-vías aumentados \mathbf{X}_{aum} .

Por otro lado, a pesar de no involucrar el fenómeno de separación de los constituyentes de la muestra, es posible obtener datos de segundo orden con características similares mediante la utilización de técnicas basadas en análisis en flujo (*FI: Flow Inje-*

ction), como el análisis por inyección en flujo (*FIA: Flow Injection Analysis*) (Musil *et al.*, 2016) o el análisis por inyección secuencial (*SIA: Sequential Injection Analysis*) (Pasamontes *et al.*, 2006).

Adicionalmente, con estas técnicas analíticas es posible obtener datos de orden superior incrementando el número de modos instrumentales, por ejemplo, cromatografía acoplada a detección de EEM, en donde para cada tiempo de retención se corresponden 2 señales espectrales (Bro, 1997; Lozano *et al.*, 2013).

MONITOREO DE REACCIONES CON DETECCIÓN ESPECTRAL

Las reacciones químicas, los equilibrios químicos e incluso los cambios estructurales de macromoléculas pueden ser monitoreados utilizando diferentes técnicas espectrales. Entre los trabajos publicados en la literatura se pueden encontrar métodos cinéticos con detección UV (Feng *et al.*, 2013) o de emisión de fluorescencia (Lukman *et al.*, 2012), métodos de cambios de pH por titulación espectrofotométrica (Meloun *et al.*, 2016), métodos de dicroísmo circular con detección UV para el análisis de proteínas (Borges *et al.*, 2005) y cambios estructurales de macromoléculas con detección infrarroja (Martí-Aluja *et al.*, 2013). Estos datos han sido ampliamente utilizados tanto para análisis cuantitativo como análisis descriptivo en virtud del gran contenido y calidad de información que ofrecen.

Las matrices de datos se pueden generar en modo continuo, en una reacción continua, por ejemplo, en una reacción cinética, o de manera discreta para reacciones que pueden ser detenidas, por ejemplo, en una titulación. Debido a que la mayoría de las reacciones cinéticas dependen de la concentración, para el análisis de múltiples muestras, generalmente, no se utilizan arreglos de 3-vías \mathbf{X} a causa de la pérdida de trilinealidad y se utilizan los arreglos de 2-vías aumentados \mathbf{X}_{aum} .

Con estas técnicas analíticas es posible, también, obtener datos de orden superior aumentando el número de modos instrumentales, por ejemplo, en reacciones cinéticas con detección de EEM (Montemurro *et al.*, 2017).

1. 5. ALGORITMOS QUIMIOMÉTRICOS

Los procedimientos matemáticos para el procesamiento de datos multidimensionales son algoritmos que permiten el análisis de los datos en base a un determinado modelo matemático. En la literatura es posible encontrar un sinnúmero de algoritmos que pueden ser utilizados para el análisis de datos, que dependen, principalmente, del tipo de dato instrumental obtenido y del propósito del análisis (Olivieri, 2012).

Los datos de orden cero, obtenidos, principalmente, con propósitos cuantitativos, se utilizan para la construcción de modelos de regresión mediante el análisis de arreglos de 1-vía por calibración univariada, modelando la relación entre la señal instrumental univariada y la concentración del analito. Las señales instrumentales se ajustan por cuadrados mínimos a un modelo matemático lineal o no-lineal, lo que requiere de la selectividad y resolución completa del componente de interés en la muestra (Olivieri, 2014).

En datos de 2-vías, la relación entre la señal instrumental y la concentración de los analitos es modelada usando estructuras bilineales basadas en variables latentes, que son combinaciones lineales de los perfiles reales y no presentan interpretación física alguna. La premisa de estos algoritmos es compensar la falta de selectividad extrayendo de la señal total la porción que se pueda utilizar. Los algoritmos de regresión en componentes principales (*PCR: Principal Component Regression*) o cuadrados mínimos parciales (*PLS: Partial Least Squares*) son los más utilizados para el modelado de este tipo de datos con fines cuantitativos. Sin embargo, otro tipo de algoritmos son utilizados para datos que presenten comportamiento no-lineal, como lo son las redes neuronales artificiales (*ANN: Artificial Neural Network*) o las máquinas de vectores soportes acopladas a ajuste por cuadrados mínimos (*LS-SVM: Least Squares-Support Vector Machine*). Los algoritmos de primer orden permiten la identificación de muestras conteniendo componentes inesperados, es decir, componentes que no se consideraron en la etapa de calibración. Si una muestra de composición desconocida no se ajusta al modelo de calibración realizado, la muestra se considera anómala, indicando que su composición difiere de las muestras representativas utilizadas en la calibración. Esta propiedad es conocida como “ventaja de primer orden” (Escandar *et al.*, 2007; Olivieri *et al.*, 2014b).

Para el análisis de datos de segundo orden es posible recurrir a una amplia variedad de algoritmos, incluyendo análisis paralelo de factores (*PARAFAC: PARAllel FACtor Analysis*) y sus variantes (Bro, 1997), el método de aniquilación de rango genera-

lizado (*GRAM: Generalized Rank Annihilation Method*), la resolución multivariada de curvas acoplada a cuadrados mínimos alternantes (*MCR-ALS: Multivariate Curve Resolution-Alternating Least Squares*) (Tauler, 1995) y la descomposición trilineal alternante (*ATLD: Alternating Trilinear Decomposition*) (Wu et al., 1998) y sus variantes, entre otros.

Los métodos basados en ALS son considerados los más eficientes para la descomposición de arreglos de 3-vías, siendo PARAFAC el algoritmo de elección para la mayoría de los análisis de arreglos de 3- o N -vías. Sin embargo, ante la pérdida de trilinealidad, MCR-ALS y PARAFAC2 son los algoritmos mayormente utilizados. Estos algoritmos se aplican con fines descriptivos y predictivos en virtud de que los perfiles obtenidos de la descomposición tienen interpretación física y pueden ser inferidos a las propiedades físico-químicas del sistema en estudio. Por otro lado, los algoritmos como cuadrados mínimos parciales desdoblados (*U-PLS: Unfolded Partial Least Squares*) o cuadrados mínimos parciales multi-vía (*N-PLS: N-way Partial Least Squares*), acoplados a procedimientos de bilinealización residual (*RBL: Residual Bilinearization*) son utilizados, específicamente, con fines cuantitativos en virtud de que los perfiles obtenidos como *scores* y *loadings* carecen de interpretación física (Olivieri, 2012). Todos estos algoritmos son de gran importancia en el análisis de muestras complejas, debido a que todos ellos alcanzan la llamada “ventaja de segundo orden”. Esta ventaja se basa en el hecho de que los analitos pueden ser determinados en presencia de constituyentes no esperados o potenciales interferentes que pueden estar presentes en las muestras de composición desconocida (Muñoz de la Peña et al., 2015; Olivieri et al., 2014b).

Para datos de orden superior, se pueden utilizar las metodologías antes descritas para los datos de segundo orden utilizando modelos de estructuras latentes no cuadrilineales o métodos basados en ALS cuadrilineales. PARAFAC es un modelo que puede ser aplicado a datos cuadrilineales e incluso a datos de orden superior. Para datos no cuadrilineales, sin embargo, los modelos basados en variables latentes acoplados a trilinealización residual (*RTL: Residual Trilinearization*) resultan de elección. Ejemplo de estos últimos son U-PLS/RTL, N-PLS/RTL y las técnicas híbridas trilineales de cuadrados mínimos acoplados a RTL (*TLLS/RTL: Technique Trilinear Least Squares*) (Olivieri, 2012). Se ha demostrado, también, que el uso de MCR-ALS para arreglos de 2-vías desdoblados aumentados es una herramienta útil para el análisis de datos de orden superior (Malik et al., 2014). Recientemente, se ha desarrollado un algoritmo que combina las propiedades de MCR-ALS y PARAFAC para el análisis de datos de tercer orden dispuestos en un arreglo de 3-vías aumentado no cuadrilineal \mathbf{X}_{aum} , denominado PARAFAC aumentado o APARAFAC (*APARAFAC: Augmented*

PARAFAC) (Bortolato *et al.*, 2015).

Aquí se desarrollaran en detalle los algoritmos utilizados en esta tesis: **1- MCR-ALS**; **2- PARAFAC**; y **3- APARAFAC**.

RESOLUCIÓN MULTIVARIADA DE CURVAS-CUADRADOS MÍNIMOS ALTERNANTES (MCR-ALS)

MCR-ALS es un método de modelado flexible de datos (*soft-modelling*) que ha sido ampliamente utilizado desde su desarrollo en el año 1995 (Tauler, 1995). Esta técnica permite obtener información acerca de un sistema multi-componente por discriminación de las contribuciones individuales de los constituyentes solapados. MCR-ALS se focaliza en la descomposición bilineal de una matriz de datos \mathbf{X} en dos sub-matrices conteniendo información de los componentes puros involucrados en el sistema. La descomposición es el resultado de la validez de la ley de Lambert-Beer y se logra por un proceso de optimización iterativa.

Para un sistema multi-componente que sigue la ley de Lambert-Beer, la señal o respuesta instrumental obtenida es producto de la contribución de todas las señales individuales de cada componente puro. En espectroscopia, esto significa que

$$d_j = \sum_{n=1}^N c_n s_{j,n} \quad 1.5$$

donde d_j representa la respuesta instrumental a la longitud de onda j , c_n es la concentración de la especie n en la mezcla y $s_{j,n}$ es factor de sensibilidad instrumental de la especie n a la longitud de onda j . Cuando se obtienen las señales instrumentales de un sistema conteniendo N constituyentes, en un rango espectral determinado de J sensores, registrándose la evolución de la concentración de cada componente, lo que se obtiene es un modelo bilineal que sigue la expresión:

$$\mathbf{D} = \mathbf{C} \mathbf{S}^T + \mathbf{E} \quad 1.6$$

donde la matriz \mathbf{D} ($J \times K$) contiene las señales instrumentales totales del sistema en estudio; \mathbf{C} ($K \times N$) contiene los K sensores del primer modo instrumental de cada N componente presente en el sistema, por ejemplo, la evolución de la concentración en

un proceso; \mathbf{S} es una matriz ($J \times N$) conteniendo el factor de sensibilidad instrumental de cada uno de los N componentes a cada j longitud de onda y \mathbf{E} ($J \times K$) representa la contribución del error de las mediciones.

La descomposición bilineal y la búsqueda de las respuestas individuales se logra mediante un proceso basado en un método iterativo, obteniéndose directamente estimaciones de las matrices \mathbf{C} y \mathbf{S} , representadas como $\hat{\mathbf{C}}$ y $\hat{\mathbf{S}}$, respectivamente. MCR-ALS se basa en el cálculo iterativo de \mathbf{S} y \mathbf{C} según

$$\hat{\mathbf{S}}^T = \mathbf{C}^+ \mathbf{D} = (\hat{\mathbf{C}}^T \hat{\mathbf{C}})^{-1} \hat{\mathbf{C}} \mathbf{D} \quad 1.7$$

$$\hat{\mathbf{C}} = \mathbf{D} \hat{\mathbf{S}}^{T+} = \hat{\mathbf{S}} \mathbf{D} (\hat{\mathbf{S}}^T \hat{\mathbf{S}})^{-1} \quad 1.8$$

donde $\hat{}$ indica el valor estimado de la variable correspondiente y $^+$ representa la operación pseudoinversa de una matriz. Se debe tener en cuenta que para comenzar el proceso de iteración se requieren parámetros que contengan información similar a la del sistema en estudio, es decir, se requiere de estimaciones iniciales tipo- \mathbf{C} o tipo- \mathbf{S} que pueden ser obtenidas experimentalmente o mediante algoritmos de descomposición bilineal auxiliares.

El proceso de optimización ALS se basa en la minimización del error residual de la descomposición bilineal mediante el uso de restricciones matemáticas o químicas. Dentro de las restricciones se encuentran aquellas que solo son aplicables en los modelos bilineales para una única muestra, o análisis de muestra individual, y aquellas que son propias de los modelos para arreglos de 2-vías aumentados.

Las restricciones se clasifican según su naturaleza o implementación. Cuando las restricciones son implementadas, la información química o matemática acerca de los perfiles debe ser traducida a una condición matemática (de Juan *et al.*, 2009). Las restricciones más utilizadas son (de Juan *et al.*, 2009; Tauler *et al.*, 2009):

No-negatividad: fuerza a los perfiles a tener valores positivos. Es aplicable a concentraciones o respuestas instrumentales, que, por naturaleza, deben ser positivas, por ejemplo, los espectros de emisión de fluorescencia o espectros de masa;

Unimodalidad: fuerza a los perfiles a tener un único máximo. Es aplicable para perfiles que presenten formas de pico, como señales cromato-

gráficas, o perfiles que muestren crecimiento o decrecimiento con valores tipo plató máximos o mínimos, respectivamente;

Condición de sistema cerrado: representa la traducción matemática de la condición de balance de masa en química. Es aplicable solo cuando la concentración total de las especies detectables es constante en cada etapa del proceso;

Modelado rígido (hard-modelling): fuerza a los perfiles de concentración o de respuesta instrumental a seguir una función matemática determinada. En dirección de la concentración, los perfiles se fuerzan a presentar la forma descrita por funciones típicas de modelos fisicoquímicos;

Trilinealidad: sigue la expresión de modelo trilineal y solo es aplicable para el modelado de múltiples muestras dispuestas en un arreglo de 2-vías aumentado. Al utilizarla se asegura que los resultados sean únicos, es decir, se asegura la obtención de resultados sin ambigüedad;

Correspondencia entre especies: es aplicable para el análisis de múltiples muestras en un arreglo de 2-vías aumentado. Se establece la presencia o ausencia de un componente en una matriz particular \mathbf{C} y/o \mathbf{S} . La información presencia/ausencia se aplica como código binario de 1 y 0, respectivamente.

En general, el proceso de MCR-ALS incluye: **1-** Determinación del número de especies espectroscópicamente activas; **2-** Obtención de estimaciones iniciales tipo- \mathbf{C} o tipo- \mathbf{S} del sistema; **3-** Cálculo de \mathbf{S}^T utilizando restricciones; **4-** Cálculo de \mathbf{C} usando restricciones; **5-** Reproducción de \mathbf{D} con los productos de \mathbf{C} y \mathbf{S}^T ; y **6-** Repetición desde **3-** hasta obtener convergencia. El criterio de convergencia se basa en la comparación del ajuste obtenido en dos iteraciones consecutivas y se logra cuando la diferencia relativa entre dos iteraciones consecutivas es menor a un valor umbral establecido. Sin embargo, es conveniente una inspección visual de la evolución de los perfiles, a fin de asegurar la obtención de la resolución óptima. En la Figura 1.3 se representa esquemáticamente la descomposición de un modelo MCR-ALS para un arreglo de 2-vías aumentado.

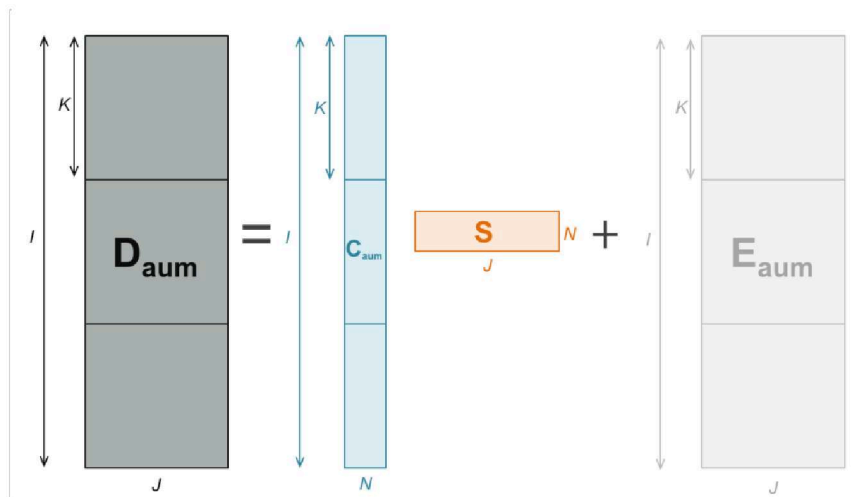


Figura 1.3. Modelo de descomposición MCR-ALS para datos de segundo orden dispuestos en un arreglo de 2-vías aumentado para un sistema de múltiples muestras conteniendo N constituyentes

Se debe tener en cuenta que con MCR es posible obtener resultados ambiguos para un mismo sistema, debido a que el modelo presenta libertad rotacional, en otras palabras, la descomposición puede dar lugar a diferentes resultados que explican el mismo sistema. Algunas de las restricciones matemáticas que se utilizan ayudan, principalmente, a la reducción de la ambigüedad rotacional, mientras que otras restricciones, como la condición de sistema cerrado, favorecen la reducción de la ambigüedad por diferencias en concentraciones. Por lo tanto, el uso de las diferentes restricciones, como también, el uso de los arreglos de 2-vías aumentados, analizando múltiples muestras simultáneamente, ayuda a disminuir la ambigüedad de los resultados, obteniéndose resultados cercanos a la situación real. Adicionalmente, los perfiles contenidos en **C** y **S** permiten inferir acerca del comportamiento de los constituyentes del sistema en el proceso y de su comportamiento espectral a causa de que los perfiles pueden ser físicamente interpretados. Por otro lado, la información contenida en **C** puede ser utilizada para fines predictivos en calibración multivariada o multi-vía.

ANÁLISIS PARALELO DE FACTORES (PARAFAC)

El modelo PARAFAC está intrínsecamente relacionado con el principio de perfiles proporcionales paralelos introducido por Cattell en el año 1944 (Cattell, 1944). Este principio establece que un mismo conjunto de perfiles describiendo la variación en más de un conjunto de datos de 2-vías, pero en diferentes proporciones, llevará a un modelo que no está sujeto a libertad rotacional (Bro, 1997). PARAFAC es un método de des-

composición de datos multidimensionales de segundo orden u orden superior. Para el análisis de arreglos de 3-vías, la descomposición se realiza en componentes trilineales o triadas, obteniéndose 3 vectores de *loadings* que son tratados matemáticamente iguales (Martí *et al.*, 2015).

Un modelo PARAFAC para un arreglo de 3-vías está definido por los elementos a_{in} , b_{jn} y c_{kn} (Bro, 1997) siguiendo un modelo de triadas, según (Ortiz *et al.*, 2015)

$$x_{ijk} = \sum_{n=1}^N a_{in} b_{jn} c_{kn} + e_{ijk} \quad i = 1, 2, \dots, I; j = 1, 2, \dots, J; k = 1, 2, \dots, K \quad 1.9$$

donde a_{in} denota el valor de los perfiles en el modo de las muestras para los N constituyentes en la muestra i ; b_{jn} y c_{kn} son los valores correspondientes a los 2 modos instrumentales para los n constituyentes en cada sensor instrumental j y k , respectivamente y e_{ijk} colecta los errores del modelo.

Asimismo, los *loadings* de todos los componentes son agrupados en matrices **A**, **B** y **C**, cuyas N columnas son los vectores \mathbf{a}_n , \mathbf{b}_n y \mathbf{c}_n , respectivamente. El modelo, entonces, puede ser expresado como producto de Khatri-Rao utilizando notación matricial

$$\underline{\mathbf{X}} = \mathbf{A} \mathbf{G} (\mathbf{C}|\otimes|\mathbf{B})^T + \underline{\mathbf{E}} \quad 1.10$$

donde $\underline{\mathbf{X}}$ representa un arreglo de datos 3-vías, \mathbf{G} es una matriz diagonal cuyos elementos no nulos indican al componente n de la matriz **C** y $\underline{\mathbf{E}}$ es la hiper-matriz conteniendo los residuos del modelo. El símbolo $|\otimes|$ es el producto Khatri-Rao. Análogamente a la ecuación 1.9, **A**, **B** y **C** corresponden a los perfiles de las diferentes muestras y los 2 modos instrumentales, respectivamente (Bro, 1997; Martí *et al.*, 2015). Comúnmente, el modelo PARAFAC es expresado como el producto Kronecker según,

$$\underline{\mathbf{X}} = \sum_{n=1}^N \mathbf{a}_n \otimes \mathbf{b}_n \otimes \mathbf{c}_n + \underline{\mathbf{E}} \quad 1.11$$

donde \otimes indica el producto de Kronecker, N es el número total de componentes y $\underline{\mathbf{E}}$ es la hiper-matriz que contiene los residuos no ajustados por el modelo (Khanmohammadi, 2015). En la Figura 1.4 se representa esquemáticamente la descomposición de un mo-

delo PARAFAC para un arreglo de 3-vías.

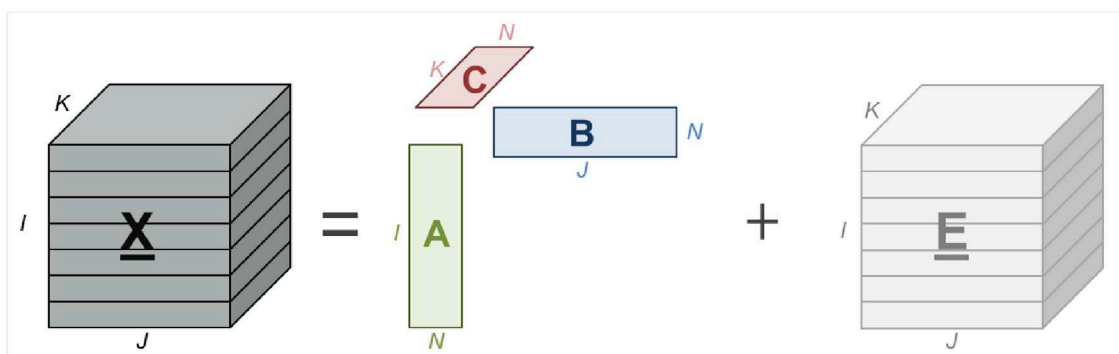


Figura 1.4. Modelo de descomposición PARAFAC para datos de segundo orden dispuestos en un arreglo de 3-vías para un sistema de múltiples muestras conteniendo N constituyentes

La principal ventaja del modelo PARAFAC es la unicidad de la solución. En los modelos bilineales basados en ALS, el mayor inconveniente es la libertad rotacional que presentan en su resolución, que puede ser significativamente reducida mediante la utilización de información externa (Bro, 1997; Martí *et al.*, 2015; Olivieri *et al.*, 2014b). Sin embargo, para los modelos PARAFAC no sería necesario el uso de restricciones en virtud de que el modelo es identificado únicamente por su estructura, si bien en la práctica pueden ser utilizadas algunas restricciones a fin de favorecer la obtención de la solución óptima. Las matrices de *loadings* **B** y **C** obtenidas de la descomposición pueden ser inferidas, entonces, a alguna propiedad física, es decir, pueden ser físicamente interpretables. Por ejemplo, para EEM de fluorescencia, los perfiles contenidos en **B** y **C** representan los espectros de excitación y emisión, mientras que la información contenida en **A** indica las concentraciones relativas de los constituyentes de la muestra. La unicidad de la solución, sin embargo, solo se alcanza si no existe deficiencia de rango en ninguno de los modos, es decir, el número de componentes es igual en cada uno de los modos instrumentales y los vectores representativos del modelo son linealmente independientes (Bro, 1997). La información contenida en **A** es utilizada, habitualmente, para análisis predictivos en calibración multivariada.

ANÁLISIS PARALELO DE FACTORES DE 3-VÍAS AUMENTADO (APARAFAC)

Este modelo ha sido desarrollado para el análisis de datos de tercer orden que presentan desvíos en la cuadrilinealidad. En principio, el modelo APARAFAC implica la construcción de un arreglo de 3-vías trilineal aumentado, donde la operación de aumentación se realiza a lo largo del modo en el que los perfiles presenten cambios entre muestras, es decir, el modo que rompe la cuadrilinealidad del arreglo (Bortolato *et al.*, 2015). La principal ventaja de esta estrategia es que el modelo puede, en principio, tener descomposición única en los 3 modos instrumentales sin aplicar ninguna restricción, análogo a un modelo PARAFAC (Bro, 1997).

Las dimensiones del arreglo de 3-vías aumentado dependerán del modo en que se realizó la aumentación, por ejemplo, para un sistema cromatográfico con detección de EEM de fluorescencia, un arreglo aumentado en el modo correspondiente a los tiempos de retención tendrá dimensiones $P \times K \times L$ ($P=IJ$), donde I , J , K y L son el número de muestras, los tiempos de retención, las longitudes de onda de excitación y de emisión, respectivamente. El correspondiente modelo APARAFAC puede ser representado, entonces, como

$$x_{ap,pkl} = \sum_{n=1}^N a_{ap,pn} b_{ap,kn} c_{ap,ln} + e_{ap,pkl} ; \quad 1.12$$

$$k = 1, 2, \dots, K; l = 1, 2, \dots, L; p = 1, 2, \dots, P = IJ$$

Así, de la descomposición de $\underline{\mathbf{X}}_{ap}$ se obtienen 3 matrices de *loadings* \mathbf{A}_{ap} , \mathbf{B}_{ap} y \mathbf{C}_{ap} de dimensiones $P \times N$, $K \times N$ y $L \times N$, respectivamente, siendo N el número de componente presentes en la muestras. Los residuos del modelo son colectados en la hiper-matriz aumentada $\underline{\mathbf{E}}_{ap}$. Considerándose el ejemplo mencionado anteriormente, la matriz \mathbf{A}_{ap} contendrá los perfiles aumentados del modo correspondiente a los tiempos de retención.

APARAFAC puede ser implementado mediante un proceso típico de cuadrados mínimos alternantes y por lo tanto requiere de la inicialización del modelo ya sea utilizando estimaciones iniciales obtenidas mediante un algoritmo auxiliar o utilizando información externa disponible.

El modelo APARAFAC para un arreglo de 3-vías aumentado descrito por la

ecuación 1.12 es único y no requeriría el uso de restricciones para lograr la descomposición. Sin embargo, ciertas restricciones son necesarias para la lograr una resolución correcta. Esta estrategia de análisis incluye la opción de limitar al sistema con restricciones similares a las utilizadas en MCR-ALS. Adicionalmente, al igual que para MCR-ALS y PARAFAC, los perfiles obtenidos de la descomposición por APARAFAC presentan interpretación física, permitiendo, así, inferir sobre las características físico-químicas del sistema. En la Figura 1.5 se representa esquemáticamente la descomposición de un modelo APARAFAC para datos de tercer orden dispuestos en un arreglo de 3-vías aumentado.

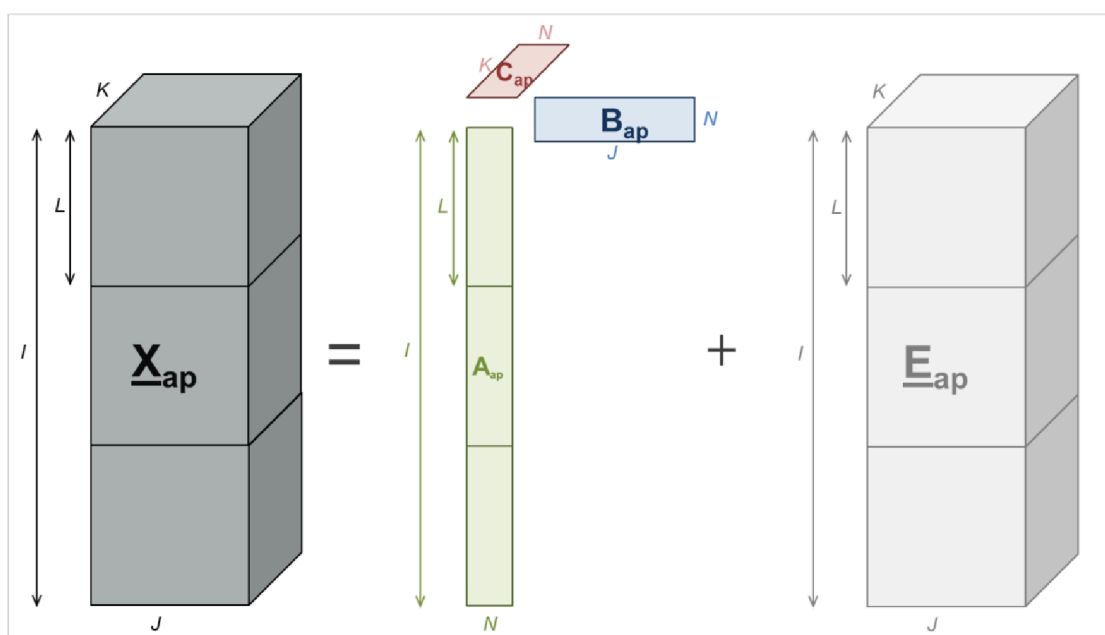


Figura 1.5. Modelo de descomposición APARAFAC para datos de tercer orden dispuestos en un arreglo de 3-vías aumentado para un sistema de múltiples muestras conteniendo N constituyentes

OTRAS HERRAMIENTAS QUIMOMÉTRICAS Y DE ANÁLISIS

Método simple interactivo de auto-modelado para el análisis de mezclas (SIMPLISMA). El método simple interactivo de auto modelado para el análisis de mezclas (*SIMPLISMA: Simple-to-use Interactive Self-modelling Mixture Analysis*) fue creado y desarrollado por Windig en el año 1991 para la resolución de datos multidimensionales de manera simple, rápida y de fácil interpretación (Windig, 2009). La idea central del algoritmo es asumir que en los datos hay variables que experimentan contribuciones de un único componente de la mezcla. Estas variables

son denominadas como “variables puras”. Si en un sistema multi-componente hay al menos 1 variable pura en sus datos, entonces el espectro del componente correspondiente puede ser resuelto en base a una resolución por cuadrados mínimos (Bogomolov *et al.*, 2005).

Para encontrar las variables puras, SIMPLISMA se basa en una función de pureza que sigue la expresión (Bogomolov *et al.*, 2005)

$$p_{f1} = \frac{\sigma_j}{\mu_j + \alpha} \quad 1.13$$

donde p_j es el primer valor de pureza de la variable j ; σ_j es el desvío estándar de la variable j ; μ_j es la media calculada para la variable j y α es una constante de desviación. El desvío estándar y la media de la variable j se estiman siguiendo las ecuaciones 1.14 y 1.15, respectivamente.

$$\sigma_j = \sqrt{\frac{1}{c} \sum_{i=1}^c (d_{ij} - \mu_j)^2} \quad 1.14$$

$$\mu_j = \frac{1}{c} \sum_{i=1}^c d_{ij} \quad 1.15$$

donde d_{ij} representa los elementos de la matriz de datos con los espectros en las filas y c es el número de espectros. El segundo subíndice de p en la ecuación 1.13 indica que es el primer valor de pureza.

Widing dedujo que la relación entre el desvío estándar y la media para la misma variable j se correlaciona con la pureza de la variable. Debido a que la pureza, el desvío estándar y la media tienen el mismo número de variables como espectros presentes, las variables puras pueden ser representadas como curvas espectrales similares a lo largo del eje espectral que pueden, entonces, ser fácilmente interpretadas. La primera variable pura es tomada como un máximo de pureza del primer espectro (Kucheryavskiy *et al.*, 2016).

De acuerdo a la expresión de pureza, cuando el valor de la media tiende a cero,

la relación entre éste y la desviación estándar tiende a infinito. Por esta razón se adiciona una constante al denominador, α , que se define como un porcentaje de la media y suprime la sobreestimación de la pureza. SIMPLISMA resuelve los componentes de manera individual. La variable pura siguiente es aquella que es más independiente de la primera variable pura. Finalmente, los componentes resueltos por SIMPLISMA pueden ser utilizados como estimaciones iniciales en MCR-ALS.

Análisis de factores evolutivos (EFA). El análisis de factores evolutivos o que evolucionan (*EFA: Evolving Factor Analysis*) es un método de análisis de rango local que detecta el aumento o decaimiento de los componentes en un conjunto de datos y provee perfiles de concentración, asumiendo un perfil de aumento-decaimiento secuencial de todos los componentes del sistema (Tauler *et al.*, 2015). Un aspecto importante de este método es el hecho de que el análisis es completamente libre de modelo, es decir, no se requieren presunciones acerca del proceso químico bajo estudio, si bien se requiere de hechos puramente físicos, tales como una relación lineal entre concentración de las especies y su señal instrumental (Maeder *et al.*, 2009). Inicialmente, se ha desarrollado para determinar constantes de equilibrio de curvas de titulación espectrofotométricas y ha demostrado ser una herramienta muy poderosa para estudios de equilibrio químico y análisis cromatográficos.

La idea fundamental de EFA es seguir el cambio o evolución del rango de una matriz de datos \mathbf{X} como función de una variable determinada (Keller *et al.*, 1992). La primera etapa se basa en la observación de la evolución dependiente del tiempo del rango total de una matriz de datos, brindando una ventana de concentraciones de los componentes involucrados. El segundo paso se basa en una regresión lineal no iterativa usando la ventana de información obtenida, resultando en perfiles de concentración y espectros para todos los componentes involucrados en el sistema (Maeder *et al.*, 1988).

El único prerequisite del método es que se obedezca la ley de Lambert-Beer. Adicionalmente, no es necesaria información acerca de la identidad de los componentes en ninguno de los modos. Una aplicación típica de este método es la utilización de la ventana de información como estimaciones iniciales a los procesos de optimización ALS (Maeder *et al.*, 2009).

Correlación optimizada de deformaciones (COW). Las herramientas de alineamiento de señales y corrección de deformaciones son utilizadas

para corregir datos de diferente naturaleza, por ejemplo, datos cromatográficos o espectrales. Los algoritmos de corrección de deformaciones aplican procesos de estiramiento y compresión en sentido horizontal, corrigiendo las señales por efecto elástico. Dentro de las técnicas desarrolladas para la corrección de deformaciones se encuentra el algoritmo de correlación optimizada de deformaciones (*COW: Correlation Optimized Warping*). Este algoritmo trabaja en segmentos de datos manteniendo los puntos cercanos consecutivos y opera sobre un vector de datos completo.

En COW se requieren 2 parámetros iniciales que son establecidos por el analista: **1- tamaño del segmento, “s”;** y **2- grado de elasticidad, “l”.** Considerándose 2 espectros, uno se selecciona como referencia, mientras que el restante es subdividido en secciones de longitud *s*, que, por efecto elástico, son iterativamente comprimidas o expandidas por interpolación horizontal hasta un máximo de “*l*” puntos, lográndose, así, la máxima correlación entre espectros. Este tipo de corrección no distorsiona la intensidad de la señal debido a que la corrección se produce en el eje horizontal. El principio del algoritmo es hacer coincidir las señales de todas las muestras a una señal común de referencia (Jellema, 2009).

1. 6. CALIBRACIÓN MULTIVARIADA

CONCEPTOS BÁSICOS

Booksh y Kowalski (Booksh *et al.*, 1994) definen a la calibración como “*un proceso matemático y estadístico que permite la extracción de información acerca de la concentración de los analitos a partir de una señal instrumental*”. Por otro lado, la Unión Internacional de Química Pura y Aplicada (*IUPAC: International Union of Pure and Applied Chemistry*) (Danzar *et al.*, 1998) establece que la calibración es “*una operación que relaciona un cantidad de salida a una cantidad de entrada para un sistema de medición en condiciones establecidas*”, que en términos de química analítica significa relacionar una respuesta analítica instrumental a la concentración del constituyente de interés en una muestra. Los métodos de calibración pueden ser clasificados depen-

diendo del tipo de dato generado o del arreglo construido.

En primer lugar, a fin de comprender las propiedades de la calibración multivariada, es importante definir algunos conceptos que refieren al tipo de muestra y al tipo de constituyente (Escandar *et al.*, 2007; Olivieri *et al.*, 2014b). Las muestras pueden ser divididas en:

Muestras de calibración: empleadas para establecer la relación entre la concentración conocida del analito y la señal instrumental registrada;

Muestras de validación: muestras compuestas de analitos con concentración conocida empleadas para evaluar el desempeño analítico de la calibración en condiciones establecidas. Deben presentar composición química cualitativamente similar a las muestras de calibración, si bien las concentraciones de los analitos deben ser diferentes a las utilizadas en la etapa de calibración;

Muestras de prueba o ensayo: muestras que pueden contener constituyentes adicionales a los presentes en las muestras de calibración y validación y son utilizadas para comprobar la capacidad de la calibración multivariada para distinguir la presencia de componentes extraños;

*Muestras reales: muestras de composición similar a las muestras de prueba de las que no se dispone información acerca de la concentración de los analitos. Ayudan a evaluar la exactitud y precisión de la calibración multivariada en comparación con técnicas de referencia. (Olivieri *et al.*, 2014b)*

En cuanto a los constituyentes que componen la muestra, la clasificación se puede realizar como (Olivieri, 2014):

Constituyentes esperados calibrados: constituyentes principales de las muestras de calibración y validación que son adicionados por el analista en concentraciones conocidas;

Constituyentes esperados no calibrados: constituyentes de las muestras de calibración y validación que presentan señal instrumental pero se carece de información acerca de su concentración;

Constituyentes no esperados: componentes presentes en las muestras de prueba y muestras reales y que pueden interferir en el análisis. En calibración univariada y calibración de datos de primer orden, estos constituyentes interfieren generando un error sistemático en la determinación del analito, mientras que para calibración multivariada de segundo orden y orden superior no siempre resultan en un inconveniente.

El proceso general de calibración relaciona el contenido de un analito único presente en la muestra a un valor único de la señal instrumental. A este proceso de calibración se lo denomina “calibración univariada”. Por otro lado, los procesos de calibración que involucran la relación entre las concentraciones de varios constituyentes en una muestra con las múltiples respuestas son comúnmente denominados procesos de “calibración multivariada” y se clasifican, como se mencionó anteriormente, según el orden de los datos multidimensionales generados, por ejemplo, calibración multivariada de primer orden para datos de primer orden, calibración multivariada de segundo orden para datos de segundo orden, y así sucesivamente.

La ventaja más evidente de los métodos basados en calibración multivariada es la determinación cuantitativa simultánea de analitos en un sistema multi-componente. Sin embargo, las ventajas más relevantes de las técnicas de calibración multivariada son las relacionadas a la presencia de constituyentes no esperados en las muestras de prueba y muestras reales, llamados interferentes. La ventaja de primer orden, que es propia de los métodos de calibración multivariada de primer orden, es aquella relacionada a la capacidad de distinguir muestras anómalas de un conjunto de muestras en virtud de las diferencias en su composición respecto a las muestras de calibración. Por otro lado, la ventaja relacionada al hecho de que los analitos pueden ser determinados en presencia de constituyentes no esperados en las muestras de prueba es llamada ventaja de segundo orden y es propia de la calibración multivariada de segundo orden u orden superior (Olivieri *et al.*, 2014b). Por último, la ventaja de tercer orden, relacionada a la calibración multivariada de tercer orden u orden superior, no ha sido aún establecida en la comunidad científica, si bien algunos autores la definen como la capacidad de resolver un problema complejo en un arreglo de datos correspondiente a una muestra única, en independencia de muestras adicionales, el aumento de la sensibilidad y la selectividad y resolución de problemas de co-linealidad entre señales (Wu *et al.*, 2015).

En otro orden, se ha demostrado que el aumento en el orden de los datos instrumentales tiene un impacto positivo en las cifras de mérito que caracterizan a los métodos analíticos (Olivieri, 2014).

CIFRAS DE MÉRITO ANALÍTICAS

En química analítica, las cifras de mérito analíticas (*FOMs: Figures Of Merit*) se

definen como parámetros numéricos necesarios para la comparación de rendimientos relativos de diferentes metodologías analíticas, cifras útiles para discriminar sus capacidades de detección (Olivieri *et al.*, 2014a) y que forman parte integral de los protocolos de análisis.

La sensibilidad (*SEN*) es clave elemental en la estimación de otras FOMs, como lo son: **1-** sensibilidad analítica: importante para la comparación de metodologías basadas en diferentes señales instrumentales; **2-** selectividad (*SEL*): ayuda a evaluar la posibilidad de cuantificación de un analito en presencia de potenciales interferentes; y **3-** predicción de incertidumbres, límite de detección (*LOD: Limit Of Detection*) y límite de cuantificación (*LOQ: Limit Of Quantitation*): necesarias para evaluar las capacidades predictivas de un método.

En calibración univariada, la predicción de la concentración de un analito (*y*) en una muestra de prueba a partir de su señal instrumental (*x*) se realiza a través de la expresión clásica obtenida de una curva de regresión

$$y = \frac{(x - n_0)}{m_0} \quad 1.16$$

donde n_0 y m_0 representan la ordenada al origen y la pendiente, respectivamente, obtenidas del gráfico de regresión de calibración univariada. En este tipo de calibración, la pendiente, m_0 , representa la *SEN* del método. Sin embargo, en métodos de calibración de orden superior el concepto de señal analítica neta (*NAS: Net Analyte Signal*) resulta útil para evaluar la *SEN* del método (Olivieri, 2014).

Una definición alternativa del concepto de *SEN* puede ser dada en términos de propagación de incertidumbres: “*la SEN es considerada como una medida del grado de ruido de salida de un sistema para un dado ruido de entrada*” (Fragoso *et al.*, 2016; Olivieri, 2014), entonces, una mayor *SEN* se obtiene si el ruido de salida es inferior al de entrada. Así, el parámetro de *SEN* se define como

$$SEN = \frac{\sigma_x}{\sigma_y} \quad 1.17$$

donde σ_x y σ_y son las incertidumbres en la señal instrumental y la concentración, respectivamente (Olivieri *et al.*, 2014a). Esta estrategia de propagación de incertidumbres asume que el ruido de entrada es independiente e idénticamente distribuido.

Se ha demostrado, además, que el valor de SEN depende del orden de los datos instrumentales y depende, también, del modelo matemático seleccionado para el análisis de los datos. La expresión general de SEN derivada del concepto de propagación de incertidumbres presenta la siguiente estructura,

$$\text{SEN} = \left\{ \mathbf{g}_n^T [\mathbf{Z}_{\text{esp}}^T (\mathbf{I} - \mathbf{Z}_{\text{noesp}} \mathbf{Z}_{\text{noesp}}^+) \mathbf{Z}_{\text{esp}}]^{-1} \mathbf{g}_n \right\}^{-1/2} \quad 1.18$$

donde cada parámetro dependerá del modelo matemático utilizado y el orden de los datos, según se detalla en la Tabla 1.1. La matriz \mathbf{Z}_{esp} y el vector especial de la identidad del analito, \mathbf{g}_n , corresponden a la fase de calibración; los subíndices “esp” y “noesp” refieren a componentes esperados y no esperados, respectivamente. La matriz \mathbf{Z}_{esp} colecta los perfiles de los constituyentes presentes en la muestra de calibración y el vector \mathbf{g}_n selecciona o combina esta información, haciéndola específica para el constituyente de interés n . El factor final de la ecuación 1.18, $(\mathbf{I} - \mathbf{Z}_{\text{noesp}} \mathbf{Z}_{\text{noesp}}^+)$, es la manifestación matemática de la ventaja de segundo orden y solo aparece en metodologías de calibración de segundo orden u orden superior.

Así, por ejemplo, en el caso de datos multidimensionales de segundo orden, dispuestos en un arreglo de 2-vías aumentado, aplicando un modelo MCR-ALS, la SEN se deriva en la expresión (Bauza *et al.*, 2012):

$$\text{SEN}_{\text{MCR}} = p_k \left[J(\mathbf{S}^T \mathbf{S})_{kk}^{-1} \right]^{-1/2} \quad 1.19$$

siendo k el índice del analito de interés en la mezcla multi-componente, J el número de datos en cada sub-matriz en el modo aumentado y p_k la pendiente de la curva estimada en la regresión pseudounivariada construida con los perfiles de concentración obtenidos posterior al modelado por MCR-ALS. Se puede observar que la SEN, en este caso, depende del perfil espectral calculado \mathbf{S} .

Por último, las estimaciones de LOD y LOQ se obtienen según las ecuaciones 1.20 y 1.21, respectivamente (Bauza *et al.*, 2012):

$$\text{LOD}_{\text{MCR}} = 2t_{\alpha, \nu} s_c = \frac{2t_{\alpha, \nu} s d_{\text{test}}}{\text{SEN}_{\text{MCR}}} \quad 1.20$$

$$\text{LOQ}_{\text{MCR}} = 10s_c = 10 \frac{\text{LOD}_{\text{MCR}}}{2t_{\alpha, \nu}} \quad 1.21$$

para una distribución t no central con $\nu = n - 2$ grados de libertad y una probabilidad del 95 % (nivel $\alpha = 0.05$). La desviación estándar de concentración predicha y la desviación estándar estimada cuando el valor real de la señal neta es cero, vienen representadas mediante s_c a sd_{test} , respectivamente. Esta última también se define como el nivel de ruido instrumental presente en el sistema.

Tabla 1.1. Parámetros utilizados en la expresión de sensibilidad de diferentes métodos de calibración

Modelo	Orden	Parámetro			Detalle
		\mathbf{g}_n	$\mathbf{Z}_{\text{noesp}}$	\mathbf{Z}_{esp}	
Univariado	0	1	m_0	–	m_0 = pendiente del gráfico univariado δ_n = vector Kronecker del analito n J = número de sensores del modo aumentado en cada sub-matriz del arreglo \mathbf{D}_{aum} m_n = pendiente del gráfico pseudounivariado ^a $\mathbf{C}_{\text{noesp}}$ = perfiles en modo no aumentado de los constituyentes esperados en las muestras de calibración
MCR-ALS	2	δ_n	$(m_n/J^{1/2})\mathbf{C}_{\text{esp}}$	$\mathbf{C}_{\text{NOESP}}$	$\mathbf{C}_{\text{noesp}}$ = perfiles en modo no aumentado de los constituyentes no esperados Segundo orden: $\mathbf{Z}_{\text{esp}} = m_0\mathbf{C}_{\text{esp}} \odot \mathbf{B}_{\text{esp}}$ Tercer orden: $\mathbf{Z}_{\text{esp}} = m_0\mathbf{D}_{\text{esp}} \odot \mathbf{C}_{\text{esp}} \odot \mathbf{B}_{\text{esp}}$
PARAFAC	2, 3, ...	δ_n			

^a Pendiente extraída de la curva de calibración pseudounivariada construida a partir de las contribuciones individuales de los analitos obtenidas del modelo quimiométrico.

Es importante tener en cuenta que para estudios cuantitativos es posible obtener información cuantitativa a partir de los perfiles obtenidos de la descomposición de los datos multidimensionales, pudiéndose construir modelos de regresión. Para MCR-ALS esta información se obtiene a partir de los perfiles de las matrices \mathbf{C} que, por ejemplo, para datos cromatográficos, se extrae como las áreas correspondientes a las señales cromatográficas o picos cromatográficos. En PARAFAC, en cambio, esta información es extraída directamente de la matriz de \mathbf{A} , obtenida del modelado quimiométrico. De

esta manera, según se obedezca la ley de Lambert-Beer, las áreas obtenidas de MCR-ALS o los valores contenidos en **A** obtenidos de PARAFAC serán proporcionales a la concentración del componente de interés, pudiéndose construir curvas de regresión para cada analito individualmente, denominándose regresión pseudounivariada.

Finalmente, otra manera de estimar las FOMs para datos de segundo orden es utilizar la información contenida en los perfiles puros obtenidos del modelado quimiométrico para cada analito del sistema multi-componente y estimarlos siguiendo una calibración univariada clásica (Saurina *et al.*, 2001). De esta manera, la SEN estará definida como el valor de la pendiente de la curva lineal ajustada de la regresión pseudounivariada, resultado de graficar las respuestas relativas obtenidas del modelado respecto a las concentraciones conocidas de los constituyentes de la calibración. Luego, las estimaciones de LOD y LOQ se obtienen siguiendo las ecuaciones 1.20 y 1.21, reemplazando s_c por s_0 , representando esta última la desviación estándar de la señal neta estimada cuando el valor real es cero. Aquí, el cálculo de las FOMs no dependen del valor de SEN (Clayton *et al.*, 1987; ICH, 2005).

E2

0 bjetivos



2. OBJETIVOS

“One is never afraid of the unknown; one is afraid of the known coming to an end”

Krishnamurti, J.

2. 1. OBJETIVOS GENERALES

Evaluar nuevas estrategias de obtención de datos multidimensionales y explorar las potencialidades de los algoritmos quimiométricos para el análisis de muestras complejas de composición variada. Desarrollar nuevos métodos analíticos para estudios descriptivos y/o predictivos de sistemas químicos de diferente naturaleza, combinando diversas metodologías analíticas con modelado quimiométrico de datos multidimensionales.

2. 2. OBJETIVOS ESPECÍFICOS

Generar datos de segundo orden utilizando técnicas separativas acopladas a detección espectral para el análisis de contaminantes emergentes en muestras acuosas.

Generar datos de segundo orden utilizando técnicas de inyección en flujo acoplada a detección espectral para la determinación de fármacos en muestras biológicas complejas.

Generar datos de segundo orden utilizando técnicas espectroscópicas infrarrojas para el análisis conformacional de macromoléculas en el proceso de desnaturalización inducido por alcohol en medio acuoso. Desarrollar y poner a punto de una nueva metodología analítica basada en espectroscopia laser infrarroja para el análisis de proteínas y polipéptidos en medio acuoso y medio deuterado.

Generar datos de segundo orden utilizando técnicas de dicroísmo circular y espectroscopia de absorción infrarroja para el análisis conformacional de macromoléculas en un proceso térmico en medio deuterado. Evaluar el comportamiento de la transición conformacional de un polipéptido en un rango amplio de concentraciones por combinación de diferentes técnicas analíticas: *dicroísmo circular, espectroscopia laser infrarroja y espectroscopia infrarroja acoplada a transformada de Fourier*.

Generar datos de tercer orden utilizando técnicas cromatográficas acopladas a detección de matrices de excitación-emisión de fluorescencia, evaluando los posibles arreglos instrumentales para el análisis.

Evaluar y combinar diferentes estrategias de desarrollo de experimentos, pre-procesamiento de datos y análisis a fin de explotar la potencialidad de los métodos analíticos desarrollados en su conjunto para obtener la máxima información relevante de los sistemas químicos bajo estudio.

03

0

**Obtención
y análisis
de datos**

3. OBTENCIÓN Y ANÁLISIS DE DATOS – APLICACIONES

“Don’t think about why you question, simply don’t stop questioning. Try to comprehend a little more each day. Have holy curiosity”

Einstein, A.

3. 1. MÉTODO CROMATOGRÁFICO PARA LA DETERMINACIÓN DE FLUOROQUINOLONAS EN MEDIO ACUOSO BASADO EN LA FORMACIÓN DE COMPLEJOS METÁLICOS Y MODELADO DE DATOS MULTIDIMENSIONALES

Enhanced fluorescence sensitivity by coupling yttrium-analyte complexes and three-way fast high-performance liquid chromatography data modelling – Alcaráz, M.R., et al., Anal Chim Acta 2016, 902, 50-58

CONSIDERACIONES GENERALES

En la década del 60', una nueva familia de fármacos sintéticos, denominada quinolonas (*QUI*), ha sido desarrollada para el tratamiento de infecciones bacterianas en humanos (Sukul *et al.*, 2007). A pesar de su efectividad, el rápido desarrollo de resistencia microbiana a los antibióticos ha obligado a la industria farmacéutica a investigar nuevas estructuras quinolónicas a fin de aumentar y mejorar su espectro de acción bactericida, como también, sus propiedades farmacocinéticas. En la década del 80', una modificación clave en la estructura química base del fármaco, basada en el agregado de un grupo flúor en el anillo aromático, dio origen a una subfamilia de QUI, denominada fluoroquinolonas (*f-QUI*), una clase importante de antibióticos que ha sido extensamente utilizada en los últimos 30 años (Uivarosi, 2013) para el tratamiento de infecciones, ya no solo en humanos sino en animales domésticos, aves de corral, animales de granja y peces. En la Figura 3.1 se representa la estructura química general básica de las f-QUI.

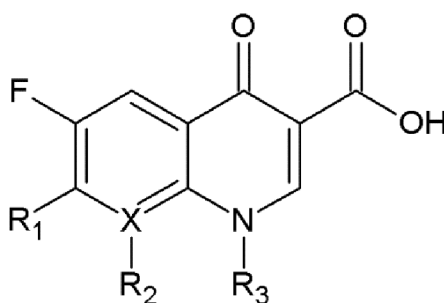


Figura 3.1. Estructura química general de las fluoroquinolonas

Las f-QUI son agentes bactericidas muy potentes utilizados generalmente como segunda línea de defensa en humanos. Actúan principalmente sobre la inhibición de las enzimas ADN girasa y ADN topoisomerasa IV bloqueando la replicación y transcripción del ADN bacteriano por interacción de sus funciones carbonilo, carboxilo y flúor con residuos de ácido aspártico, serina y lisina de las enzimas, llevando a una rápida muerte celular (Sadeek *et al.*, 2011; Uivarosi, 2013). Dentro de esta familia de antibióticos, ciprofloxacina (*CPF*), ofloxacina (*OFL*) y norfloxacina (*NRF*) han sido aprobadas para el uso en humanos, mientras que enrofloxacina (*ENF*) y sarafloxacina (*SRF*) se aplican exclusivamente en medicina veterinaria. Es importante destacar que el Centro para la Medicina Veterinaria (*CVM: Center for Veterinary Medicine*), en dependencia de la Administración de alimentos y drogas de los Estados Unidos (*US FDA: United*

State-Food and Drug Administration), ha prohibido el uso de SRF en ganillas ponedoras y ha propuesto eliminar el uso de ENF en animales de granja debido a que contribuye al desarrollo de *Campylobacter* resistente a f-QUI en infecciones en humanos (Schaefer, 2002), como así también, el retiro absoluto de esta última en el agua de aves de corral (Crawford, 2005).

Las f-QUI se utilizan ampliamente para el tratamiento de diversas afecciones y, en medicina veterinaria, son intensamente administradas en grandes cantidades. Estos compuestos no son completamente metabolizados por el organismo y se descargan al ambiente como compuesto parental o como subproductos de ellos. En consecuencia, y debido a la gran ocurrencia en fuentes ambientales, estos compuestos son denominados *contaminantes emergentes* y representan un gran interés social y medioambiental. Su interés radica, principalmente, en el desconocimiento de los efectos que pudiesen ocasionar en la salud si persisten en el ambiente incluso en bajos niveles de concentración (Khetan *et al.*, 2007), teniendo en cuenta, además, que estos fármacos sintéticos presentan una alta resistencia a la biodegradación (He *et al.*, 2015). Es por todo esto que el monitoreo y la determinación de estos compuestos a nivel de trazas en fuentes ambientales resulta, hoy en día, indispensable para el control de la salud humana y medioambiental.

En la literatura es posible encontrar un gran número de métodos para la determinación de f-QUI en diversas matrices ambientales, incluyendo métodos cromatográficos con detección de fluorescencia (*FD: Fluorescence Detection*), detección por espectroscopia de masas o detección UV-Vis con DAD, obteniéndose datos de segundo orden u orden superior (Escandar *et al.*, 2014). En este contexto, las técnicas de calibración multivariada han sido implementadas para el modelado de datos multidimensionales, lográndose una considerable mejora en las propiedades analíticas del método. Adicionalmente, es posible encontrar métodos basados en sondas luminiscentes que involucran el uso de iones metálicos (Rodríguez-Díaz *et al.*, 2003; Wang *et al.*, 2009; Wang *et al.*, 2006) o nanopartículas metálicas (Liu *et al.*, 2014; Yáñez-Jácome *et al.*, 2015), aprovechándose sus particulares características para lograr aumentar la sensibilidad y la selectividad del método, entre las que se destacan, bandas espectrales estrechas, tiempo de vida luminiscente prolongado, amplios corrimientos de Stokes, fuerte capacidad de combinación con otros compuestos (Lakowicz, 2006; Tong *et al.*, 2010a). En los últimos años, las sondas luminiscentes han demostrado ser una excelente alternativa para la determinación de diferentes compuestos a nivel de trazas en matrices de diversa complejidad.

Una propiedad química interesante de las QUI es su capacidad de unirse a iones

metálicos de valencia superior para formar complejos. La presencia de un grupo carboxilo contiguo a un carbonilo cetónico permite la formación de quelatos metálicos mediante acción de la QUI como ligando bidentado, si bien, a través de otros grupos funcionales, puede actuar también como ligando unidentado o como ligando puente (Sadeek *et al.*, 2011; Uivarosi, 2013). Por análisis espectroscópico (UV-Vis, IR y NMR) y análisis teórico computacional (*TD-DFT: Time-Dependent-Density Functional Theory*) se ha demostrado que la formación del complejo metálico entre las QUI e iones metálicos como Co^{2+} , Ni^{2+} , Y^{3+} y Zn^{4+} , se genera a través del átomo de oxígeno deprotonado del grupo carboxílico y el átomo de oxígeno carbonilo del anillo de las QUI en estado zwitteriónico (Drevenšek *et al.*, 2006; Patel *et al.*, 2008; Sadeek *et al.*, 2011). Esto sugiere que la formación del complejo metálico se encuentra altamente afectada por las modificaciones del pH del medio en virtud de las propiedades acido-base del fármaco. En la Figura 3.2 se muestra el modelo de coordinación propuesto para los complejos entre las QUI con metales polivalentes, en el ejemplo de CPF- Y^{3+} .

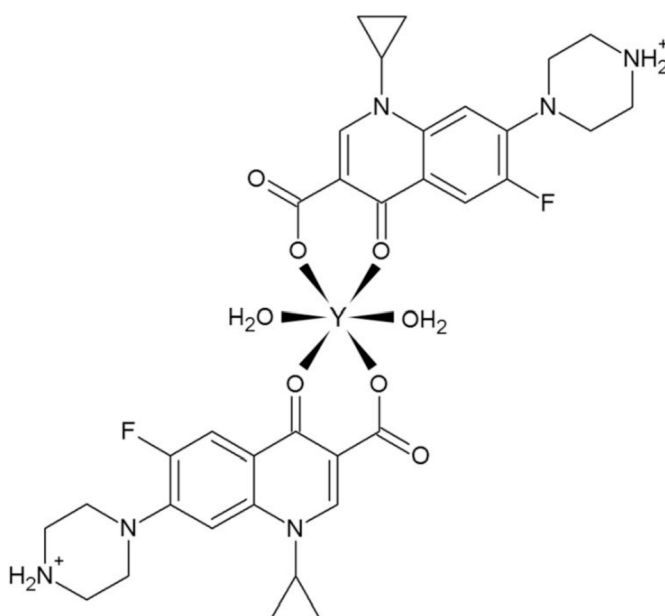


Figura 3.2. Modelo de coordinación propuesto de la formación del complejo metálico entre Y^{3+} y CPF

La formación de estos complejos metálicos no solo afecta la actividad antibacteriana y farmacocinética del fármaco, sino que también modifica sus propiedades ópticas (Sadeek *et al.*, 2011). En trabajos recientes se ha demostrado que la presencia del metal itrio, en su estado oxidado, Y^{3+} , puede incrementar notablemente la fluorescencia nativa de varios de estos compuestos y modificar su espectro UV y de fluorescencia.

El ion Y^{3+} en solución no presenta fluorescencia nativa, pero en presencia de ciertos compuestos, como las QUI, es capaz de exaltar la señal de fluorescencia nativa del cromóforo (Han *et al.*, 2005; Tong *et al.*, 2010a; Tong *et al.*, 2010b). La formación del complejo QUI- Y^{3+} conlleva tanto un aumento de la señal fluorescente del analito como un cambio en la posición del máximo de intensidad (Han *et al.*, 2005; Tong *et al.*, 2010a; Tong *et al.*, 2010b). Esto también se ha demostrado para los espectros de absorbancia, en donde la presencia del metal induce un cambio en la posición del máximo de absorbancia correspondiente (Sadeek *et al.*, 2011).

Con base en lo mencionado, aprovechándose la particularidad de las QUI de formar complejos metálicos y las características de las sondas luminiscentes, se propone un método analítico para la determinación de 7 f-QUI a nivel de trazas en agua de recursos naturales, incluyendo un sistema basado en complejos f-QUI- Y^{3+} , por cromatografía líquida de alto rendimiento (*HPLC: High Performance Liquid Chromatography*) y detección de fluorescencia de barrido rápido (*HPLC-FSFD*). Además, a fin de obtener mejoras en las propiedades analíticas, se utilizaron herramientas quimiométricas de calibración multivariada para el modelado de los datos obtenidos.

MATERIALES Y MÉTODOS

Reactivos. Todos los estándares utilizados fueron de calidad grado analítico. Las drogas sólidas de enoxacina (*ENO*), norfloxacin (NRF), ofloxacin (OFL) y sarafloxacin (SRF) fueron provistas por Sigma-Aldrich (Steinheim, Alemania). Ciprofloxacin (CPF), difloxacin (DIF) y enrofloxacin (ENF) fueron adquiridas de Fluka (Buchs, Suiza). El acetonitrilo (ACN) calidad cromatografía líquida fue obtenido de LiChrosolv (Merk Millipore Co., Darmstadt, Alemania). El ácido acético glacial (HAc) fue adquirido de Merk (Darmstadt, Alemania) y el acetato de sodio trihidrato (NaAc) calidad p.a. fue provisto por Anedra (La Plata, Argentina). El nitrato de itrio hexahidrato ($Y(NO_3)_3 \cdot 6H_2O$) fue adquirido de Sigma-Aldrich (Steinheim, Alemania). El agua de laboratorio ultrapura fue obtenida por un sistema de purificación de agua Milli-Q Milli-Pore (Billerica, Estados Unidos).

Soluciones de trabajo. Se preparó una solución madre de cada f-QUI mediante disolución de la cantidad apropiada de cada analito en metanol alcalinizado

(pH~9.0) a fin de obtener una concentración final de $200.0 \mu\text{g mL}^{-1}$. La solución madre de itrio (Y^{3+}) de concentración 0.1 mol L^{-1} de Y^{3+} se preparó disolviendo la cantidad apropiada de $\text{Y}(\text{NO}_3)_3 \cdot 6\text{H}_2\text{O}$ en agua ultrapura. Para su conservación, todas las soluciones se almacenaron a $4 \text{ }^\circ\text{C}$ protegidas de la luz por un periodo no mayor a 2 meses.

Para la determinación cromatográfica, se preparó una solución reguladora de HAc/NaAc disolviendo una cantidad apropiada de NaAc en agua ultrapura y ajustando el pH a un valor de 4.0 con HAc glacial, lográndose una concentración formal igual a 0.02 mol L^{-1} . A esta solución se le adicionó 1.0 mL de solución madre de Y^{3+} , obteniéndose una concentración final de Y^{3+} de $1.0 \times 10^{-4} \text{ mol L}^{-1}$. Finalmente, se completó el volumen de 1000.0 mL con agua ultrapura (esta solución será referida en el texto como AcYB). Previamente a ser usada como fase móvil cromatográfica, la solución se filtró a través de una membrana de nylon con tamaño de poro de $0.45 \mu\text{m}$ (Sartorius, Goettingen, Alemania).

Muestras de calibración y validación. Para la calibración se preparó un set de 5 soluciones estándares en diferentes niveles de concentración para cada f-QUI individualmente. Las soluciones se prepararon transfiriendo una alícuota apropiada de solución madre del analito correspondiente y $5.0 \mu\text{L}$ de solución madre de Y^{3+} a matraces aforados de 5.0 mL, completándose el volumen con agua ultrapura. Los rangos de concentración en los sets de calibración para cada f-QUI estuvieron comprendidos según se detallan a continuación: **1- CPF, ENF y NRF:** 0.0 ng mL^{-1} y 30.0 ng mL^{-1} ; **2- ENO:** 0.0 ng mL^{-1} y 200.0 ng mL^{-1} ; **3- DIF:** 0.0 ng mL^{-1} y 90.0 ng mL^{-1} ; **4- OFL:** 0.0 ng mL^{-1} y 36.0 ng mL^{-1} ; y **5- SRF:** 0.0 y 84.0 ng mL^{-1} . Considerando que, en las condiciones experimentales utilizadas, NRF/CPF y SRF/DIF presentan espectros de emisión de fluorescencia idénticos y resolución cromatográfica incompleta, se llevó a cabo una calibración conjunta siguiendo el procedimiento que se explica a continuación: *se preparó un set de calibración de 5 soluciones mezcla en diferentes niveles de concentración de NRF y CPF transfiriendo alícuotas apropiadas de solución madre de los respectivos analitos y $5.0 \mu\text{L}$ de solución madre de Y^{3+} a matraces aforados de 5.0 mL y completándose el volumen con agua ultrapura.* Siguiendo el mismo procedimiento, se llevó a cabo la calibración conjunta de SRF y DIF.

Para la validación, se preparó un set de 11 soluciones mezcla conteniendo todas las f-QUI analizadas en diferentes niveles de concentración, considerando valores distintos a los utilizados para la calibración, si bien comprendidos en los mismos rangos

de concentración. Las soluciones se prepararon siguiendo un modelo aleatorio, transfiriendo alícuotas apropiadas de cada solución madre de f-QUI y 5.0 μL de solución madre de Y^{3+} a matraces aforados de 5.0 mL y completándose el volumen con agua ultrapura.

La concentración final de Y^{3+} en cada solución estándar de calibración y solución mezcla de validación fue de $1.0 \times 10^{-4} \text{ mol L}^{-1}$. Todas las soluciones se prepararon diariamente, por triplicado.

Finalmente, previamente a la inyección al sistema cromatográfico, todas las soluciones se filtraron usando membranas de nylon con tamaño de poro de 0.45 μm (MSA, Osmonics Inc, Minnesota, Estados Unidos) y se transfirieron a viales de inyección de 2 mL. Seguidamente, se inyectó un volumen de 100 μL de solución en el sistema cromatográfico.

Muestras ambientales de matrices acuosas. Las muestras de agua analizadas se obtuvieron de 3 fuentes naturales distintas: **1-** *Muestras de agua superficial recogidas del arroyo Las Prusianas (Santa Fe, Argentina)*; **2-** *Muestras de agua de efluente recolectadas de diversos puntos de desagüe de la Facultad de Bioquímica y Ciencias Biológicas (Santa Fe, Argentina)*; y **3-** *Muestras de agua obtenidas de napa subterránea de la ciudad de Santa Fe apta para consumo humano (barrio distrito Colastiné, Santa Fe, Argentina)*. Las muestras se colectaron en botellas de 1 L de color ámbar y se procesaron de inmediato a su llegada al laboratorio; en los casos en los que no fue posible analizarlas de manera inmediata, se almacenaron a 4 °C protegidas de la luz por un lapso no mayor a 1 semana. A fin de remover sedimentos que pudiesen estar presente, las muestras se centrifugaron a una velocidad de 4000 r.p.m. durante 10 min y se filtraron por vacío a través de membranas de nylon con tamaño de poro de 0.45 μm .

Se prepararon 3 sets de 12 soluciones muestra que incluyeron a las soluciones muestra no fortificadas (o blanco de muestra) y las soluciones muestra fortificadas en 3 niveles de concentración para cada f-QUI. Las soluciones se prepararon siguiendo el mismo procedimiento descrito para las soluciones de validación, si bien el volumen se completó con agua de las muestras recolectadas.

A partir de este momento, las muestras ambientales de matrices acuosas serán referidas como muestras reales.

Instrumentación y procedimiento. Todos los experimentos cromatográficos se llevaron a cabo usando un instrumento Agilent 1100 LC (Agilent Technologies, Waldbronn, Alemania) equipado con desgasificador, bomba cuaternaria, inyector automático, compartimento termostatzado para columna, DAD UV-Vis, FSD y el programa informático ChemStation (Agilent Technologies, Waldbronn, Alemania) requerido para el control del instrumento y la adquisición de los datos.

La separación cromatográfica se realizó utilizando una columna analítica 3.5 μm Zorbax Eclipse XDB-C18 (75 mm \times 4.6 mm) (Agilent Technologies, Waldbronn, Alemania) como fase estacionaria, operando en modo isocrático con una velocidad de flujo de 2.20 mL min^{-1} , con una temperatura de compartimento para columna de 45 $^{\circ}\text{C}$, durante 16.0 min. La fase móvil consistió en una mezcla de solución AcYB y ACN en una proporción 91:9. Las matrices de datos tiempo de retención-espectro de emisión de fluorescencia (*RTEM: Retention Time-Emission Matrix*) se registraron en el rango espectral comprendido entre 380.0 nm y 510.0 nm, con la longitud de onda de excitación fijada en 280.0 nm, en el tiempo de elución comprendido entre 0.0 min y 16.0 min. El detector se operó con una ganancia de fotomultiplicador de 18 y una velocidad de barrido igual a 180 nm s^{-1} . De esta manera, las RTEM consistieron en 1012 \times 132 datos para la dimensión temporal y espectral, respectivamente.

Las mediciones espectroscópicas de fluorescencia se realizaron utilizando un espectrofotómetro de fluorescencia Perkin-Elmer LS-55 (Waltham, Estados Unidos) y una celda de cuarzo de 1 cm \times 1 cm de paso óptico. Los espectros se registraron cubriendo el rango espectral de entre 350.0 nm y 520.0 nm para emisión y 250.0 nm y 380.0 nm para excitación, utilizando una longitud de onda excitación y emisión correspondiente al máximo de intensidad, respectivamente. Para el control del instrumento y la adquisición de los datos se utilizó el programa informático FL WinLab (Perkin-Elmer Inc., Waltham, Estados Unidos).

Las mediciones de pH se realizaron utilizando un potenciómetro Orion 410A (Massachusetts, Estados Unidos) equipado con un electrodo de vidrio combinado Boeco BA17 (Hamburgo, Alemania).

Programas informáticos y algoritmos para el análisis de datos. El procesamiento de datos y el análisis quimiométrico con MCR-ALS se realizaron en MatLab 7.10 (MathWorks Inc., Natick, Estados Unidos). Los algoritmos necesarios para operar MCR-ALS se encuentran asequibles en internet accediendo a la

página web <http://www.mcrals.info/>. Todos los algoritmos necesarios para el pre-procesamiento de los datos fueron desarrollados y escritos en MatLab por integrantes de nuestro laboratorio.

RESULTADOS

Optimización de la concentración de Y^{3+} . Se evaluaron los cambios en la intensidad de fluorescencia de las 7 f-QUI promovidos por la presencia de Y^{3+} en una solución reguladora de concentración formal igual a 0.02 mol L^{-1} de NaAc/HAc pH 4.00. Los resultados demostraron la existencia de una fuerte dependencia de la intensidad de fluorescencia con la concentración de Y^{3+} . La intensidad de fluorescencia de todas las f-QUI mostró un significativo incremento frente al aumento de la concentración del metal, hasta alcanzar un valor máximo y constante cuando la concentración del metal superó el valor de $1.0 \times 10^{-4} \text{ mol L}^{-1}$. Sobre la base de estos resultados, se estableció un nivel de concentración de Y^{3+} igual a $1.0 \times 10^{-4} \text{ mol L}^{-1}$ para el desarrollo de todos los experimentos subsiguientes. En estas condiciones, la relación molar de Y^{3+} en la solución final respecto a los analitos a determinar fue de 1000:1 (Y^{3+} :f-QUI, considerando la mayor concentración de todas las f-QUI en una solución mezcla), asegurando el exceso de Y^{3+} en todos los experimentos.

En la Figura 3.3 se muestran los espectros de excitación y emisión normalizados para SRF en solución acuosa en presencia y ausencia de Y^{3+} . Se puede observar que la presencia del metal induce cambios espectrales debidos a la formación del complejo metálico. El espectro de emisión de fluorescencia del sistema SRF- Y^{3+} muestra un fuerte corrimiento de la longitud de onda de máxima intensidad a longitudes de ondas menores. Si bien el espectro de excitación de fluorescencia también muestra una perturbación ante la presencia del metal, el corrimiento hacia longitudes de ondas menores no es de la misma magnitud que el mostrado para la emisión. Este fenómeno se presentó para todas las f-QUI analizadas en las mismas condiciones. Estos resultados coinciden con los reportados en trabajos anteriores para NRF (Han *et al.*, 2005), ENF (Tong *et al.*, 2010b) y CPF (Tong *et al.*, 2010a), en similares condiciones, en los que se asegura que el cambio espectral observado se da por fluorescencia del ligando perturbada por Y^{3+} .

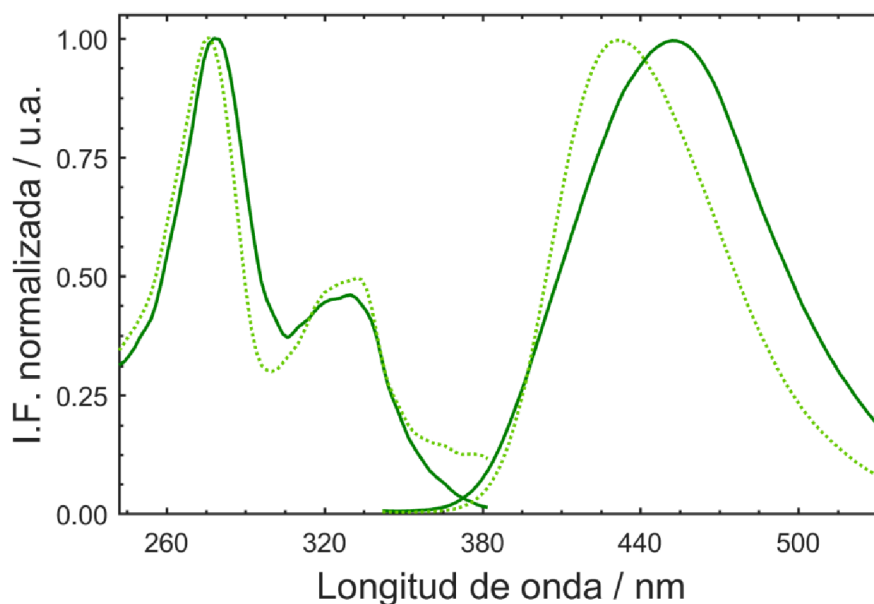


Figura 3.3. Espectros de excitación y emisión de fluorescencia normalizados de SRF en presencia (línea de puntos) y ausencia (línea continua) de Y^{3+} en una solución reguladora de NaAc/HAc pH 4.00

Para evaluar el efecto del incremento de la señal fluorescente de las f-QUI ante la presencia del metal en las condiciones cromatográficas experimentales, se realizaron determinaciones cromatográficas para cada f-QUI individualmente en presencia y ausencia de Y^{3+} . Los análisis cromatográficos en ausencia del metal se realizaron utilizando una fase móvil libre de Y^{3+} y sin agregado previo del metal a las muestras a analizar. En la Figura 3.4 puede apreciarse el efecto del Y^{3+} sobre la señal de fluorescencia obtenida para SRF por análisis cromatográfico. Una señal imperceptible de fluorescencia nativa (recuadro de Figura 3.4) de una solución de 60.0 ng mL^{-1} de SRF experimentó un incremento de 460 veces el valor de su señal ante la presencia de Y^{3+} .

Se observaron incrementos de la señal fluorescente en órdenes comprendidos entre 100 a 1740 veces mayores a la señal nativa. Este incremento en la señal analítica condujo a importantes reducciones en los límites de cuantificación y detección, haciendo al método propuesto comparable con aquellos que requieren sistemas de detección más complejos o metodologías que involucren procedimientos de pre-procesamiento de muestras que pudiesen resultar tediosos o largos.

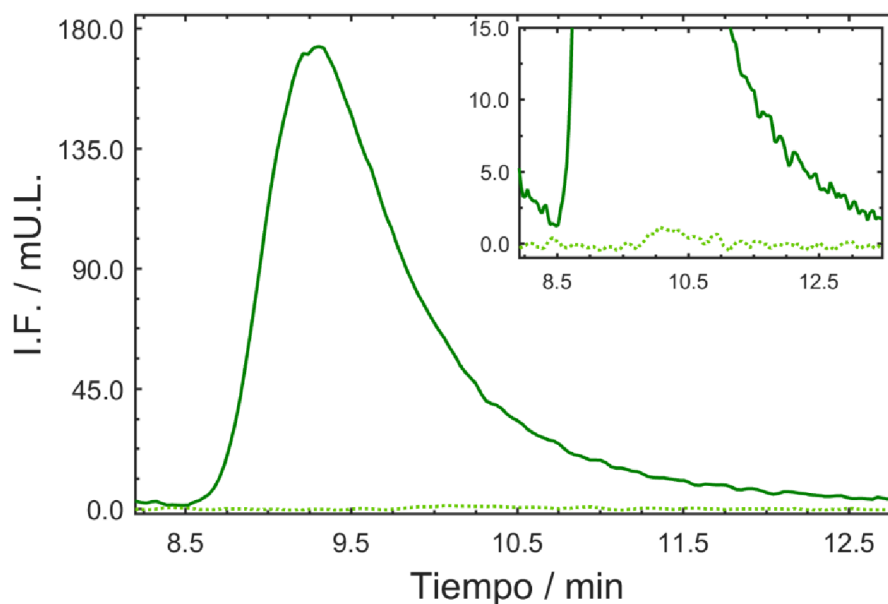


Figura 3.4. Señal cromatográfica obtenida para una solución conteniendo 60.0 ng mL^{-1} de SRF en ausencia (línea de puntos) y presencia (línea continua) de $1 \times 10^{-4} \text{ mol L}^{-1}$ de Y^{3+} . En el recuadro se amplifica la señal obtenida para SRF en ausencia de Y^{3+} . Detección: $\lambda_{\text{ex}} = 280.0 \text{ nm}$; $\lambda_{\text{em}} = 440.0 \text{ nm}$

En la Tabla 3.1 se detallan las áreas cromatográficas obtenidas para las 7 f-QUI evaluadas en presencia y ausencia de Y^{3+} .

Tabla 3.1. Áreas obtenidas de las señales cromatográficas correspondientes para las f-QUI analizadas en ausencia y presencia de Y^{3+}

<i>Analito</i> ^a	<i>Área sin Y^{3+}</i> ^b	<i>Área con Y^{3+}</i> ^c	<i>Aumento de intensidad</i> ^d
ENO	ND ^e	424.5	>424
NRF	5.5	2918.1	531
OFL	11.3	1895.4	168
CPF	17.7	1825.6	103
ENF	8.3	1340.5	162
SRF	27.6	12690.1	460
DIF	1.5	2614.5	1743

^a Concentraciones de CPF, ENF y NRF: 10.0 ng mL^{-1} ; OFL: 12.0 ng mL^{-1} ; ENO y DIF: 50.0 ng mL^{-1} ; SRF: 60.0 ng mL^{-1} ;

^b Área cromatográfica obtenida para cada f-QUI en ausencia de Y^{3+} . Detección: $\lambda_{\text{ex}} = 280.0 \text{ nm}$ y $\lambda_{\text{em}} = 440.0 \text{ nm}$ para CPF, NRF, ENF, SRF y DIF; $\lambda_{\text{em}} = 400.0 \text{ nm}$ para ENO; y $\lambda_{\text{em}} = 490.0 \text{ nm}$ para OFL;

^c Área cromatográfica obtenida para cada f-QUI en presencia de Y^{3+} . Detección: $\lambda_{\text{ex}} = 280.0 \text{ nm}$ y $\lambda_{\text{em}} = 440.0 \text{ nm}$ para CPF, NRF, ENF, SRF y DIF; $\lambda_{\text{em}} = 400.0 \text{ nm}$ para ENO; y $\lambda_{\text{em}} = 490.0 \text{ nm}$ para OFL;

^d Relación entre áreas obtenidas, calculada como $\frac{\text{Área con } \text{Y}^{3+}}{\text{Área sin } \text{Y}^{3+}}$;

^e ND: no detectable.

Selección de condiciones cromatográficas. Durante el desarrollo del método cromatográfico se observó baja reproducibilidad en el valor de las áreas cromatográficas en virtud de, principalmente, la formación del complejo metálico. La adición de Y^{3+} únicamente como componente de la fase móvil generó una variabilidad significativa para réplicas de inyección, tanto en los tiempos de retención como en las áreas de las señales cromatográficas. Este fenómeno sugirió una falta de reproducibilidad debida a la deficiencia en la formación del complejo metálico dentro de la columna, a causa de que la interacción con la fase estacionaria imposibilitaría la correcta interacción entre los analitos y el Y^{3+} . Ante este inconveniente, se optó por generar el complejo previamente a la resolución cromatográfica agregando Y^{3+} a las soluciones muestra durante la preparación, en un nivel de concentración igual al presente en la fase móvil. Adicionalmente, la presencia de Y^{3+} en la fase móvil favorecería a la estabilización en la columna del complejo pre-formado.

Posteriormente, se evaluó la manera de minimizar el porcentaje de solvente orgánico constituyente de la fase móvil, a fin de beneficiar la formación del complejo f-QUI- Y^{3+} y, así, desarrollar un método acorde con los lineamientos de la química analítica verde (Lakowicz, 2006; Tistaert *et al.*, 2012), pese a causar un detrimento en la forma del pico cromatográfico de los analitos. Además, se optimizó el pH de la fase móvil considerando la naturaleza de la fase estacionaria y su correcta interacción con las f-QUI, sin alterar la formación y estabilidad del complejo metálico formado. El valor de pH que generó los mejores resultados fue 4.00.

Modelado quimiométrico por MCR-ALS. En las condiciones experimentales analizadas existen dos grupos de analitos con idénticos espectros de emisión de fluorescencia: NRF/CPF/ENF y SRF/DIF. Esta característica se describe como deficiencia de rango espectral entre componentes de un sistema analítico (Tistaert *et al.*, 2012). Los espectros correspondientes a cada f-QUI analizada en este trabajo en presencia de Y^{3+} se muestran en la Figura 3.5.

En este trabajo, se seleccionó MCR-ALS, en su versión extendida (*ext-MCR-ALS*), como algoritmo quimiométrico para el modelado de datos. Para ello, se construyó una matriz aumentada en columnas, es decir, apilada en sentido espectral, llevándose a cabo la resolución en el sentido temporal.

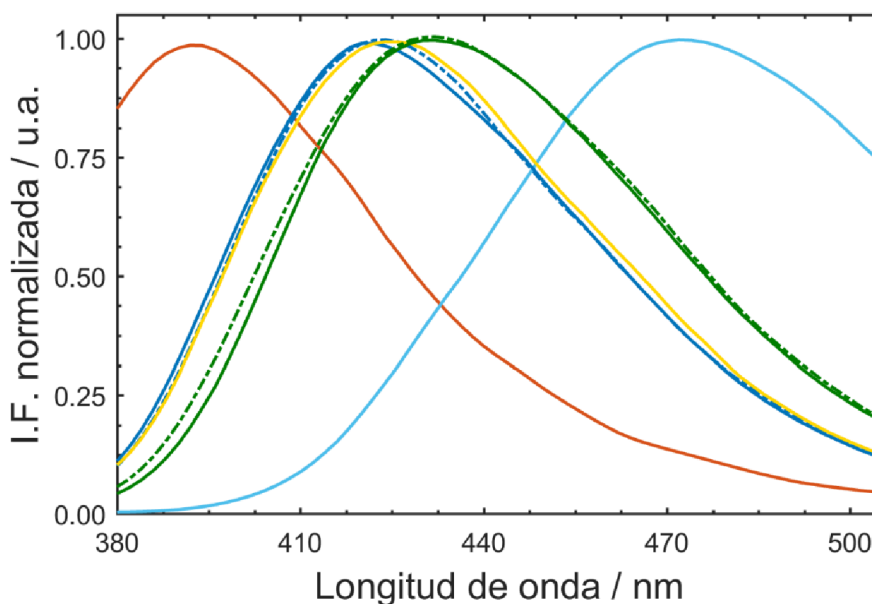


Figura 3.5. Espectros de emisión de fluorescencia normalizados de ENO (rojo), CPF (línea azul continua), NRF (línea de rayas y puntos azul), OFL (celeste), ENF (amarillo), SRF (línea continua verde) y DIF (línea de rayas y puntos verde) en presencia de Y^{3+} en solución reguladora de NaAc/HAc pH 4.0

En este caso, ext-MCR-ALS, al igual que otros algoritmos quimiométricos (PARAFAC, UPLS-RBL), no resulta el algoritmo ideal para la resolución quimiométrica, ya que no puede superar el inconveniente de deficiencia de rango espectral. Sin embargo, mediante una estrategia de análisis aplicada a la resolución quimiométrica se logró favorecer el modelado y la resolución de los datos. La estrategia aquí utilizada para el análisis de datos se resume en los siguientes pasos:

Previo al modelado quimiométrico, las matrices de datos se dividieron en 2 regiones en el sentido temporal. La primera región estuvo comprendida entre 2.2 min y 6.2 min incluyendo a ENO, NRF, OFL y CPF; la segunda región estuvo comprendida entre 6.2 min y 16 min incluyendo a ENF, SRF y DIF. No obstante, cada región contuvo a un par de analitos con idénticos espectros de emisión e incompleta resolución cromatográfica, manteniéndose aún la deficiencia de rango espectral.

Durante el ajuste por ALS del modelado quimiométrico, se consideró a cada par de analitos como componente único, no aplicando la restricción de unimodalidad para el factor correspondiente, obteniéndose como resultado un perfil temporal con dos picos cromatográficos.

Posterior al modelado quimiométrico y obtenidos los perfiles correspondientes a cada componente del sistema, se dividieron los perfiles temporales que presentaban 2 picos cromatográficos en 2 regiones, a fin de

obtener las contribuciones individuales de cada analito. Las contribuciones de NRF y CPF se obtuvieron dividiendo el perfil temporal en las regiones comprendidas entre 2.2 min y 3.2 min y 3.2 min y 6.2 min, respectivamente. El mismo procedimiento se siguió para obtener las contribuciones individuales de SRF y DIF, considerando las regiones comprendidas entre 8.2 min y 11.2 min y entre 11.2 min y 16 min, respectivamente.

La Figura 3.6 muestra el perfil cromatográfico obtenido para una muestra de validación conteniendo a las 7 f-QUI, indicándose las regiones seleccionadas previo al modelado quimiométrico.

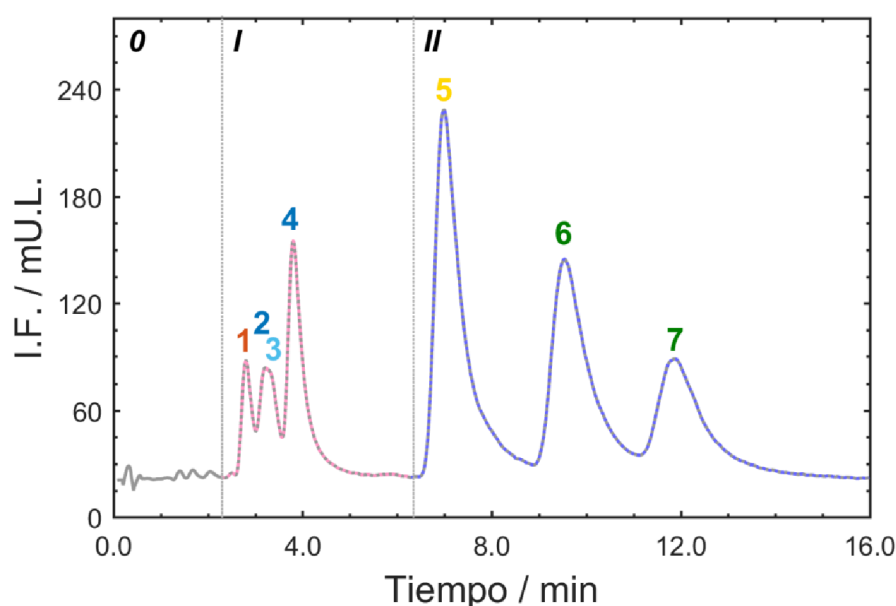


Figura 3.6. Cromatograma obtenido para una muestra conteniendo las 7 f-QUI en presencia de Y^{3+} , en el que se muestran las regiones seleccionadas para el modelado quimiométrico. *Región 0* (gris): frente de corrida, sin presencia de analitos; *Región I* (rosa): (1) ENO, (2) NRF, (3) OFL y (4) CPF; *Región II* (azul): (5) ENF, (6) SRF y (7) DIF

Para mejorar la calidad de los datos obtenidos y ayudar a la resolución quimiométrica, se procedió a la aplicación de técnicas de pre-procesamiento de datos, como ser corrección de línea de base y eliminación/disminución de ruido instrumental. Con el objetivo de reducir el ruido instrumental presente en los perfiles cromatográficos, se utilizó un procedimiento de suavizado basado en el método de Savitzky-Golay (Savitzky *et al.*, 1964). La corrección de la línea de base se llevó a cabo mediante la substracción de la línea de base cromatográfica correspondiente a cada cromatograma, considerándose ésta como una región de tiempo que no presenta señal analítica apreciable en todo el rango espectral evaluado, por ejemplo, la región final de un cromatograma.

Con base en lo explicado anteriormente, se construyeron dos matrices aumentadas en columnas (\mathbf{D}_{aumI} y \mathbf{D}_{aumII}) conteniendo las matrices de datos correspondientes a la validación (o a las muestras ambientales, según corresponda) y las matrices de datos correspondientes a la calibración de cada analito. Posteriormente, previo a iniciar la descomposición bilineal de los datos, se determinó el número de componentes espectralmente activos en cada matriz \mathbf{D}_{aum} aplicando descomposición en valores singulares (*SVD: Singular Values Decomposition*) y se obtuvieron las estimaciones espectrales iniciales mediante una metodología basada en SIMPLISMA (Windig *et al.*, 1991).

Con el objetivo de obtener resultados químicamente interpretables, se aplicaron distintas restricciones matemáticas durante el proceso iterativo ALS: **1- correspondencia entre especies**; **2- no negatividad en ambos modos**; y **3- unimodalidad en modo temporal**. Cabe mencionar que la restricción de unimodalidad se aplicó solo a aquellos componentes que no presentaron deficiencia de rango espectral. En la Figura 3.7 se muestran los perfiles espectrales y temporales obtenidos como resultado de la descomposición bilineal por MCR-ALS para una de las muestras ambientales fortificadas.

Para obtener las contribuciones individuales de los analitos y proceder al análisis cuantitativo, se dividieron los perfiles temporales correspondientes a CPF/NRF y SRF/DIF como se muestra en la Figura 3.8.

Análisis cuantitativo. De la descomposición bilineal se obtuvieron las matrices \mathbf{C} y \mathbf{S} conteniendo información relacionada a las abundancias y a los espectros de todos los componentes analizados, respectivamente. Considerando que la información contenida en \mathbf{C} para cada componente de las diferentes matrices de calibración, área o altura de pico cromatográfico, se encuentra directamente correlacionada con las concentraciones conocidas de los analitos, fue posible construir una curva de calibración tipo univariada (considerada aquí como pseudounivariada) a fin de poder predecir las concentraciones de los componentes presentes en muestras de composición desconocida.

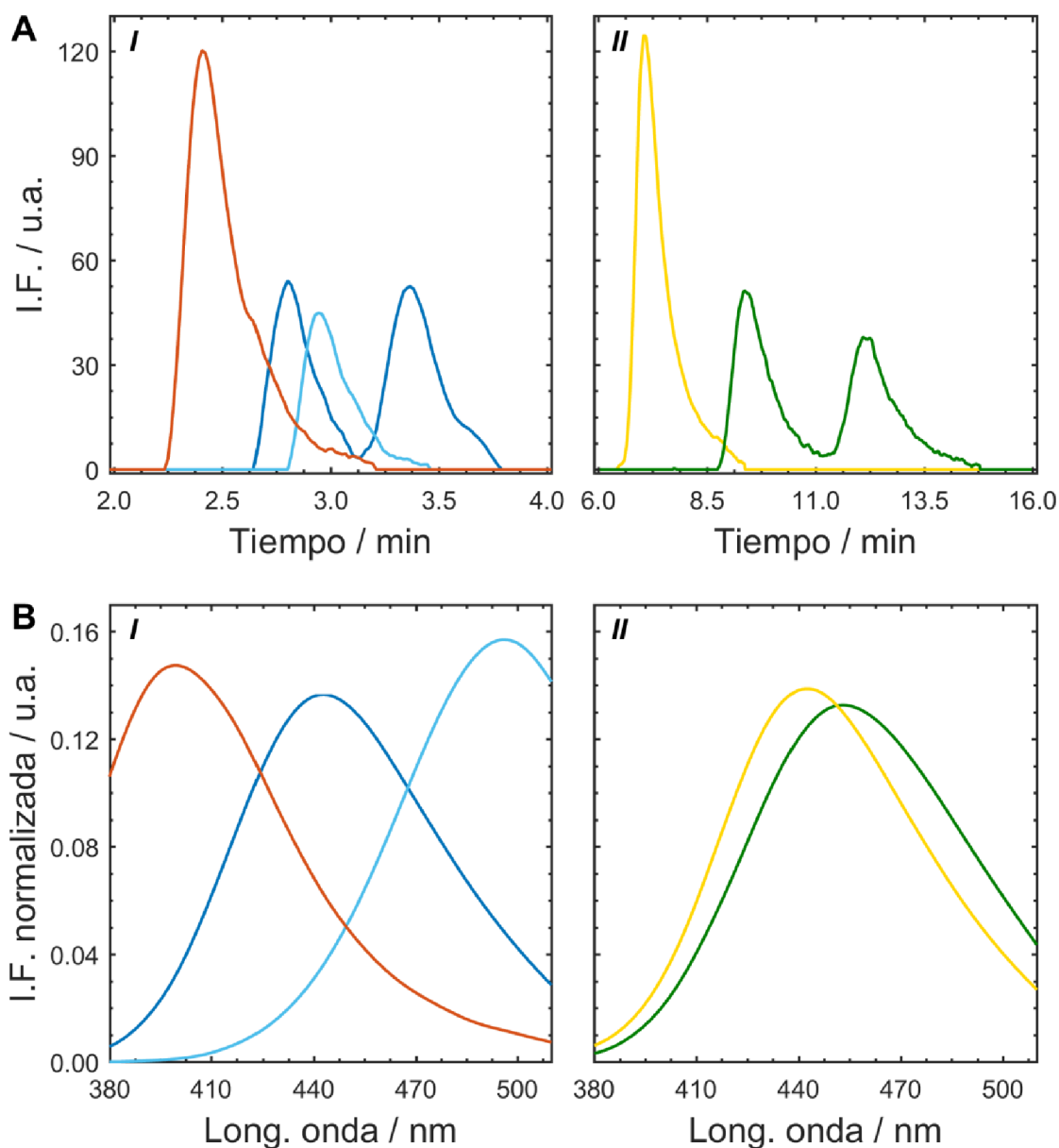


Figura 3.7. Perfiles temporales (A) y espectrales (B) de las 7 f-QUI analizadas obtenidos por modelado quimiométrico con MCR-ALS de cada matriz D_{aum} . La *Región I* contiene los analitos ENO (rojo), CPF/NRF (azul) y OFL (celeste) y la *Región II* los analitos ENF (amarillo) y SRF/DIF (verde)

En la Tabla 3.2 se muestran los resultados de predicción obtenidos para las 7 f-QUI en las distintas muestras de validación. En todos los casos, los valores de errores relativos de predicción (*REP %: Relative Error of Prediction*) fueron menores a 10 %. Para asegurar la exactitud del método desarrollado, se evaluaron las recuperaciones de manera estadística mediante un test de hipótesis, donde la aceptación de la hipótesis nula, H_0 : *la recuperación experimental media es igual a 100 %* (Olivieri, 2015), valida los valores de recuperación obtenidos experimentalmente. Los valores de t para datos experimentales (t_{exp}) obtenidos para las 7 f-QUI resultaron inferiores al valor crítico tabulado $t_{(0.05,10)} = 2.228$, para un nivel de confianza del 95 %, (nivel $\alpha = 0.05$) y $\nu = n-1$

grados de libertad ($\nu = 10$). Con base en estos resultados, se aceptó la H_0 , indicando que la recuperación experimental media no fue estadísticamente diferente a 100 %.

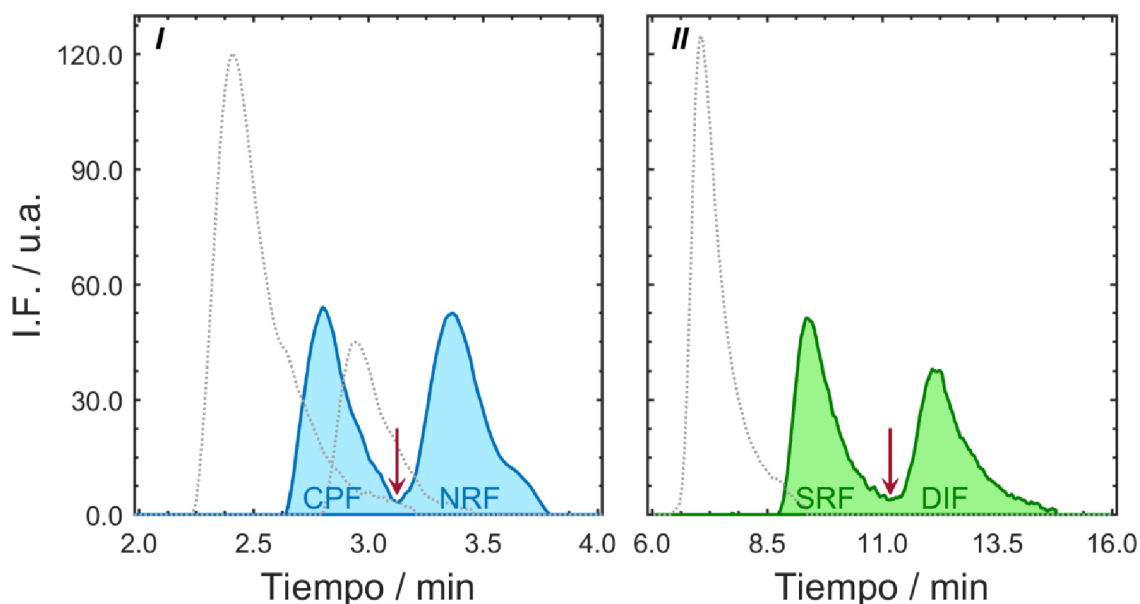


Figura 3.8. Perfiles temporales obtenidos por modelado con MCR-ALS para una de las muestras de validación analizada. El perfil de CPF y NRF, presente en la Región I, se indica en azul, mientras que el obtenido para SRF y DIF se indica en verde en la Región II. Las flechas rojas señalan el punto de división seleccionado para obtener las contribuciones individuales de los analitos. En líneas de punto gris se muestran los perfiles obtenidos para los demás analitos presentes en la muestra

Los resultados obtenidos del estudio de recuperación para la determinación simultánea de las 7 f-QUI en las tres muestras ambientales analizadas, para blancos y muestras fortificadas, se detallan en la Tabla 3.3. Es importante remarcar que con el método propuesto y desarrollado en esta tesis fue posible determinar la presencia de 3 f-QUI (*NRF*, *CPF* y *ENF*) en las muestras de agua del arroyo Las Prusianas. Si bien las concentraciones de los 3 analitos encontrados en las muestras de agua estaban a niveles muy bajos (0.4 , 1.0 y 3.6 ng mL^{-1} para *CPF*, *NRF* y *ENF*, respectivamente), resulta interesante remarcar, que no solo fue posible detectarlos, sino también cuantificarlos sin utilizar técnicas de pre-concentración o de pre-tratamiento de muestra. Para el análisis de muestras ambientales, también se evaluaron las recuperaciones, estadísticamente, a través del mismo test de hipótesis utilizado para las muestras de validación. Los resultados demostraron que las recuperaciones experimentales medias (\bar{R}_{exp}) y el valor real de 100 %, para un nivel de confianza del 95 % (nivel $\alpha = 0.05$) y $\nu = n - 1$ grados de libertad ($\nu = 8$), ($t_{(0.05,8)} = 2.306$), presentaron diferencias estadísticamente no significativas para todas las muestras ambientales fortificadas analizadas en todos los niveles de concentración.

Tabla 3.2. Valores de concentración predicha en muestras de validación para las 7 f-QUI determinadas simultáneamente ^a

Muestra	ENO		NRF		OFL		CPF		ENF		SRF		DIF	
	NOM.	PRED.	NOM.	PRED.	NOM.	PRED.	NOM.	PRED.	NOM.	PRED.	NOM.	PRED.	NOM.	PRED.
M01	125.0	119.5	15.0	13.6	30.0	33.2	25.0	25.7	15.0	13.7	24.0	24.6	80.0	76.3
M02	175.0	159.2	5.0	5.4	6.0	5.4	5.0	4.6	25.0	22.6	48.0	48.0	40.0	40.7
M03	75.0	74.2	5.0	5.0	30.0	29.0	25.0	22.8	0.8	0.8	48.0	44.4	80.0	82.4
M04	25.0	26.1	0.8	0.6	18.0	19.8	15.0	14.2	0.8	0.7	9.0	8.5	60.0	65.2
M05	175.0	158.0	5.0	5.5	18.0	18.0	15.0	13.6	25.0	25.6	48.0	50.0	60.0	66.0
M06	125.0	115.1	15.0	14.5	18.0	17.7	15.0	14.9	15.0	13.8	24.0	25.7	60.0	64.2
M07	75.0	72.7	25.0	26.1	6.0	6.6	5.0	5.4	5.0	4.8	72.0	77.7	60.0	57.9
M08	25.0	26.2	5.0	5.4	6.0	6.1	5.0	4.5	0.8	0.8	9.0	9.4	40.0	40.9
M09	175.0	162.4	5.0	5.3	6.0	5.5	5.0	4.5	25.0	22.8	48.0	47.6	40.0	40.8
M10	125.0	130.2	15.0	15.9	0.8	0.8	0.8	0.8	15.0	13.8	24.0	25.0	20.0	21.3
M11	75.0	72.7	25.0	26.0	0.8	0.8	0.8	0.8	5.0	5.4	72.0	79.0	40.0	40.2
REP % ^b	8.70		5.98		8.06		7.26		9.73		7.74		6.99	
\bar{R}_{exp} ^c	97.1		100.8		100.8		96.2		95.5		102.4		103.0	

^a Las concentraciones están dadas en ng mL⁻¹. Nom.: nominal; Pred.: predicha;

^b REP %: error relativo de predicción, expresado en %; $REP \% = 100 \times \frac{\sqrt{\frac{1}{7} \sum_{i=1}^7 (c_{nom} - c_{pred})^2}}{c}$, para l= 11;

^c \bar{R}_{exp} : recuperación media.

Tabla 3.3. Valores de concentración encontrados de f-QUI en muestras acuosas ambientales ^a

<i>Muestra</i> ^c	<i>ENO</i>		<i>NRF</i>		<i>OFL</i>		<i>CPF</i>		<i>ENF</i>		<i>SRF</i>		<i>DIF</i>	
	ADIC.	ENC.	ADIC.	ENC.	ADIC.	ENC.	ADIC.	ENC.	ADIC.	ENC.	ADIC.	ENC.	ADIC.	ENC.
<i>We</i>	0.0	ND ^b	0.0	ND	0.0	ND	0.0	ND	0.0	ND	0.0	ND	0.0	ND
<i>We_01</i>	40.0	36.3	5.0	4.6	12.0	12.6	5.0	4.7	5.0	4.6	60.0	58.3	50.0	45.1
<i>We_02</i>	160.0	145.5	25.0	23.1	36.0	38.0	10.0	10.4	10.0	9.2	12.0	11.8	25.0	22.9
<i>We_03</i>	100.0	92.8	10.0	10.3	36.0	35.3	25.0	24.6	25.0	22.8	36.0	36.7	50.0	54.9
<i>Wa</i>	0.0	ND	0.0	ND	0.0	ND	0.0	ND	0.0	ND	0.0	ND	0.0	ND
<i>Wa_01</i>	40.0	37.8	5.0	4.7	12.0	13.1	5.0	4.9	5.0	4.6	60.0	55.1	50.0	46.0
<i>Wa_02</i>	160.0	159.4	25.0	23.0	36.0	34.1	10.0	10.7	10.0	10.9	12.0	11.8	25.0	23.5
<i>Wa_03</i>	100.0	107.9	10.0	10.8	36.0	36.4	25.0	25.2	25.0	25.4	36.0	32.6	50.0	51.0
<i>Wsp</i>	0.0	ND	0.0	1.0	0.0	ND	0.0	0.4	0.0	3.6	0.0	ND	0.0	ND
<i>Wsp_01</i>	40.0	43.6	5.0	5.4 ^d	12.0	12.9	5.0	4.7 ^d	5.0	5.3 ^d	60.0	55.5	50.0	50.0
<i>Wsp_02</i>	160.0	159.4	25.0	23.9 ^d	36.0	35.0	10.0	10.8 ^d	10.0	10.8 ^d	12.0	12.4	25.0	26.7
<i>Wsp03</i>	100.0	103.3	10.0	10.1 ^d	36.0	38.9	25.0	22.9 ^d	25.0	24.7 ^d	36.0	34.4	50.0	50.1
\bar{R}_{exp} ^e	98.7		98.4		102.9		99.5		99.0		96.6		98.5	

^a Las concentraciones están dadas en ng mL⁻¹. Cada valor de concentración encontrado se expresa como el valor medio de tres réplicas de preparación. Adic.: adicionada; Enc.: encontrada;

^b ND, No Detectable;

^c We: muestra de agua de consumo obtenida de napa subterránea de la ciudad de Santa Fe; Wa: muestra de agua de efluente recolectada de diversos puntos de desagüe de la Facultad de Bioquímica y Ciencias Biológicas (Santa Fe, Argentina); Wsp: muestra de agua del arroyo Las Prusianas (Santa Fe, Argentina);

^d Diferencia entre el valor de concentración encontrado en la muestra fortificada y el valor de concentración encontrado en la muestra blanco;

^e \bar{R}_{exp} , Valor de recuperación media, expresado en %.

Por otro lado, con el objetivo de comparar cuantitativamente la calidad de los resultados espectrales obtenidos, se calculó el grado de solapamiento espectral (s_{12}) entre los espectros puros de cada analito (\mathbf{s}_1) y los perfiles espectrales obtenidos por modelado quimiométrico (\mathbf{s}_2), utilizando la siguiente la expresión:

$$s_{12} = \frac{\|\mathbf{s}_1^T \mathbf{s}_2\|}{\|\mathbf{s}_1\| \|\mathbf{s}_2\|} \quad 3.1$$

donde el valor de s_{12} puede tomar valores de entre 1 y 0, correspondiéndose a solapamiento completo o no solapamiento, respectivamente. Los espectros de los analitos puros se obtuvieron a partir de las señales cromatográficas puras de los analitos evaluados individualmente en las muestras de calibración al máximo nivel de concentración. Los valores s_{12} obtenidos en todos los casos fueron mayores a 0.9998 demostrándose la eficiencia del modelado quimiométrico y la capacidad de obtener información espectral de alta calidad.

Cifras de mérito analíticas. Uno de los objetivos fundamentales de esta tesis, además de evaluar la efectividad de los algoritmos quimiométricos para modelar datos complejos, es demostrar la importancia de considerar el orden de los datos multidimensionales al momento de obtener cifras de mérito analíticas representativas del sistema y, también, demostrar las mejoras obtenidas para dichas cifras al aumentar el orden dimensional de los datos.

La Tabla 3.4 resume los valores de LOD y LOQ obtenidos para las 7 f-QUI, estimados según 3 estrategias diferentes. La *Estrategia I* se basa en lo propuesto por Bauza *et al.* (Bauza *et al.*, 2012) que, para las estimaciones de las cifras de mérito, considera el orden dimensional de los datos y el tipo de algoritmo utilizado para el modelado quimiométrico. La *Estrategia II* sigue el principio de calibración univariada clásica (Saurina *et al.*, 2001), estimando las cifras de mérito según se detalla en las guías ICH (ICH, 2005). Por último, se propone como *Estrategia III* a la recomendada por las guías ICH para métodos analíticos que exhiben línea de base (ICH, 2005), por ejemplo, métodos cromatográficos y electroforéticos, que considera a la relación señal-a-ruido (*SNR: Signal-to-Noise Ratio*) entre la señal analítica de una solución blanco y la de un analito en la mínima concentración detectable para las estimaciones de LOD y LOQ, donde valores de SNR iguales a 3 y a 10 son, generalmente, considerados aceptables para la

estimación de LOD y LOQ, respectivamente.

Tabla 3.4. Valores de límite de detección y cuantificación estimados para las f-QUI analizadas siguiendo las tres estrategias de cálculo ^a

Analito	LOD			LOQ		
	Estrategia I ^b	Estrategia II ^c	Estrategia III ^d	Estrategia I	Estrategia II	Estrategia III
ENO	4.7	5.2	4.1	14.3	14.7	10.3
NRF	0.05	2.6	0.2	0.1	7.2	0.5
OFL	0.1	3.3	0.2	0.3	9.3	0.4
CPF	0.07	1.3	0.2	0.2	3.6	0.5
ENF	0.06	3.4	0.2	0.2	9.6	0.4
SRF	0.5	2.7	0.9	1.6	7.5	2.2
DIF	0.4	6.9	2.3	1.1	19.4	5.9

^a Valores dados en ng mL⁻¹;

^b *Estrategia I*, valores estimados según referencia (Bauza *et al.*, 2012);

^c *Estrategia II*, valores estimados según lo propuesto en por (Saurina *et al.*, 2001);

^d *Estrategia III*, valores estimados según recomendaciones de las guías ICH Q2R1, (ICH, 2005).

Puede observarse que las mejores estimaciones fueron obtenidas utilizando la *Estrategia I*. Estos resultados remarcan la importancia de considerar el orden dimensional de los datos en la estimación de las cifras de mérito representativas para el método desarrollado. Se demuestra, también, que al aumentar el orden de los datos se logran notables mejoras en las estimaciones de LOD y LOQ. Este fenómeno es en virtud del incremento de la SEN frente al incremento del orden de los datos, lográndose una disminución en los valores de LOD y LOQ (Olivieri, 2014). Por otro lado, los valores de LOD y LOQ estimados según la *estrategia I* son coincidentes en el orden de magnitud con los estimados según la *Estrategia III*, demostrándose que de esta manera es posible obtener estimaciones realistas de las cifras de mérito que caracterizan al método analítico. La diferencia entre los valores obtenidos mediante estas dos estrategias radica, principalmente, en el procesamiento de los datos previo al modelado por MCR-ALS, que incluye, entre otras, una etapa de suavizado mediante el método de Savitzky-Golay.

Los LOD estimados siguiendo la *Estrategia I* resultaron menores a 0.5 ng mL⁻¹ para todas las f-QUI evaluadas, excepto para ENO. Estos valores hacen al método propuesto comparable con aquellos que involucran equipamiento sofisticado y de alto costo e incluso con métodos que requieren técnicas de pre-concentración para lograr

valores similares a los reportados en este trabajo. Los LOD aquí obtenidos son comparables a los reportados por Herrera-Herrera *et al.* (Herrera-Herrera *et al.*, 2013) para un método desarrollado para la determinación de f-QUI en muestras de agua utilizando cromatografía líquida de ultra-alto rendimiento acoplado a DAD (*UHPLC-DAD: Ultra-High Performance Liquid Chromatography-Diode Array Detector*), e involucrando, además, pasos previos de pre-concentración mediante micro-extracción líquido-líquido dispersiva (*DLLME: Dispersive Liquid-Liquid Micro-Extraction*). Una metodología recientemente publicada para la determinación de f-QUI en diferentes matrices acuosas reporta valores de LOD en el mismo orden de magnitud que los presentados en este trabajo, utilizando un método basado en *salting-out* asistido por extracción líquido-líquido (*SALLE: Salting-out Assisted Liquid-Liquid Extraction*) como estrategia de pre-concentración y resolución cromatográfica por UHPLC acoplado a detección fluorescente (Arroyo-Manzanares *et al.*, 2015). Asimismo, los valores de LOD resultaron ser menores a los reportados por Cañada-Cañada *et al.* (Cañada-Cañada *et al.*, 2009) para la determinación de 8 f-QUI en muestras acuosas.

Todos los resultados obtenidos y mostrados anteriormente remarcan la idoneidad del método propuesto para la detección y cuantificación simultánea de diferentes f-QUI en matrices acuosas de origen natural a nivel de trazas, sin requerir de pasos de pre-concentración y/o tratamiento de muestras.

CONCLUSIÓN

Se desarrolló un método que combina cromatografía líquida de alto rendimiento con detección fluorescente, basado en la formación de complejos de f-QUI- Y^{3+} y modelado quimiométrico de datos de 2-vías aumentado. Este método permitió la determinación de 7 f-QUI en diferentes muestras de agua de fuentes naturales de una manera rápida, sencilla y eficiente. El incremento de la intensidad de fluorescencia de los analitos evaluados, promovido por la presencia de Y^{3+} , demostró ser una simple pero eficiente manera de lograr excelentes valores de sensibilidad, permitiendo la detección y cuantificación de los analitos en muestras complejas a nivel de trazas.

Adicionalmente, el uso de algoritmos quimiométricos para modelar datos de segundo orden, como MCR-ALS, permitió resolver problemas de análisis de mezclas, incluso en presencia de componentes con señales altamente solapadas. En este caso, el uso del algoritmo MCR-ALS, acoplado a una estrategia de pre- y post-análisis, per-

mitió superar la complejidad del sistema en virtud de la deficiencia de rango espectral, lográndose obtener las contribuciones individuales de todos los analitos.

La complejidad mostrada por el sistema analítico presente resulta un desafío para aquellos algoritmos, tales como PARAFAC, que no permiten desvíos de trilinealidad en los datos o presentan deficiencia de rango en alguno de sus modos. Las estrategias de procesamiento de datos propuestas en este trabajo demandan obtener perfiles espectrales y temporales químicamente interpretables como resultado de la descomposición bilineal, de manera de obtener las contribuciones individuales de todos los analitos evaluados de manera precisa. Por ello, los algoritmos basados en variables latentes, como PLS, no demuestran ser eficientes para superar la complejidad de este tipo de datos, al menos, de una manera sencilla.

Con base en lo expuesto, se puede concluir que el método propuesto resulta altamente adecuado para la cuantificación de f-QUI en matrices acuosas de muestras ambientales de una forma simple, rápida y eficiente, sin requerir instrumentación sofisticada y/o de alto costo, evitando largas y tediosas etapas de pre-concentración y tratamiento de muestras. Además, de acuerdo con la bibliografía consultada (El-Didamony, 2011; Zhu *et al.*, 2008) y considerando que las f-QUI actúan con Y^{3+} vía quelación a través de grupos específicos, el método propuesto resultaría idóneo para la determinación de otras f-QUI que no se analizaron en este trabajo.

3. 2. MÉTODO ELECTROFORÉTICO PARA LA DETERMINACIÓN DE QUINOLONAS EN AGUA DE CONSUMO CON MODELADO QUIMIOMÉTRICO DE DATOS MULTIDIMENSIONALES

Ultrafast quantitation of six quinolones in water samples by second-order capillary electrophoresis data modeling with multivariate curve resolution–alternating least squares – Alcaráz, M.R., et al., Anal Bioanal Chem 2014, 406, 2571–2580

CONSIDERACIONES GENERALES

El uso no controlado y el descarte desmesurado de fármacos al ambiente han llevado a una emergencia global por el surgimiento de una inminente contaminación ambiental, principalmente, por antibióticos. Un gran número de fármacos son descargados al ambiente como producto inalterado o como metabolitos por diferentes medios que terminan luego acumulándose en agua o en sedimentos sólidos (Rehman *et al.*, 2015; Rusu *et al.*, 2015). En la industria ictícola, las drogas utilizadas para el tratamiento en peces son adicionados directamente en el agua de criadero en grandes cantidades, si bien la US FDA prohibió el uso o el exceso de determinadas drogas para el tratamiento de peces destinados a consumo, como por ejemplo, cloranfenicol, nitrofuranos y f-QUI (FDA *et al.*, 2011). Por su parte, el tratamiento de plantas y vegetales con antibióticos se realiza por rociado directo con la droga sobre el espécimen. Estas prácticas impactan directamente sobre el ambiente y, en consecuencia, sobre la salud humana y medioambiental. Sin embargo, el caso más alarmante y de mayor incidencia sobre la contaminación ambiental con antibióticos es el impacto de la industria farmacéutica que, en algunos países desarrollados, es la principal fuente de contaminación ambiental debido a su alto caudal de desechos (Rehman *et al.*, 2015).

Dentro de los antibióticos más utilizados en las últimas décadas, debido a su alta efectividad, se encuentran las QUI, un grupo de antibióticos sintéticos de amplio espectro con un alto valor terapéutico. Se ha demostrado que estos compuestos pertenecen a un grupo de antibióticos de gran persistencia en sedimentos sólidos y que

muestran baja capacidad de biodegradación (Schaefer, 2002; Sturini *et al.*, 2012). Los residuos de estos antibióticos en el ambiente, en alimentos de origen animal o en agua de consumo pueden perjudicar notablemente la salud del consumidor y, al mismo tiempo, favorecer a la no deseada resistencia bacteriana a estas sustancias (Rusu *et al.*, 2015). El desarrollo de dicha resistencia y, principalmente, la resistencia a QUI han encendido las alarmas de las autoridades internacionales, como la Organización Mundial de la Salud (OMS o WHO: *World Health Organization*) o la US FDA (Crawford, 2005; Rusu *et al.*, 2015), quienes demostraron la presencia de este fenómeno y aseguran que su incidencia se encuentra gradualmente en aumento. La frecuente aparición de QUI en sedimentos sólidos, alimentos y en agua de consumo (Webb *et al.*, 2003; Ye *et al.*, 2007) ha intensificado la búsqueda y el desarrollo de nuevas actividades de investigación con foco en la identificación de posibles riesgos ambientales, como así también en el desarrollo de métodos analíticos sensibles para su determinación.

Dentro del gran número de métodos analíticos que se pueden hallar en la bibliografía para la determinación de QUI en matrices acuosas, se encuentran aquellos basados en técnicas separativas (He *et al.*, 2015; Rusu *et al.*, 2014), espectroscópicas (Tong *et al.*, 2010b) y electroquímicas (Pellegrini *et al.*, 2004), entre otras. Como método separativo, la electroforesis capilar (CE: *Capillary Electrophoresis*) ha demostrado ser una poderosa herramienta para la determinación de diversos compuestos en matrices complejas debido a su simplicidad, al requerimiento de pequeñas cantidades de muestras, al uso de solventes no contaminantes y a su rapidez y versatilidad (Le *et al.*, 2016; Rusu *et al.*, 2014).

En CE, la separación se basa, principalmente, en las propiedades acido-base de los analitos y en su relación carga/masa, por lo que las similitudes estructurales entre compuestos resulta siempre un desafío analítico, fundamentalmente, cuando se desea lograr la separación completa de las señales analíticas en un sistema multi-componente. Sin embargo, en los casos donde no es posible lograr la completa separación, la generación de datos de segundo orden con posterior modelado quimiométrico resulta una excelente alternativa para la resolución y cuantificación de los analitos, incluso cuando sus señales analíticas se encuentran fuertemente solapadas (García *et al.*, 2009; Vera-Candiotti *et al.*, 2008).

Una característica química significativa de las QUI es que presentan carácter ácido-base, que incide en su mecanismo de acción y en sus propiedades farmacológicas. Estas drogas son compuestos ionizables que, dependiendo del pH del entorno, pueden encontrarse en forma cargada o neutra. Con base en su comportamiento ácido-base y sus características estructurales, las QUI pueden ser clasificadas en dos

tipos generales: **1-** las piperizinil-QUI que contienen un grupo piperizinil en su núcleo y se puede encontrar como 3 especies iónicas diferentes, catiónica (carga neta +; qn^+), zwitteriónica (carga neta nula, qn^0) y aniónica (carga neta -; qn^-); y **2-** las QUI ácidas que solo pueden encontrarse en estado neutro (qn^0) o aniónico (qn^-) (Babić *et al.*, 2007). En consecuencia, las primeras muestran 2 constantes de disociación principales, $pK_{a,1}$ y $pK_{a,2}$, mientras que las QUI ácidas solo muestran 1 valor de pK_a . En la Figura 3.9 se esquematizan los equilibrios ácido-base principales correspondientes a los dos grupos de QUI mencionados.

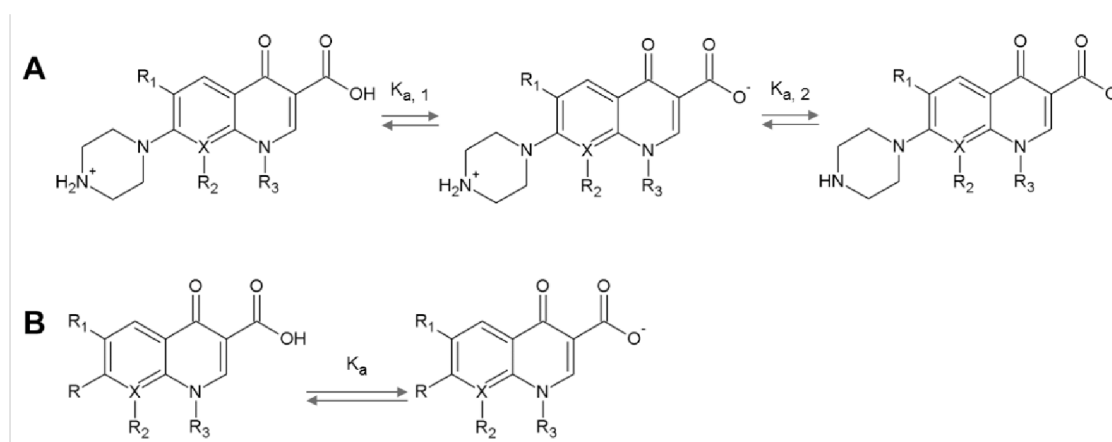


Figura 3.9. Esquema propuesto para los equilibrios ácido-base de las piperizinil-QUI (A) y de las QUI ácidas (B)

Aprovechando las propiedades estructurales y ácido-base de las QUI y las bondades de las técnicas electroforéticas, se propone un método simple para la determinación simultánea de 6 QUI en agua de consumo, basado en electroforesis capilar zonal con detección DAD-UV, acoplado a herramientas de calibración multivariada. Además, con el objetivo de disminuir los límites de detección y cuantificación, se propone una metodología de extracción en fase sólida como etapa de pre-concentración. Para la resolución de las señales analíticas y obtención de las contribuciones individuales de cada componente, se propone una estrategia de análisis y procesamiento de datos que utiliza herramientas de diseño experimental y algoritmos quimiométricos de calibración multivariada.

MATERIALES Y MÉTODOS

Reactivos. Todos los estándares utilizados fueron de calidad grado analítico. Las drogas sólidas de flumequine (*FLU*), enoxacina (*ENO*), ofloxacina (*OFL*), sarafloxacin (*SRF*), cinoxacina (*CIN*), difloxacina (*DIF*) y fenitoína (*PHT*) fueron adquiridas de Sigma-Aldrich (Steinheim, Alemania). Enrofloxacin (*ENF*), ciprofloxacina (*CPF*) y danofloxacina (*DNF*) fueron provistas por Fluka (St. Gallen, Suiza) y marbofloxacina (*MRF*) fue obtenida de Molekula (Gillingham, Reino Unido). El metanol (*MeOH*) calidad cromatografía líquida, fue obtenido de J.T.Baker (Deventer, Países Bajos). El ácido clorhídrico (*HCl*), hidróxido de sodio (*NaOH*), fosfato de sodio monobásico (NaH_2PO_4), citrato de sodio (*NaCit*) y tetraborato de sodio ($Na_2B_4O_7$), todos de calidad grado analítico, fueron obtenidos de Cicarelli (San Lorenzo, Argentina). El agua de laboratorio ultrapura fue obtenida por un sistema de purificación de agua Milli-Q Milli-Pore (Billerica, Estados Unidos).

Soluciones de trabajo. Se preparó una solución madre de cada QUI por disolución de una cantidad apropiada de cada analito en metanol alcalinizado (pH~9.0) a fin de obtener una concentración final de cada QUI de $200.0 \mu\text{g mL}^{-1}$. Para su conservación, todas las soluciones se almacenaron a 4°C protegidas de la luz por un periodo no mayor a 2 meses.

Todas las soluciones reguladoras necesarias para este trabajo se prepararon por disolución de la cantidad apropiada de la sal correspondiente en agua ultrapura, ajustando el pH con $\text{HCl } 1.0 \text{ mol L}^{-1}$ o $\text{NaOH } 1.0 \text{ mol L}^{-1}$, según corresponda.

Muestras de calibración y validación. Para la calibración se preparó un set de 5 soluciones estándares en diferentes niveles de concentración para cada QUI individualmente. Las soluciones se prepararon transfiriendo una alícuota apropiada de la solución madre correspondiente a matraces aforados de 25.0 mL , completándose el volumen con agua ultrapura. Los rangos de concentración en los sets de calibración para cada QUI estuvieron comprendidos entre 0.0 ng mL^{-1} y 100.0 ng mL^{-1} para *ENO*, *CPF*, *OFL*, *ENF* y *FLU*, y entre 0.0 ng mL^{-1} y 330.0 ng mL^{-1} para *CIN*.

Para la validación, se preparó un set de 5 soluciones mezcla en diferentes niveles de concentración, considerando valores distintos a los utilizados para la calibra-

ción, si bien comprendidos en el mismo rango de concentraciones. Las soluciones se prepararon siguiendo un modelo aleatorio transfiriendo alícuotas apropiadas de cada solución madre de QUI a matraces aforados de 25.0 mL y completándose el volumen con agua ultrapura.

Todas las soluciones se prepararon diariamente por triplicado y se sometieron al mismo procedimiento de pre-concentración.

Muestras de agua de consumo. Las muestras de agua se obtuvieron directamente de la red de agua corriente apta para consumo de la ciudad de Santa Fe (Santa Fe, Argentina), recolectadas en distintos puntos de la ciudad. Las muestras se colectaron en botellas de 1 L color ámbar y se procesaron de inmediato a su llegada al laboratorio; en los casos en los que no fue posible analizarlas de manera inmediata, se almacenaron a 4 °C protegidas de la luz, por un lapso no mayor a 1 semana.

Se preparó 1 set compuesto por 8 soluciones muestra que incluyeron a las soluciones muestra no fortificadas, o blanco de muestra, y las soluciones muestra fortificadas en 3 niveles de concentración para cada QUI. Las soluciones se prepararon siguiendo el mismo procedimiento descrito para las soluciones de validación, si bien el volumen se completó con agua de las muestras recolectadas. A 3 de las muestras fortificadas se le adicionó, además, 50.0 ng mL⁻¹ de MRF, DNF y PHT, que fueron considerados como componentes esperados no calibrados. La solución muestra restante se preparó en un nivel de concentración de QUI igual a la concentración media del rango de concentraciones utilizadas para la calibración por dilución de las alícuotas correspondientes de cada QUI en matraces aforados de 25.0 mL y completándose el volumen con agua de consumo. Esta muestra contenía, además, 50.0 ng mL⁻¹ de MRF, DNF, PHT, SRF y DIF, que fueron considerados como componentes esperado no calibrados.

Todas las soluciones se prepararon por triplicado y se sometieron al mismo procedimiento de pre-concentración. Finalmente, las soluciones se filtraron a través de membranas de nylon con tamaño de poro de 0.45 µm (Sartorius, Goettingen, Alemania) y se desgasificaron por ultrasonido (Cole Palmer 8891, Abbott, Temecula, CA, Estados Unidos).

Extracción en fase sólida. La extracción en fase sólida (*SPE: Solid Phase Extraction*) se realizó en cartuchos de pre-concentración Oasis HLB comercia-

les de $1 \text{ cm}^3 / 30 \text{ mg}$ de capacidad (Waters Corporation, Milford, Estados Unidos). Para su acondicionamiento, se transfirió 1.0 mL de MeOH seguido de 2.0 mL de agua ultrapura, con una velocidad de flujo de, aproximadamente, 3.0 mL min^{-1} . Para la pre-concentración, se transfirieron 25.0 mL de la muestra a analizar respetándose la misma velocidad de flujo utilizada para el acondicionamiento. Posteriormente, se transfirieron 2.0 mL de agua ultrapura con el objetivo de remover trazas de muestra que hayan podido quedar retenidas en la matriz del cartucho. Los analitos adsorbidos se eluyeron, finalmente, con 250.0 μL de MeOH y se colectaron directamente en viales de inyección. De esta manera, se obtuvo un factor de pre-concentración igual a 100, que en otras palabras significa lograr concentrar 100 veces el nivel de concentración inicial de los analitos. Finalmente, una alícuota del extracto metanólico se inyectó directamente en el sistema electroforético.

Instrumentación y procedimiento analítico. Todos los experimentos electroforéticos se llevaron a cabo usando un instrumento Agilent EC (Agilent Technologies, Waldbronn, Alemania) equipado con inyector automático, compartimento para capilar termostatzado, DAD-UV-Vis y el programa informático ChemStation (Agilent Technologies, Waldbronn, Alemania) requerido para el control del instrumento, la adquisición y el análisis de datos.

La separación electroforética se realizó bajo polaridad positiva, utilizando un capilar de sílice fundida no recubierto de 50 cm de longitud y 75 μm de diámetro interno (MichroSolv Technology Corp., Eatontown, Estados Unidos), operando a un voltaje de 25 kV, bajo una corriente inducida de aproximadamente 65 μA y a una temperatura de 25 $^{\circ}\text{C}$. Para el análisis electroforético, en primer lugar, se acondicionó y se activó el capilar mediante lavados sucesivos con NaOH 1.0 mol L^{-1} , agua ultrapura y solución de electrolito de corrida (*BGE: BackGround Electrolyte*), en periodos de 10 min cada una. La misma secuencia utilizada para el acondicionamiento y activación del capilar, pero en periodos de 3 min, se utilizó para remover sustancias que pudiesen quedar adsorbidas en la pared del capilar entre el análisis de muestras. La inyección de la muestra se realizó de manera hidrodinámica en polo positivo, aplicándose una presión igual a 50 mbar, durante 10 s. Las matrices de datos correspondientes a tiempo de migración-espectro UV (*MTUM: Migration Time-UV Matrix*) se registraron en el rango espectral comprendido entre 190.0 nm y 400.0 nm, en el tiempo de elución de entre 0.0 min y 3.0 min. Sin embargo, para favorecer la resolución y el modelado quimiomé-

trico, las matrices de datos se restringieron en la dimensión temporal, evitándose la señal correspondiente al flujo electroosmótico (*EOF: ElectroOsmotic Flow*) y en dimensión espectral en el rango comprendido entre 260.0 nm y 365.0 nm. De esta manera, las MTUM consistieron en 121 × 71 datos para la dimensión temporal y espectral, respectivamente.

Las mediciones de pH se realizaron utilizando un potenciómetro Orion 410A (Massachusetts, Estados Unidos) equipado con un electrodo de vidrio combinado Boeco BA17 (Hamburgo, Alemania).

Optimización de condiciones experimentales. A fin de obtener las condiciones experimentales adecuadas que permitan la correcta cuantificación de los analitos, se utilizó un diseño experimental de optimización evaluando las respuestas generadas por la combinación de dos variables de proceso (voltaje aplicado y pH de la solución BGE), analizando, a su vez, diferentes mezclas de electrolitos constituyentes de la solución BGE. Las respuestas evaluadas en el diseño experimental fueron: **1- tiempo de análisis;** y **2- resolución entre las señales analíticas de CPF y ENF.** Una vez seleccionado el modelo que mejor se ajuste a las respuestas obtenidas, se seleccionaron los valores óptimos de los factores evaluados mediante la función de deseabilidad de Derringer (Derringer *et al.*, 1980).

Programas informáticos y algoritmos para el análisis de datos. El procesamiento de datos y el análisis quimiométrico con MCR-ALS se realizaron en MatLab 7.10 (MathWorks Inc., Natick, Estados Unidos). Los algoritmos necesarios para operar MCR-ALS se encuentran asequibles en internet accediendo a la página web <http://www.mcrals.info/>. Todos los algoritmos necesarios para el preprocesamiento de los datos fueron desarrollados y escritos en MatLab por integrantes de nuestro laboratorio. El diseño experimental y la función de deseabilidad se realizaron utilizando el programa Design-Expert 8.0.5 (Stat-Ease Inc., Minneapolis, Estados Unidos).

RESULTADOS

Optimización de las condiciones experimentales. Al igual que lo ocurrido para el método cromatográfico expuesto en la sección 3.1, en las condiciones experimentales analizadas, los analitos CPF y ENF presentan idénticos espectros de absorbancia. En consecuencia, el sistema presentó deficiencia de rango espectral entre factores, por lo que la descomposición bilineal por MCR-ALS basada, principalmente, en las diferencias espectrales de los componentes, no resultaría suficientemente eficiente para lograr la satisfactoria resolución de los analitos. Por consiguiente, se propone una estrategia operativa que consiste en una etapa de pre-análisis basada en la búsqueda de valores óptimos de los factores experimentales que permitan la completa resolución de los analitos con idéntico espectros de absorbancia y una etapa de post-procesamiento de datos que permita obtener las contribuciones individuales de cada componente.

En este trabajo se aprovecharon las ventajas del diseño experimental de optimización a fin de posibilitar la resolución electroforética completa entre las señales analíticas de los componentes que presentan deficiencia de rango espectral, pero, al mismo tiempo, reducir el tiempo de análisis electroforético. El diseño experimental utilizado para la optimización de las respuestas consistió en una combinación entre un diseño Lattice *simplex* y un diseño factorial en dos niveles. Las respuestas evaluadas para la optimización fueron: **1- tiempo de análisis;** y **2- resolución electroforética entre las señales analíticas de CPF y ENF.**

Por un lado, se evaluó el efecto de la cantidad y la naturaleza de los electrolitos constituyentes de la solución BGE mediante un diseño Lattice *simplex*, teniendo en cuenta los siguientes criterios: **1- naturaleza de electrolitos:** PO_4^{3-} , $B_4O_7^{2-}$, y Cl^- ; **2- concentración formal de la solución BGE:** comprendida entre 0.0 mol L^{-1} y 0.015 mol L^{-1} ; **3- la solución BGE debe contener al menos dos electrolitos de distinta naturaleza;** y **4- la concentración máxima de cada electrolito en la solución BGE mezcla no debe ser mayor a 0.010 mol L^{-1} .** De esta manera, el diseño experimental consistió de 7 combinaciones, que incluyeron un punto central, 3 puntos internos y 3 puntos en el centro de los bordes. Adicionalmente, se evaluó el efecto de las variables de proceso para cada una de las 7 combinaciones del diseño descrito mediante un diseño factorial en 3 niveles. Las variables de proceso evaluadas fueron: **1- pH de la solución BGE:** entre 7.0 y 9.0; y **2- voltaje aplicado:** evaluado entre 15 kV y 25 kV. El diseño

experimental combinado completo consistió, entonces, en un total de 63 experimentos separados en 2 bloques, considerando a cada bloque como un día de trabajo diferente. La Figura 3.10 esquematiza gráficamente el diseño experimental combinado utilizado para la optimización de las condiciones experimentales de trabajo.

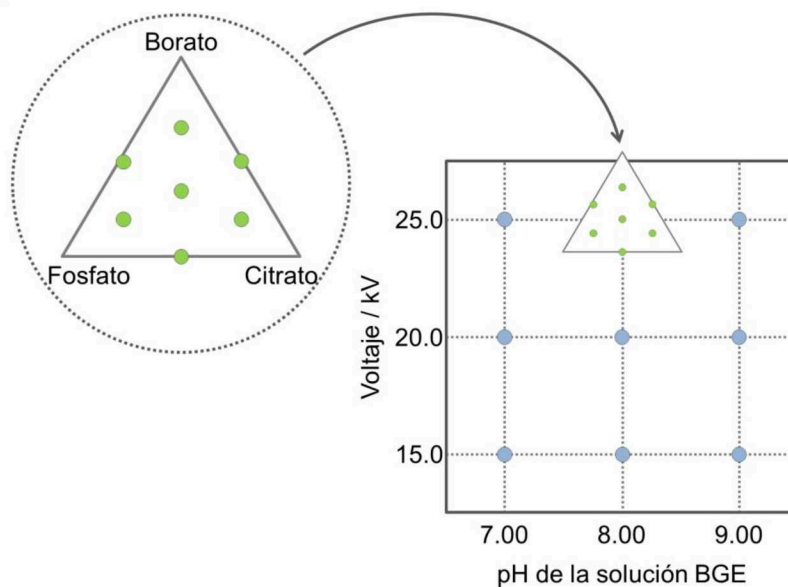


Figura 3.10. Representación esquemática del diseño experimental combinado Lattice simplex y factorial en 3 niveles. Cada círculo celeste representa los puntos del diseño factorial, mientras que los círculos verdes indican los puntos correspondientes al diseño Lattice simplex

Los factores y rangos evaluados en el diseño experimental se seleccionaron de acuerdo a los resultados obtenidos en experimentos preliminares, por ejemplo, variaciones en la temperatura, en la longitud de capilar y en el tiempo de inyección mostraron ser no significativas en el resultado experimental.

Obtenidas las 2 respuestas de interés correspondientes a los 63 experimentos, se evaluaron los diferentes modelos matemáticos disponibles que mejor describiesen el comportamiento de las respuestas, seleccionándose aquel que arroje mejor valor de ajuste. Los coeficientes obtenidos para los modelos matemáticos utilizados se calcularon y se evaluaron mediante regresión múltiple por eliminación progresiva y test de hipótesis de análisis de varianza (*ANOVA: ANalysis Of VAriance*) (Myers *et al.*, 2009), respectivamente. El modelo matemático de elección se consideró satisfactorio al aceptar la hipótesis nula del test de ANOVA para un nivel de confianza del 95 % ($\alpha = 0.05$). Adicionalmente, para la selección del modelo matemático de ajuste se consideró el principio de parsimonia, o simplicidad. De esta manera, las respuestas consideradas

para la optimización de los factores fueron ajustadas a un modelo combinado cuadrático \times lineal ($r^2 = 0.9283$) y combinado lineal \times cuadrático ($r^2 = 0.8746$) para tiempo de análisis y resolución electroforética, respectivamente.

Los valores óptimos de los factores experimentales se obtuvieron mediante el uso de la función de deseabilidad de Derringer, con base en los siguientes criterios para cada respuesta evaluada: **1- tiempo de análisis mínimo**; y **2- resolución electroforética entre las señales analíticas correspondientes a CPF y ENF igual a 2**. Finalmente, se seleccionaron aquellas condiciones experimentales que, cumpliendo ambos criterios, arrojasen un valor de deseabilidad cercano al valor máximo, es decir, $D = 1.00$. De esta manera, se seleccionaron las condiciones experimentales óptimas y se procedió según se detalla a continuación: *bajo el modo de electroforesis capilar zonal, se utilizó una solución BGE compuesta por una mezcla de 0.010 mol L^{-1} de NaB_4O_7 y 0.005 mol L^{-1} de Na_3PO_4 , a pH 8.2, utilizando un voltaje de operación igual a 25 kV.*

Los valores teóricos predichos por el modelo matemático seleccionado se verificaron experimentalmente mediante el análisis de una solución mezcla que contenía 50.0 ng mL^{-1} de ENO, CPF, OFL, ENF y FLU y 167.0 ng mL^{-1} de CIN, sometida a SPE, utilizando las condiciones experimentales óptimas. Los resultados experimentales obtenidos mostraron diferencias estadísticamente no significativas frente a los valores de las respuestas predichas por el modelo, exponiéndose la correcta selección del modelo matemático de ajuste.

Análisis descriptivo del sistema. A partir de los valores teóricos de las constantes de disociación y de la masa molar de los analitos evaluados, fue posible estimar *a priori* el orden de migración electroforética en función de la relación carga/masa que presentan bajo el pH de la solución de trabajo, considerando la polaridad empleada en el sistema. En la Tabla 3.5 se resumen las constantes de disociación y las masas molares de las QUI aquí analizadas.

En las condiciones experimentales óptimas, es decir, a pH = 8.2 y polaridad positiva, todos los analitos presentaron tiempos de migración mayores al tiempo del EOF, respetándose el siguiente orden de migración: ENO, CPF, OFL, ENF, FLU y CIN.

Tabla 3.5. Constantes de disociación y masa molar de las 6 QUI analizadas

PARÁMETRO	ENO	CPF	OFL	ENF	FLU	CIN
$pK_{a,1}$ ^a	6.25	5.86	6.20	5.88	6.35	5.20
$pK_{a,2}$ ^a	8.80	8.24	8.20	7.74	-	-
Masa Molar ($g\ mol^{-1}$) ^b	320.32	331.25	361.37	259.40	261.25	262.22

^a Datos obtenido de (Babić *et al.*, 2007; Speltini *et al.*, 2010);

^b Calculado a partir de la correspondiente estructura química neutra no ionizada.

La Figura 3.11 muestra el perfil electroforético obtenido para una solución mezcla sometida a SPE y analizada en las condiciones experimentales óptimas.

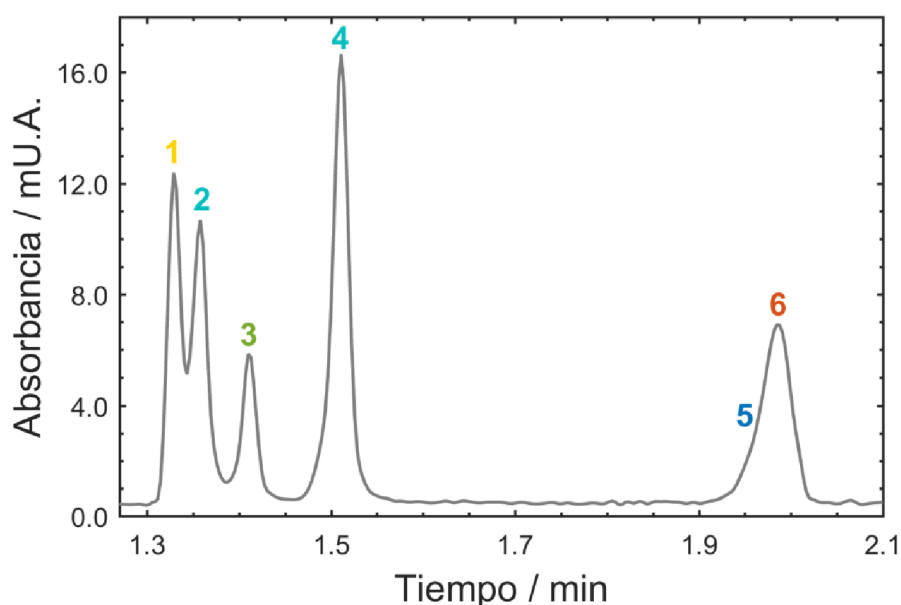


Figura 3.11. Electroferograma obtenido para una solución mezcla de $50.0\ ng\ mL^{-1}$ de (1) ENO, (2) CPF, (3) OFL, (4) (5) ENF y (6) FLU, y $167.0\ ng\ mL^{-1}$ de CIN, sometida a SPE y analizada en condiciones experimentales óptimas. Detección: 280 nm

En la Figura 3.11 puede observarse claramente que los analitos migran en 2 grupos bien definidos. Los 4 que migran primero, ENO, CPF, OFL y ENF, son analitos que a pH 8.2 se encuentran disociados en sus especies aniónica y zwitteriónica, en distintas proporciones dependiendo de su valor de $pK_{a,2}$. Por la otra parte, los 2 analitos que migran a tiempos mayores, FLU y CIN, presentan señales analíticas completamente solapadas. Esto se debe a que, al valor de pH seleccionado, estos compuestos se encuentran completamente ionizados con qn^- y a que presentan relación carga/masa muy similares. Es de notar que el grupo de QUI que migra primero pertenece

a las piperizinil-QUI, mientras que las otras son compuestos pertenecientes al grupo de QUI acidas. Cabe destacar que si bien una de las particularidades que presenta el método electroforético es la baja reproducibilidad en los tiempos de migración, se respetó el orden de migración predicho teóricamente en todas las experiencias realizadas en idénticas condiciones.

Los analitos evaluados en este trabajo son compuestos que pertenecen a una misma familia de fármacos y que presentan ciertas diferencias en su estructura química. Estas diferencias estructurales, como por ejemplo, presencia de flúor en la estructura base o presencia de grupo piperazinil en su núcleo, se ven reflejadas, en parte, en sus espectros de absorbancia UV. Aquí, algunos de los compuestos analizados muestran idénticos espectros de absorbancia, debido en parte a la gran similitud estructural que presentan, como es el caso de CPF y ENF. En la Figura 3.12 se muestran las estructuras químicas de las sustancias analizadas en este trabajo, apreciándose las similitudes y las diferencias estructurales presentes entre ellos.

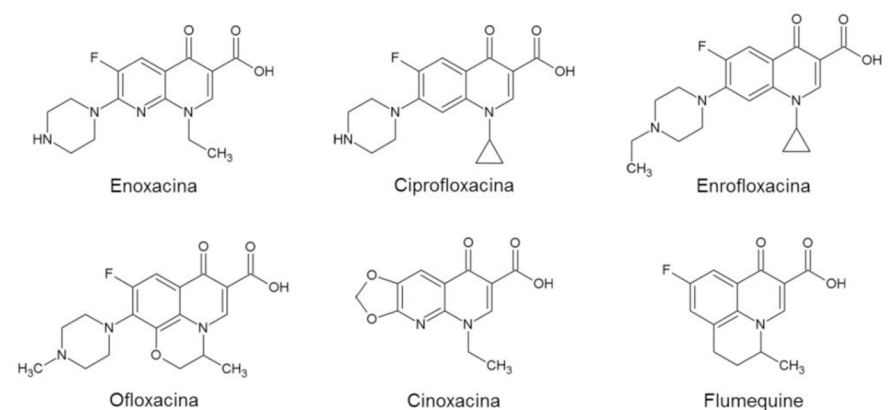


Figura 3.12. Estructura química de las QUI evaluadas

Se puede apreciar que CPF y ENF comparten la misma estructura química, a diferencia de un grupo etil lateral presente en la ENF, demostrándose que la igualdad de sus espectros de absorbancia, se debe, en parte, a la gran similitud estructural que presentan. En la Figura 3.13 se muestran los espectros de absorbancia de los 6 compuestos analizados en condiciones experimentales óptimas.

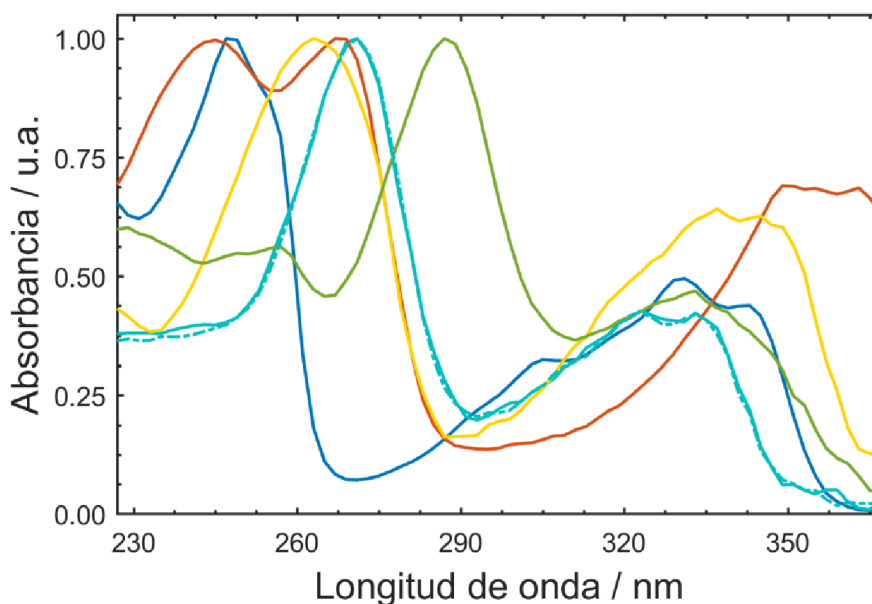


Figura 3.13. Espectros de absorción UV normalizados de ENO (amarillo), CPF (línea continua celeste), ENF (línea celeste de rayas y puntos), OFL (verde), FLU (azul) y CIN (rojo) en 0.015 mol L^{-1} de solución $\text{PO}_4^{-3}:\text{B}_2\text{O}_7^{-2}$ pH 8.2

Modelado quimiométrico por MCR-ALS. El sistema analítico expuesto se presenta como un sistema complejo, cuya complejidad radica principalmente en el solapamiento de las señales electroforéticas, la similitud de los espectros de absorción y la baja reproducibilidad en los tiempos de migración. Esta complejidad se ve aún incrementada ante la presencia de potenciales interferentes que puedan estar presentes en muestras de composición desconocida. Esto puede apreciarse en la Figura 3.14 que muestra el electroferograma obtenido para una muestra conteniendo los 6 QUI en presencia de potenciales interferentes o constituyentes esperados no calibrados (Olivieri *et al.*, 2014b).

En términos quimiométricos, las características del sistema mencionadas se ven reflejadas en una deficiencia de rango espectral y en una falta de trilinealidad para un arreglo de datos de 3 vías de múltiples muestras. En consecuencia, y evaluando los posibles algoritmos disponibles para el modelado y descomposición bilineal, MCR-ALS en su versión extendida (*ext-MCR-ALS*) resultó el algoritmo de elección para lograr la correcta resolución de los componentes y obtener las contribuciones individuales para el análisis predictivo. Ext-MCR-ALS, además de superar el inconveniente de falta de trilinealidad, permite obtener perfiles temporales y espectrales con interpretación física, permitiendo la identificación analítica y visual de los componentes de la muestra analizada.

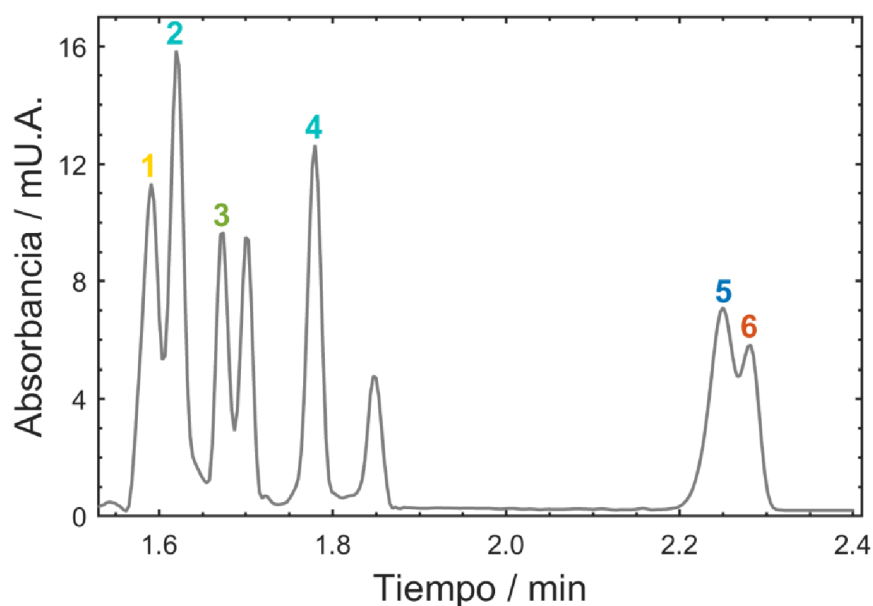


Figura 3.14. Electroferograma obtenido para una muestra de agua de consumo conteniendo 50.0 ng mL^{-1} de (1) ENO, (2) CPF, (3) OFL, (4) ENF y (5) FLU, y 167.0 ng mL^{-1} de (6) CIN, en presencia de potenciales interferentes, sometida a SPE. Detección: 280 nm

En innumerables trabajos, se ha demostrado que MCR-ALS logra explotar satisfactoriamente la llamada ventaja de segundo orden, posibilitando la resolución de los componentes de interés en presencia de potenciales interferentes. Sin embargo, en semejanza a lo ocurrido en el sistema analítico propuesto en la sección 3.1, aplicar un modelado quimiométrico por MCR-ALS a los datos originales no resultó lo suficientemente eficiente para superar la deficiencia de rango espectral, siendo necesario desarrollar una estrategia de análisis que posibilite obtener las contribuciones individuales de cada analito evaluado. La estrategia aquí propuesta se basa, fundamentalmente, en 3 etapas:

*Pre-análisis: aprovechando las ventajas de los diseños experimentales de optimización, se buscaron las condiciones experimentales óptimas que permitiesen la completa resolución de los analitos que tuviesen idénticos espectros de absorbancia (CPF y ENF). Esta etapa se explicó en detalle en el apartado “**Optimización de las condiciones experimentales**” de la presente sección.*

Modelado quimiométrico: durante el ajuste por ALS del modelado quimiométrico, los analitos con idénticos espectros de absorbancia se consideraron como componente único, no aplicándose la restricción de unimodalidad para el factor correspondiente, obteniéndose como resultado un perfil temporal con dos picos electroforéticos.

Post-análisis: obtenidos los perfiles correspondientes a cada componente del sistema, se dividió el perfil temporal con los 2 picos electroforéticos correspondientes a CPF y ENF a fin de obtener las contribuciones individuales de cada componente. Aquí, no fue posible establecer un tiempo único de división del perfil temporal, dado la baja reproducibilidad en los tiempos de migración, por lo que el tiempo seleccionado para la división del perfil temporal fue establecido para cada matriz de datos individualmente y se consideró como el tiempo que represente la máxima resolución entre los componentes involucrados.

Seleccionadas las condiciones experimentales óptimas del sistema electroforético en la etapa de pre-análisis, se procedió al análisis de las muestras para la determinación simultánea de las 6 QUI. A cada matriz de datos MTUM obtenida se la sometió a una etapa de pre-procesamiento que consistió en corrección de la línea de base y reducción del ruido instrumental. La corrección de línea de base se llevó a cabo aplicándose una metodología basada en ajuste por cuadrados mínimos asimétricos propuesto por Eilers (Eilers, 2005), modificada y adaptada para datos de segundo orden, es decir, considerándose la estructura espacial de los datos como arreglo bidimensional. La reducción de ruido instrumental, por su parte, se realizó aplicándose un método de suavizado simple basado en la metodología propuesta por Savitzky-Golay (Savitzky *et al.*, 1964).

Posteriormente, para el modelado quimiométrico por MCR-ALS, se construyeron matrices de datos aumentadas en columnas (D_{aum}) que incluyeron a las matrices de datos correspondientes a las muestras de validación o muestras de agua de consumo, según corresponda, y a las matrices de datos de calibración. Seguidamente, se determinó el número de componentes espectralmente activos presentes en la matriz D_{aum} mediante SVD. Es de destacar, que el número de especies contribuyentes en el sistema analítico evaluado resultó, en todos los casos, igual al número de analitos espectralmente diferentes presentes en la muestra, es decir, se obtuvo un número de factores igual a 5 para una muestra compuesta por las 6 QUI en agua ultrapura. Sin embargo, a fin de mejorar la calidad del ajuste, en todos los casos se consideraron 2 componentes adicionales a los obtenidos por SVD. Así, para una muestra que contenía los 6 compuestos de interés, se consideraron 7 factores. Esta adición en el número de componentes se basa en la presencia de línea de base residual producto del proceso de SPE. Finalmente, se obtuvieron las estimaciones espectrales iniciales mediante la metodología basada en SIMPLISMA (Windig *et al.*, 1991).

Con el objetivo de disminuir la ambigüedad de los resultados finales y obtener perfiles con interpretación física, se aplicaron distintas restricciones durante el proceso iterativo ALS: **1- correspondencia entre especies**; **2- no negatividad en ambos modos**; y **3- unimodalidad en el modo temporal**. Cabe mencionar que la restricción de unimodalidad solo fue aplicada a aquellos componentes que no presentaron deficiencia de rango espectral. En la Figura 3.15 se muestran los perfiles temporales y espectrales obtenidos como resultado de la descomposición bilineal por MCR-ALS para una de las muestras de agua de consumo conteniendo los 6 componentes de interés en presencia de 5 componentes esperados no calibrados.

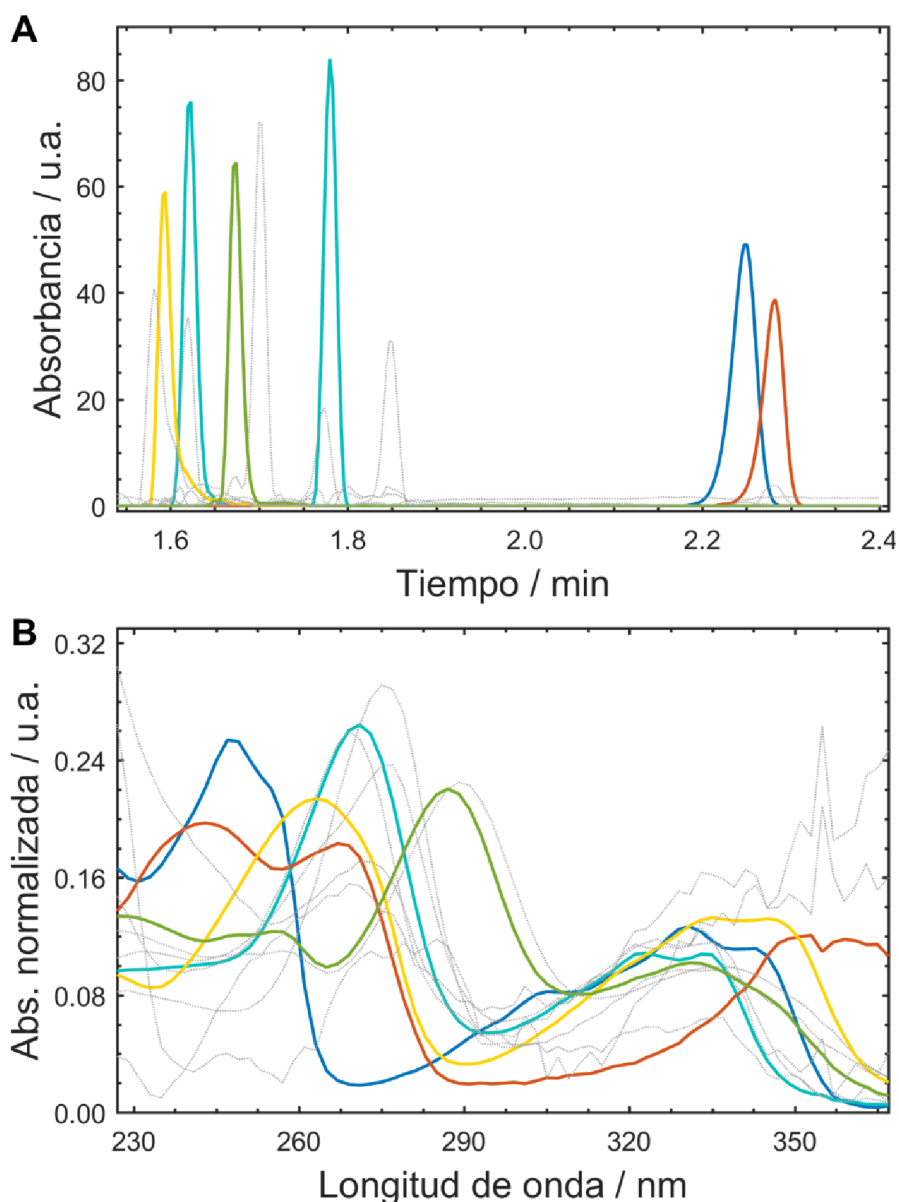


Figura 3.15. Perfiles temporales (A) y espectrales (B) obtenidos por modelado quimiométrico con MCR-ALS para una muestra de agua de consumo conteniendo ENO (amarillo), CPF/ENF (celeste), OFL (verde), FLU (azul) y CIN (rojo), en presencia de interferentes (línea de puntos gris)

Para obtener las contribuciones individuales de cada componente y proceder al correcto análisis predictivo, se dividieron los perfiles temporales correspondientes a CPF/ENF, como se indica en la Figura 3.16.

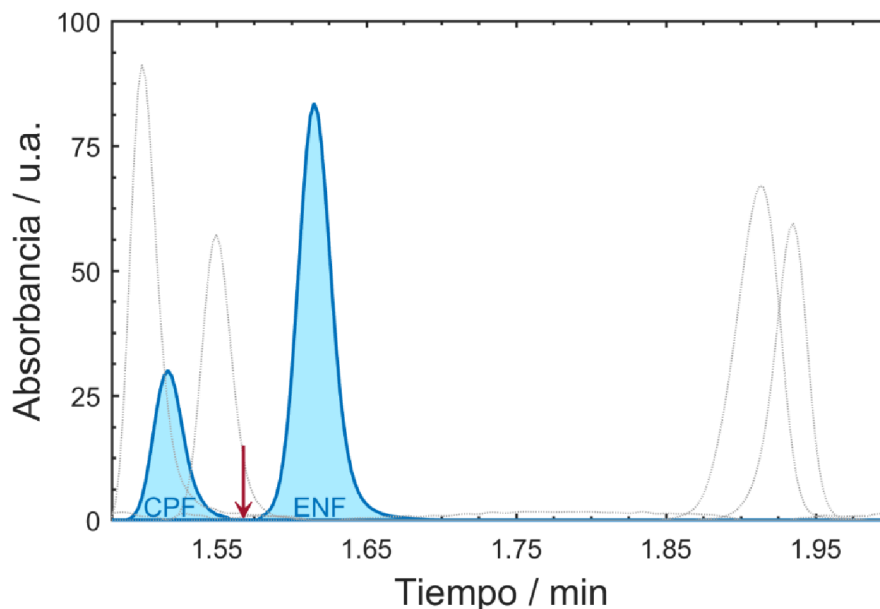


Figura 3.16. Perfiles temporales obtenidos por modelado con MCR-ALS para una muestra de validación conteniendo los 6 compuestos de interés. El perfil de CPF y ENF se indica en azul, mientras que en líneas de punto gris se muestran los perfiles obtenidos para los demás analitos presentes en la muestra. La flecha roja señala el punto de división seleccionado para obtener las contribuciones individuales de CPF y ENF

Análisis cuantitativo. Obtenidas las matrices **C** y **S** correspondientes al resultado de la descomposición bilineal, se utilizó la información contenida en **C** para la construcción de una curva de calibración tipo univariada, o pseudounivariada, a fin de predecir las concentraciones de los componentes presentes en muestras de validación y de agua de consumo.

La Tabla 3.6 resume los valores de predicción obtenidos para las 6 QUI analizadas presentes en las muestras de validación. En todos los casos, los valores de REP % resultaron menores al 20 % y los valores de recuperación media oscilaron entre 80 % y 110 %. Con el test de hipótesis realizado para evaluar la exactitud de las recuperaciones se demostró que los valores de recuperación media obtenidos experimentalmente y el valor ideal de recuperación 100 % presentaron diferencias estadísticamente no significativas, para un nivel de confianza del 95 % ($\alpha = 0.05$) y $\nu = n-1$ grados de libertad ($\nu = 4$) ($t_{(0.05,4)} = 2.776$) (Olivieri, 2015).

A fin de evaluar la capacidad predictiva del método en matrices acuosas de composición variable, se analizaron muestras de agua de consumo que contenían

todos los analitos en presencia de potenciales interferentes. Los blancos de muestra de agua de consumo analizados no mostraron señales detectables de ninguna de las QUI estudiadas.

Tabla 3.6. Valores de concentración predicha en muestras de validación para las 6 QUI determinadas simultáneamente ^a

Muestra	ENO		CPF		OFL		ENF		FLU		DIF	
	NOM	PRED	NOM	PRED	NOM	PRED	NOM	PRED	NOM	PRED	NOM	PRED
M01	100.0	80	20.0	21	40.0	44	60.0	57	20.0	17	333.0	287
M02	80.0	76	20.0	17	100.0	80	40.0	50	100.0	110	67.0	80
M03	60.0	47	80.0	88	100.0	103	20.0	21	40.0	44	200.0	204
M04	40.0	40	80.0	83	60.0	62	100.0	103	80.0	98	133.0	147
M05	50.0	50	50.0	60	50.0	48	50.0	52	50.0	49	167.0	150
REP %^b	18.0		10.1		15.5		8.3		15.8		11.6	
\bar{R}_{exp}^c	90.6		104.8		98.5		106.4		105.1		101.6	

^a Las concentraciones nominales (Nom) y predichas (Pred) están dadas en ng mL⁻¹;

^b REP %: error relativo de predicción, expresado en %; $REP \% = 100 \times \frac{\sqrt{\frac{1}{l} \sum_1^l (c_{nom} - c_{pred})^2}}{\bar{c}}$, para l= 5;

^c \bar{R}_{exp} , Valor de recuperación media, expresado en %.

Los valores de concentración encontrados en las muestras de agua de consumo fortificadas se resumen en la Tabla 3.7. Se puede observar que los valores de recuperación media oscilan entre 85 % y 125 %, cumpliendo con los estándares recomendados para el análisis de fármacos en muestras de agua de consumo (USEPA, 2007). Además, se evaluaron estadísticamente las recuperaciones experimentales medias mediante el test de hipótesis *t* (Olivieri, 2015), dejando expuesto que los valores de recuperación media obtenidos y el valor ideal 100 % presentaron diferencias estadísticamente no significativas para todas las QUI analizadas, para un nivel de confianza del 95 % ($\alpha = 0.05$) y $\nu = n-1$ grados de libertad ($\nu = 3$). Con base en los resultados obtenidos es posible asegurar la exactitud del método propuesto para la determinación simultánea de 6 QUI en muestras de validación y de agua de consumo, en presencia de potenciales interferentes.

La robustez del método desarrollado se evaluó introduciendo pequeñas variaciones en las condiciones experimentales optimizadas, utilizando un diseño experimental Plackett-Burman en 2 niveles. Las variaciones introducidas en las condiciones óptimas

fueron: **1-** $pH = 8.15$ y 8.25 ; **2-** $voltaje = 24\text{ kV}$ y 26 kV ; y **3-** $proporci3n\ PO_4^{3-}:B_4O_7^{3-}$ de la soluci3n $BGE = 0.0045:0.0105\text{ mol L}^{-1}$ y $0.0055:0.0095\text{ mol L}^{-1}$. El an3lisis se llev3 a cabo analizando las concentraciones predichas de la muestra T_01 en las distintas condiciones establecidas por el dise1o experimental. Los resultados obtenidos demostraron que posibles variaciones de los factores experimentales no generan efectos significativos que pudiesen alterar la calidad predictiva del m3todo propuesto.

Tabla 3.7. Valores de concentraci3n encontrados de QUI en muestras de agua de consumo ^a

Muestra	ENO		CPF		OFL		ENF		FLU		DIF	
	ADIC.	ENC.	ADIC.	ENC.	ADIC.	ENC.	ADIC.	ENC.	ADIC.	ENC.	ADIC.	ENC.
T_01	80.0	68	40.0	69	60.0	58	80.0	91	60.0	49	133.0	117
T_02	60.0	51	80.0	79	40.0	47	60.0	48	40.0	36	267.0	233
T_03	40.0	47	60.0	53	80.0	70	40.0	36	80.0	61	200.0	163
T_04	50.0	53	50.0	49	50.0	53	50.0	49	50.0	46	167.0	162
\bar{R}_{exp}^b	88.6		114.4		101.9		95.5		89.2		90.3	

^a Los valores de concentraci3n est3n dados en $ng\text{ mL}^{-1}$. Cada valor de concentraci3n encontrado se expresa como el valor medio de tres r3plicas de preparaci3n. Adic.: adicionada; Enc.: encontrada;

^b \bar{R}_{exp} , Recuperaci3n media, expresada en %.

Finalmente, se evalu3 cuantitativamente la calidad y eficiencia del modelado quimiom3trico calculando el grado de solapamiento espectral (s_{12}) entre los espectros puros de cada analito (s_1) y los perfiles espectrales obtenidos por MCR-ALS (s_2), siguiendo la ecuaci3n 3.1. Los espectros de los analitos puros se obtuvieron a partir de las se1ales electrofor3ticas puras de los analitos evaluados individualmente en las muestras de calibraci3n al m3ximo nivel de concentraci3n. Para todas las QUI se obtuvieron valores de s_{12} cercanos al valor de m3ximo solapamiento, es decir, $s_{12}=1.000$, asegur3ndose, en consecuencia, la eficiencia del modelado quimiom3trico y la capacidad de obtener informaci3n espectral de alta calidad.

Es importante mencionar que a pesar de la complejidad del sistema an3lítico presente, la separaci3n electrofor3tica se realiz3 mediante un m3todo libre de solventes org3nicos contaminantes, en un tiempo no mayor a 3.0 min. De esta manera, el m3todo desarrollado queda enmarcado dentro de los lineamientos de la qu3mica an3lítica verde (Armenta *et al.*, 2008; Gałuszka *et al.*, 2013).

Cifras de mérito analíticas. La Tabla 3.8 muestra las FOMs analíticas obtenidas para las 6 QUI evaluadas utilizando el método analítico desarrollado y estimadas según 3 estrategias diferentes. Aquí, se utilizaron las mismas estrategias de estimación de FOMs que se propusieron en la sección 3.1 para el método cromatográfico. La *Estrategia I* se basa en lo postulado por Bauza *et al.* (Bauza *et al.*, 2012) considerándose la dimensión de los datos y el tipo de algoritmo utilizado para el modelado quimiométrico; la *Estrategia II* sigue el principio de calibración univariada clásica (Saurina *et al.*, 2001) estimándose las FOMs según sugieren las guías ICH (ICH, 2005); y la *Estrategia III* se apoya en las recomendaciones de las guías ICH para métodos analíticos que exhiben línea de base (ICH, 2005), considerándose la SNR entre la señal analítica de una solución blanco y la de un analito en su mínima concentración detectable.

Tabla 3.8. Valores de límite de detección y cuantificación obtenidos para las QUI analizadas siguiendo tres estrategias de estimación ^a

Analito	LOD			LOQ		
	Estrategia I ^b	Estrategia II ^c	Estrategia III ^d	Estrategia I	Estrategia II	Estrategia III
ENO	6.9	11.2	6.7	20.9	33.9	22.3
CIP	11.0	17.5	10.4	33.3	53.0	34.6
OFL	3.9	5.9	3.7	11.8	17.9	12.3
ENF	10.3	16.5	9.8	31.2	50.0	32.6
FLU	4.2	5.0	4.3	12.7	15.2	14.3
CIN	21.9	30.4	23.5	66.4	85.9	78.3

^a Valores de concentración dados en ng mL⁻¹;

^b *Estrategia I*, valores estimados según referencia (Bauza *et al.*, 2012);

^c *Estrategia II*, valores estimados según lo propuesto en por (Saurina *et al.*, 2001);

^d *Estrategia III*, valores estimados según recomendaciones de las guías ICH Q2R1, (ICH, 2005).

Puede apreciarse que las mejores estimaciones se lograron utilizando la *Estrategia I*, remarcando nuevamente la importancia de considerar el orden dimensional de los datos y el algoritmo utilizado para el modelado quimiométrico. Adicionalmente, los valores de LOD y LOQ estimados según la *Estrategia I* son coincidentes en orden de magnitud con los estimados según la *Estrategia III*, exponiéndose la posibilidad de obtener estimaciones realistas de las FOMs que caracterizan a un método analítico utilizando herramientas de calibración multivariada de segundo orden.

Con base en los resultados mostrados se propone el presente método como una alternativa a los métodos de rutina utilizados en los laboratorios para el monitoreo de

contaminantes emergentes en agua de consumo. Los valores alcanzados resultan comparables a los reportados por Cañada-Cañada *et al.* (Cañada-Cañada *et al.*, 2009) para la determinación de 8 f-QUI en muestras acuosas, utilizando, aquí, un método que demanda un tiempo de análisis significativamente menor y sin involucrar el uso de solventes orgánicos. Es de destacar que si bien los niveles de concentración reportados en otros trabajos previos, que utilizan detectores más sensibles y selectivos, resultan inferiores a los alcanzados en este trabajo (Speltini *et al.*, 2010; Ye *et al.*, 2007), el volumen utilizado en la etapa de pre-concentración puede ser incrementado, disminuyéndose, consecuentemente, los límites de concentración detectables. Experiencias preliminares demostraron que volúmenes de muestra de hasta 250.0 mL no superan el volumen de ruptura del cartucho utilizado para la SPE, lo que posibilitaría lograr factores de pre-concentración aún mayores a los aquí reportados y así llegar a obtener valores de LOD y LOQ inferiores a los informados.

CONCLUSIÓN

Se desarrolló un método electroforético con detección DAD-UV acoplado a modelado quimiométrico por MCR-ALS para la determinación simultánea de 6 QUI en muestras de agua de consumo. Como etapa de pre-concentración se propuso una metodología basada en extracción en fase sólida, obteniéndose un factor de pre-concentración igual a 100, quedando sujeto al volumen de muestra analizado. Se demostró, también, que muestras de agua de consumo conteniendo un alto número de componentes pudieron ser analizadas en un tiempo de electroforesis no mayor a 3.0 min, sin involucrar solventes orgánicos que resultan contaminantes para el medioambiente.

El sistema analítico expuesto presentó una alta complejidad basada en la similitud espectral de los analitos, el solapamiento de las señales electroforéticas, la baja reproducibilidad de los tiempos de migración y la presencia de interferentes en las muestras. La selección del diseño experimental apropiado, la aplicación del algoritmo correcto para el modelado quimiométrico de los datos y el desarrollo de una minuciosa estrategia de post-análisis mostraron ser una excelente combinación para superar la complejidad expuesta por el sistema y obtener resultados altamente satisfactorios para el análisis simultáneo de 6 QUI presentes en muestras acuosas.

Es preciso mencionar, también, que la complejidad mostrada por el sistema ana-

lítico presente resulta un desafío para aquellos algoritmos que no permiten desvíos de trilinealidad en los datos o presentan deficiencia de rango en alguno de sus modos. Asimismo, debido a la necesidad de obtener perfiles con interpretación física, los algoritmos basados en variables latentes no resultarían eficientes para superar la complejidad de este tipo de datos. MCR-ALS demostró una vez más su capacidad para superar situaciones que presenten desviaciones de trilinealidad y deficiencia de rango espectral, resolviendo de manera eficiente las señales de los compuestos de interés en presencia de componentes no modelados, o potenciales interferencias presentes en muestras de composición desconocida.

Por consiguiente, el método propuesto de pre-concentración se muestra como una metodología simple y eficiente para lograr altos factores de pre-concentración, que pueden ser aun incrementados aumentando del volumen de muestra utilizado. Las cifras de mérito logradas, en adición a los resultados predictivos obtenidos en las muestras analizadas, convierten al método propuesto en una excelente alternativa a las técnicas utilizadas en laboratorios de rutina para la determinación de 6 QUI en agua de consumo.

3. 3. MÉTODO DE GRADIENTE DE PH POR SISTEMA DE INYECCIÓN EN FLUJO ACOPLADO A MODELADO DE DATOS DE SEGUNDO ORDEN PARA EL ANÁLISIS DE DOXORRUBICINA EN PLASMA

Modelling of second-order spectrophotometric data generated by a pH-gradient flow injection technique for the determination of doxorubicin in human plasma –Alcaráz, M.R., et al., Microchem J 2014, 112, 25–33

CONSIDERACIONES GENERALES

El crecimiento de la población mundial y los nuevos estilos de vida adoptados por los habitantes ha llevado consigo, entre otras cosas, un rápido crecimiento de epidemias, como también el desarrollo de nuevos tipos de enfermedades. Un tópico particularmente alarmante con base en este fenómeno es el constante aumento de nuevos casos asociados a tumores y enfermedades cancerígenas.

El cáncer es una de las principales causas de deceso a nivel mundial y es una enfermedad caracterizada por el crecimiento y distribución no controlados de células anómalas y que es causado por factores internos o externos, tales como mutaciones adquiridas por radiación o tabaquismo, e incluso por el tipo de alimentación (Sriraman *et al.*, 2016; Tacar *et al.*, 2013). El tratamiento y control de esta enfermedad dependen del tipo de cáncer, el lugar donde se radica y el estadio de crecimiento, entre otros factores. Las técnicas basadas en quimioterapia son utilizadas como tratamiento primario y han sido globalmente empleadas para tratar a pacientes que padecen esta enfermedad. Estas metodologías se basan en la administración de drogas que permiten la erradicación o reducción del número de células neoplásicas en el paciente (Sriraman *et al.*, 2016).

La US FDA ha aprobado una amplia variedad de drogas para el tratamiento de cáncer, siendo las antraciclinas la familia de antibióticos más utilizada en las últimas décadas (Carvalho *et al.*, 2009). Una de las principales antraciclinas administradas para quimioterapia en pacientes enfermos es la doxorubicina (DXR), un compuesto que se aisló de *Streptomyces peacetius*, por primera vez, en la década del 60' y que aún en la

actualidad se sigue empleando (Carvalho *et al.*, 2009; Hortobágyi, 1997; Trevisan *et al.*, 2003). La DXR actúa por unión a enzimas asociadas al ADN, tales como topoisomerasa I y II, que pueden intercalarse entre los pares de bases de la doble hélice del ADN. En consecuencia, se produce un desencadenamiento de efectos citotóxicos que llevan consigo una anti-proliferación de las células tumorales, resultado del daño ocasionado en el ADN (Hortobágyi, 1997; Tacar *et al.*, 2013). Se demostró, también, que las moléculas de DXR son capaces de intercalarse en el ADN nuclear y mitocondrial con la consecuente inhibición de la ADN y ARN polimerasa, impidiendo la replicación del ADN, como así también, la transcripción del ARN (Carvalho *et al.*, 2009; Hortobágyi, 1997; Tacar *et al.*, 2013).

Si bien la DXR ha demostrado ser una droga muy eficiente para el tratamiento de diferentes tipos de cáncer y tumores, también se demostró que, al ser una droga que no actúa específicamente sobre células tumorales, puede afectar el crecimiento de otro tipo de células del organismo humano (Hajian *et al.*, 2016; Tomoaia *et al.*, 2015). Esta acción desencadena una serie de efectos colaterales indeseables que conducen a un organismo inmunodeprimido, reduciéndose el número de células inmunes y aumentando la susceptibilidad del paciente a infecciones microbianas u otras afecciones. La gravedad de estos efectos depende, principalmente, de la dosis suministrada de la droga y de la capacidad de regeneración de médula ósea del paciente (Tacar *et al.*, 2013). En virtud de sus efectos citotóxicos, la determinación de DXR en plasma humano resulta de gran interés a fin de reducir los riesgos a las acciones colaterales en pacientes bajo tratamiento.

Entre los métodos analíticos reportados en la bibliografía para la determinación y cuantificación de DXR en plasma pueden encontrarse metodologías basadas en cromatografía líquida (Pérez-Blanco *et al.*, 2014), electroforesis con detección fluorescente (Yang *et al.*, 2016), sensores electroquímicos (Hajian *et al.*, 2016) y técnicas espectroscópicas de fluorescencia, acopladas a herramientas quimiométricas de calibración multivariada (Trevisan *et al.*, 2003). Aquí, se propone un método simple basado en un análisis por inyección en flujo y detección espectral de absorbancia visible, utilizando un detector de arreglo de diodos, para la determinación de DXR en plasma humano, obteniéndose datos de segundo orden por generación de un gradiente de pH. Para el modelado quimiométrico, la cuantificación del analito y la obtención de información cualitativa acerca del sistema en estudio se utilizaron diferentes herramientas quimiométricas, como MCR-ALS y PKFIT.

MATERIALES Y MÉTODOS

Reactivos. La droga sólida de doxorubicina hidrocloreto (*DXR*, 94.0%) fue provista por los laboratorios Richmond S.A.C.I.F. (Pilar, Argentina) y la de rifampicina (*RFP*, 99.4%) fue gentilmente provista por el Laboratorio de Control de Calidad de Medicamentos de la Facultad de Bioquímica y Ciencias Biológicas de la Universidad Nacional del Litoral (Santa Fe, Argentina). El ácido clorhídrico (*HCl*), hidróxido de sodio (*NaOH*), ácido fosfórico (*H₃PO₄*) y carbonato de sodio (*Na₂CO₃*), todos de calidad grado analítico, fueron adquiridos de Cicarelli (San Lorenzo, Argentina). El ácido tricloroacético (*TCA*) de calidad grado analítico fue provisto por Anedra (San Fernando, Argentina). El agua de laboratorio ultrapura fue obtenida por un sistema de purificación de agua Milli-Q MilliPore (Billerica, Estados Unidos).

Soluciones de trabajo. Se preparó una solución madre conteniendo 70.0 $\mu\text{g mL}^{-1}$ de *DXR* por disolución de una cantidad apropiada de la droga sólida en agua ultrapura. Paralelamente, se preparó una solución madre de *RFP* disolviendo la cantidad correspondiente de la droga sólida en *H₃PO₄* 1.0 %, a fin de obtener una concentración final de 620.0 $\mu\text{g mL}^{-1}$. Para su conservación, las soluciones se almacenaron a 4 °C protegida de la luz por un periodo no mayor a 3 días.

Para el análisis por inyección en flujo (*FIA: Flow Injection Analysis*) se preparó una solución reguladora de *Na₂CO₃/HNaCO₃* que luego se utilizó como solución reguladora transportadora (*SRT*), disolviendo una cantidad apropiada de *Na₂CO₃* en agua ultrapura y ajustando el pH a un valor de 9.5, lográndose una concentración formal de 0.02 mol L⁻¹. La solución reguladora utilizada como solución de inyección (*SRI*), o solución de preparación de muestras, se preparó disolviendo una cantidad apropiada de *Na₂CO₃* y ajustando el pH a un valor de 6.0, obteniéndose una concentración formal de 0.02 mol L⁻¹.

Las muestras de plasma heparinizado de pacientes voluntarios sanos no tratados fueron provistas gentilmente por el Hospital Iturraspe de la ciudad de Santa Fe, Argentina, y se almacenaron a -20 °C hasta el momento del análisis.

Muestras de calibración, validación y de prueba. Para la calibración se preparó 1 set compuesto por 5 soluciones estándares de *DXR* en distintos niveles de concentración y para la validación se preparó un set de 11 soluciones

estándares de DXR siguiendo un diseño aleatorio, en diferentes niveles de concentración, con valores diferentes a los utilizados para la calibración, si bien comprendidos en el mismo rango de concentraciones.

La preparación de todas las soluciones se realizó transfiriendo una alícuota apropiada de solución madre de DXR a matraces aforados de 2.0 mL, completándose el volumen con plasma humano heparinizado. Para garantizar el contacto y la interacción del analito con la matriz, las soluciones preparadas se sometieron a una agitación mecánica durante 1 min y se dejaron en reposo durante 30.0 min a temperatura ambiente para favorecer la estabilización de la muestra. La desproteinización del plasma se logró adicionando 200.0 μL de una solución de TCA 50 % a la solución preparada, centrifugándose, posteriormente, por un periodo de 20 min a una velocidad de 3000 r.p.m. Una alícuota de 600.0 μL de la solución resultante y 200.0 μL de solución SRI se transfirieron a tubos Eppendorf® y se centrifugaron durante un periodo de 5.0 min a una velocidad de 10000 r.p.m. Al final del procedimiento, las concentraciones de las muestras de calibración y validación estuvieron contenidas en el rango de concentración comprendido entre 0.0 $\mu\text{g mL}^{-1}$ y 12.0 $\mu\text{g mL}^{-1}$ de DXR, y el pH de todas las soluciones resultó igual a 6.0.

Para evaluar la capacidad predictiva del método para la determinación de DXR en presencia de potenciales interferentes se prepararon 2 soluciones mezcla de DXR y RFP en distintos niveles de concentración, considerándose a RFP como constituyente esperado no calibrado, una potencial interferencia que pudiese estar presente en muestras de plasma de pacientes tratados.

Todas las soluciones se prepararon diariamente por triplicado siguiendo el mismo procedimiento de preparación.

Finalmente, se transfirió un alícuota de 500.0 μL de muestra a viales de inyección, inyectándose 100.0 μL de muestra en el sistema FIA.

Instrumentación y procedimiento. Se desarrolló un sistema de inyección en flujo adaptando un instrumento Agilent 1100 LC (Agilent Technologies, Waldbronn, Alemania) equipado con desgasificador, bomba cuaternaria, inyector automático, compartimento de control de temperatura, DAD-UV-Vis, FSFD y el programa informático ChemStation (Agilent Technologies, Waldbronn, Alemania) requerido para el control del instrumento y la adquisición de los datos.

Los experimentos de FIA se llevaron a cabo utilizando un capilar flexible de ace-

ro no recubierto de 80.0 cm de largo y 0.5 mm de diámetro interno, operando en modo isocrático con una velocidad de flujo de 0.5 mL min^{-1} , durante 1 min. Las matrices de datos correspondientes a tiempo de elución-espectro UV-Vis (*ETUV: Elution Time-UV-Vis*) se registraron en el rango espectral comprendido entre 350.0 nm y 650.0 nm, en el tiempo de elución comprendido entre 0.0 min y 1.0 min. De esta manera, las ETUVM consistieron en 174×251 datos para la dimensión temporal y espectral, respectivamente.

La titulación espectrofotométrica se realizó sobre una solución conteniendo $23.0 \mu\text{g mL}^{-1}$ de DXR en $0.1 \text{ mol L}^{-1} \text{ H}_3\text{PO}_4$ pH 4.0. El procedimiento general involucró la adición de alícuotas de $5.0 \mu\text{L}$ de NaOH de concentración 0.01 mol L^{-1} o 5.0 mol L^{-1} , según corresponda para lograr pequeños incrementos de pH, a 50.0 mL de solución de DXR bajo agitación constante y con monitoreo continuo de pH. Adicionada una alícuota de NaOH correspondiente y registrado el respectivo valor de pH, se tomó una alícuota de 2.0 mL de solución, se transfirió a una cubeta de cuarzo 1 cm de paso óptico y se colocó en un espectrofotómetro Perkin-Elmer Lambda 20 (Waltham, Estados Unidos) para la adquisición del espectro correspondiente, cubriéndose el rango espectral comprendido entre 350.0 nm y 650.0 nm. Posteriormente, la alícuota se transfirió cuantitativamente a la solución original a fin de no introducir cambios en la concentración final del analito. Este procedimiento se realizó hasta alcanzar un valor de pH de la solución de aproximadamente 14.0. Finalmente, la matriz de datos obtenida consistió en 40×150 datos para la dimensión correspondiente al pH y la dimensión espectral, respectivamente. Para el control del instrumento y la adquisición de los datos se utilizó el programa informático UV WinLab (Perkin-Elmer Inc., Waltham, Estados Unidos).

Las mediciones de pH se realizaron utilizando un potenciómetro Orion 410A (Massachusetts, Estados Unidos) equipado con un electrodo de vidrio combinado Boeco BA17 (Hamburgo, Alemania).

Programas informáticos y algoritmos para el análisis de datos. El procesamiento de datos y el análisis quimiométrico con MCR-ALS y PKFIT se realizaron en MatLab 7.10 (MathWorks Inc., Natick, Estados Unidos). Los algoritmos necesarios para operar MCR-ALS se encuentran asequibles en internet accediendo a la página web <http://www.mcrals.info/>. Las constantes de disociación en condiciones experimentales se determinaron usando algoritmos quimiométrico compilados en PKFIT (Ibañez *et al.*, 2007).

RESULTADOS

Propiedades ácido-base de DXR. La DXR es un compuesto que, en su estructura química, presenta diferentes átomos de hidrógeno que pueden ser ionizados y también un átomo de nitrógeno de fácil protonación. Consecuentemente, en un equilibrio protolítico en medio acuoso se pueden distinguir 4 especies iónicas principales dependientes del pH del medio. A bajos valores de pH, la deprotonación del grupo amino, cargado positivamente (DXR^+), ubicado en posición 4', lleva consigo la formación de la especie neutra del compuesto (DXR), que, ante el incremento de pH, se ioniza por deprotonación del grupo hidroxilo fenólico en posición 11 (DXR^-) con posterior deprotonación de un segundo grupo hidroxilo fenólico ubicado en posición 6 (DXR^{2-}). En la Figura 3.17 se muestran los principales equilibrios ácido-base presentes en todo el rango de pH para DXR. En la literatura consultada (Abraham *et al.*, 2005; Sturgeon *et al.*, 1977; Tomoaia *et al.*, 2015) se reporta que las constantes de disociación asociadas a la formación de estas 4 especies iónicas principales presentan los siguientes valores: **1- equilibrio $\text{DXR}^+ / \text{DXR}$, $pK_{a,1}=8.15$; 2- equilibrio $\text{DXR} / \text{DXR}^-$, $pK_{a,2}=10.16$; y 3- equilibrio $\text{DXR}^- / \text{DXR}^{2-}$, $pK_{a,3}=13.2$.**

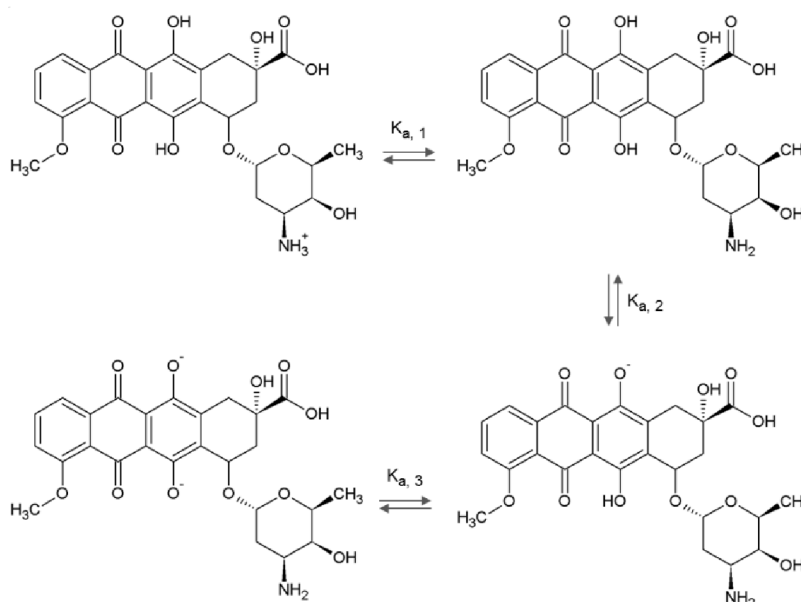


Figura 3.17. Esquema propuesto para los principales equilibrios ácido-base de DXR en medio acuoso, con $pK_{a,1} = 8.15$, $pK_{a,2} = 10.16$ y $pK_{a,3} = 13.2$

Los valores de las constantes de disociación ácido-base pueden presentar pequeñas diferencias según la metodología analítica empleada para su determinación y

las condiciones experimentales utilizadas.

En este trabajo, con base en la información obtenida de la literatura y con el objetivo de estudiar el comportamiento ácido-base del compuesto en virtud de sus características espectrales, se realizó una titulación espectrofotométrica, registrándose los espectros Vis correspondientes a los distintos valores de pH comprendidos en el rango de 4.0 y 14.0. La obtención de las constantes de disociación, $pK_{a,1}$, se logró por análisis de los datos de la titulación espectrofotométrica con PKFIT por ajuste por cuadrados mínimos alternantes para datos de segundo orden. Los perfiles espectrales y las curvas de distribución de especies obtenidas se muestran en la Figura 3.18, en la que se pueden apreciar 3 especies principales en el equilibrio ácido-base con valores de constantes de disociación igual a $pK_{a,1}=8.84$ y $pK_{a,2}=11.36$. Estas cifras son consistentes con las reportadas en la literatura, si bien el valor de $pK_{a,2}$ presenta una diferencia por exceso frente al valor de referencia.

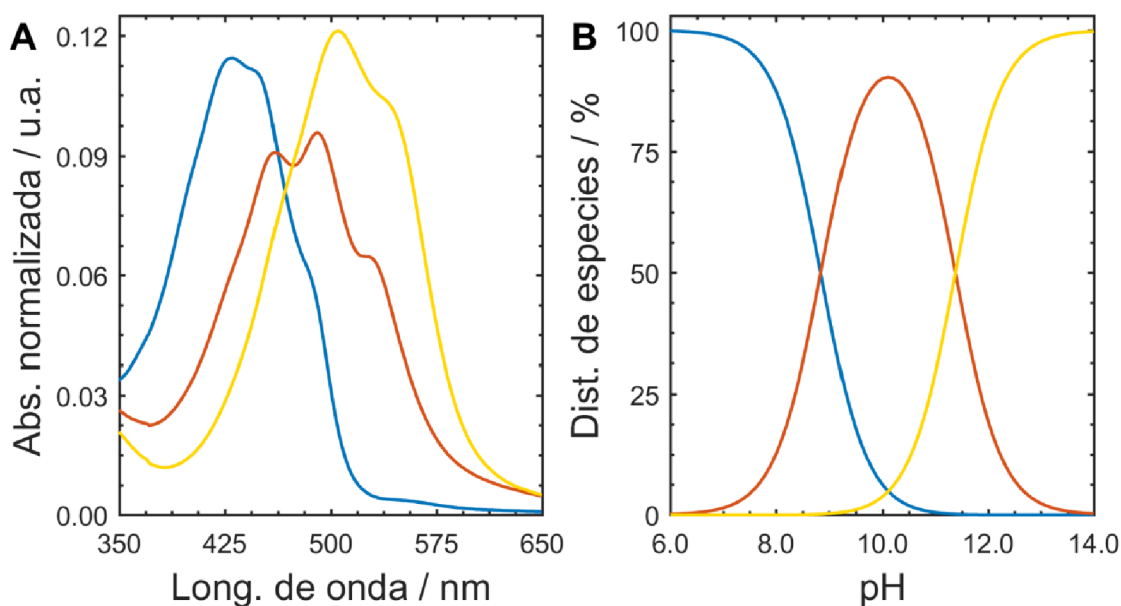


Figura 3.18. Perfiles espectrales (A) y de distribución de especies (B) obtenidos por PKFIT para DXR en el rango de pH de entre 6.0 y 14.0. Se distinguen 3 especies: DXR⁺ (azul), DXR (rojo) y DXR⁻ (amarillo)

Es de remarcar, además, que con esta metodología no fue posible identificar el valor correspondiente a la constante de disociación $pK_{a,3}$, el que es cercano al valor límite superior del rango de pH. La diferencia entre los valores de $pK_{a,2}$ experimental y teórico, como así también la deficiencia para identificar el valor de $pK_{a,3}$, puede deberse a que a pH mayores de 10.0, la presencia de iones alcalinos, principalmente Na⁺, conduce a una sobreestimación de los valores de pH cuando se utiliza un electrodo de

vidrio para la medición de pH, es decir, los valores obtenidos experimentalmente resultan mayores a los valores de pH reales de la solución (Oláh, 1960).

Sobre la base de los resultados obtenidos y teniendo en cuenta que el instrumento disponible para llevar a cabo el sistema de inyección de flujo no permite trabajar a valores de pH superiores a 9.5, todos los experimentos de gradiente de pH por FIA se realizaron en el rango de pH comprendido entre 6.0 y 9.5. En consecuencia, solo fue posible identificar a las especies involucradas en equilibrio $\text{DXR}^+ / \text{DXR}$, con un valor experimental de $\text{pK}_{a,1}$ asociado igual a 8.84.

Modelado quimiométrico por MCR-ALS y análisis descriptivo del sistema FIA. Para determinar el gradiente de pH generado en el sistema de FIA se registraron matrices ETUV para 3 réplicas de inyección en 3 niveles de concentración de DXR en plasma, analizándolas individualmente por MCR-ALS. Para obtener resultados confiables y llegar a conclusiones acertadas respecto al comportamiento del sistema, fue importante considerar, en primer lugar, algunas características inherentes del sistema.

El concepto base de un sistema FIA es la inyección de una muestra líquida en un flujo líquido transportador continuo que genera una zona, o bolo de muestra, que es transportada a través de un capilar y que es luego detectada por un sistema de detección (Saurina, 2010). Dependiendo de las características de la solución transportadora y de la muestra, es posible obtener un sistema lineal o un sistema de gradiente. En este trabajo, debido a que el pH de la SRT y el de la SRI fueron diferentes (9.5 y 6.0, respectivamente) se obtuvo un sistema de doble gradiente de pH, generado por la dispersión de la muestra dentro del flujo transportador. En la Figura 3.19 se esquematiza el sistema FIA utilizado, representándose el perfil de elución obtenido de la inyección de una muestra ácida en un flujo transportador alcalino.

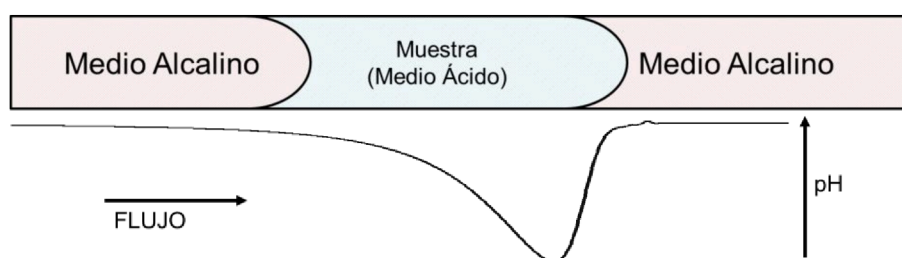


Figura 3.19. Representación esquemática de un sistema FIA para el análisis de una muestra ácida en un flujo transportador alcalino

En una reacción química en equilibrio se obedece la ley de balance de masa o conservación de masa, es decir, que la masa total de un compuesto, que es igual a la sumatoria de todas las especies involucradas en el equilibrio, se mantiene invariable en todo el proceso. En el caso del equilibrio ácido-base de DXR en estudio, la concentración analítica total del analito, C_{DXR} , es igual a la suma de las concentraciones correspondientes a la especie DXR^+ y DXR ; así,

$$C_{\text{DXR}} = [\text{DXR}^+] + [\text{DXR}] \quad 3.2$$

Por otro lado, en el sistema FIA, la inyección de la muestra en el flujo transportador conlleva la dispersión del analito, de concentración C_{DXR} , a lo largo del bolo de muestra. Es importante remarcar que si bien se genera una dispersión del bolo, la concentración total C_{DXR} se conserva, obedeciéndose la ley de balance de masa a cada tiempo de elución. Una excelente estimación del perfil de C_{DXR} en el bolo de muestra se logra analizando las variaciones de intensidad de la señal de absorción de la muestra a la longitud de onda correspondiente al punto isobéptico (Ando *et al.*, 1997; Saurina *et al.*, 2000). En el punto isobéptico, las especies involucradas en el equilibrio presentan idénticos valores de absorptividad molar, por lo que la señal de absorción total registrada dependerá de la concentración total C_{DXR} y no de las contribuciones individuales de las especies involucradas. De esta manera, se puede obtener el perfil de dispersión de la muestra en el flujo transportador independiente del valor de pH.

Experimentalmente, el valor de longitud de onda igual al punto isobéptico se obtuvo de los perfiles espectrales obtenidos por PKFIT, siendo $\lambda_{\text{pi}} = 530 \text{ nm}$. Posteriormente, el perfil de dispersión de la C_{DXR} en el punto isobéptico para DXR se extrajo de la matriz ETUV obtenida por FIA. La Figura 3.20 expone una matriz de datos obtenida por FIA en la que se indica el perfil de dispersión generado por la muestra en la longitud de onda correspondiente al punto isobéptico.

Finalmente, teniendo en cuenta todas las consideraciones mencionadas y a fin de minimizar la ambigüedad de los resultados obtenidos, se procedió a la descomposición bilineal por MCR-ALS aplicando las siguientes restricciones durante la optimización ALS: **1-** no negatividad en modo espectral y temporal; y **2-** condición de balance de masa o condición de sistema cerrado. Esta última restricción se aplicó considerando la dispersión del bolo de la muestra generado en el sistema FIA, utilizando el perfil de dispersión de C_{DXR} obtenido en el punto isobéptico como valor de referencia.

El número de componentes espectroscópicamente activos y sus estimaciones espectrales iniciales se obtuvieron por SVD y una metodología basada en SIMPLISMA (Windig *et al.*, 1991), respectivamente.

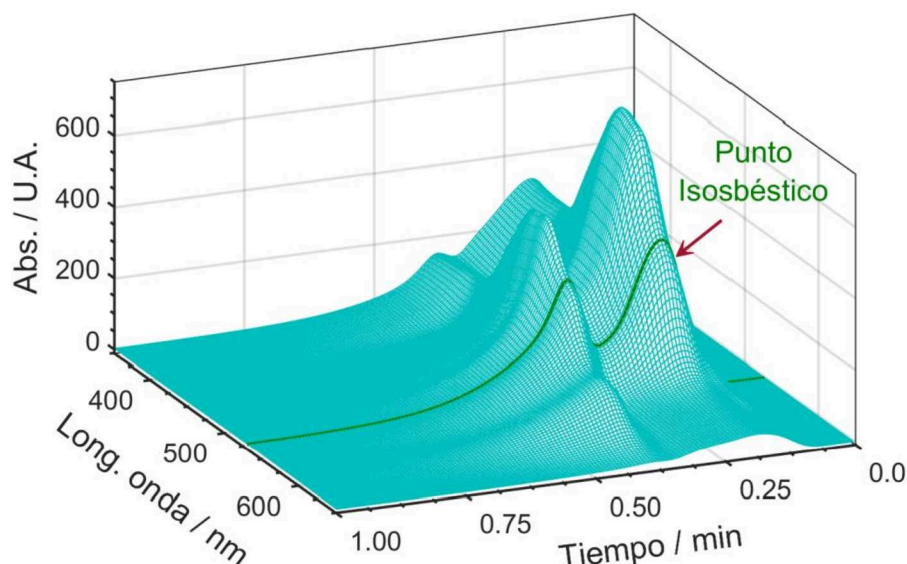


Figura 3.20. Matriz de datos obtenida por FIA para una solución de DXR de C_{DXR} de $7.0 \mu\text{g mL}^{-1}$ indicándose el perfil de dispersión del bolo de muestra generado en el punto isosbético $\lambda_{\text{pi}} = 530 \text{ nm}$

Obtenidos los resultados de la resolución quimiométrica con las correspondientes matrices **C** y **S** conteniendo los perfiles temporales y espectrales, respectivamente, se utilizó la información contenida en la matriz **C** a fin de establecer la relación entre el tiempo de elución de la muestra y los valores de pH del gradiente generado. Teniendo en cuenta la restricción de balance de masa y considerando el valor de $\text{pK}_{\text{a},1} = 8.84$, obtenido por PKFIT, se estimó el gradiente de pH generado en el sistema FIA utilizando la ecuación Henderson-Hasselbarlch para el cálculo de pH en un equilibrio ácido-base (Saurina *et al.*, 2000):

$$\text{pH} = \text{pK}_{\text{a}} + \log \frac{[\text{DXR}]}{[\text{DXR}^+]} = \text{pK}_{\text{a},1} - \log \frac{c_{\text{DXR}}}{c_{\text{DXR}^+}} \quad 3.3$$

donde $\text{pK}_{\text{a},1} = 8.84$ y c_{DXR^+} y c_{DXR} es el perfil temporal correspondiente a la especie DXR^+ y DXR, respectivamente, extraídos de la matriz **C**. Es de remarcar que, en virtud de las características del sistema FIA, se obtuvo un doble gradientes de pH: **1-** entre 0.00 min y 0.25 min; y **2-** entre 0.25 min y 1.00 min. La Figura 3.21 expone los resulta-

dos espectrales y temporales obtenidos de la resolución quimiométrica por MCR-ALS, mostrándose las curvas de distribución de especies obtenidas para los 2 gradientes de pH generados en el sistema FIA, calculados según la ecuación 3.3.

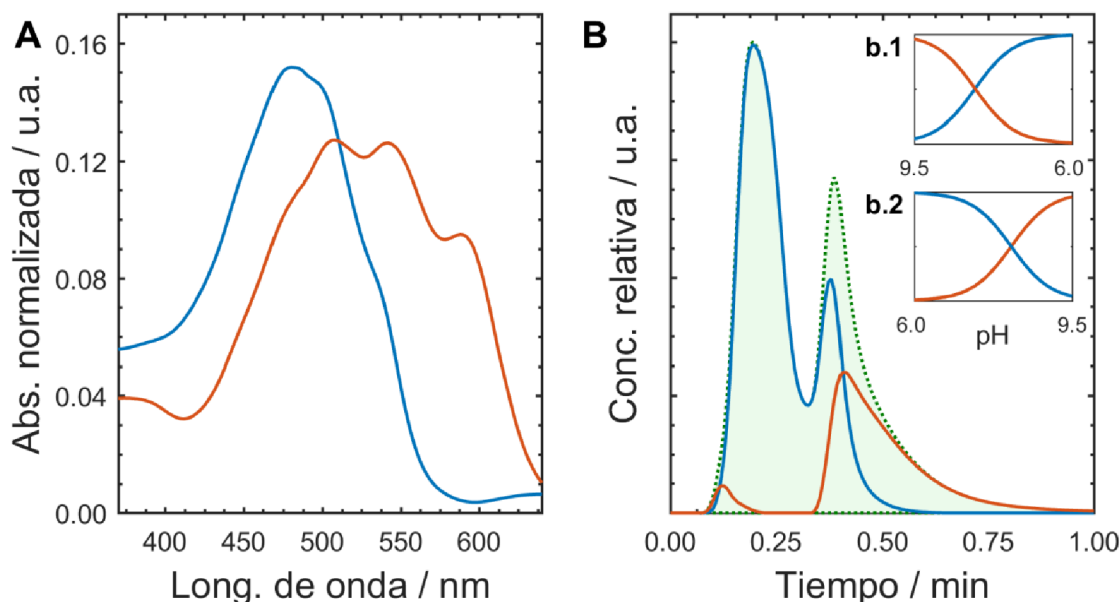


Figura 3.21. Perfiles espectrales (A) y temporales (B) obtenidos por MCR-ALS para una muestra de DXR de concentración $C_{\text{DXR}} = 7.0 \mu\text{g mL}^{-1}$, mostrándose la especie DXR^+ (azul), DXR (rojo) y el perfil de dispersión del bolo de muestra (línea de puntos verde) ante la restricción de condición de balance de masa. En los recuadros se muestran las curvas de distribución para la especie DXR^+ (azul) y DXR (rojo) obtenidas por el gradiente de pH generado en el sistema FIA para los rangos de tiempo comprendidos entre 0.00 min y 0.25 min (b.1) y entre 0.25 min y 1.00 min (b.2)

A fin de conocer la reproducibilidad de la generación del gradiente de pH, se evaluó el tiempo al que el valor de pH resultó igual al valor del $\text{pK}_{\text{a},1}$ para distintas réplicas de muestras de DXR de diferentes niveles de concentración. Los tiempos correspondientes al $\text{pK}_{\text{a},1}$ resultaron igual a 0.120 ± 0.003 min y 0.392 ± 0.001 min para el gradiente comprendido entre 0.00 min y 0.25 min y 0.25 min y 1.00 min, respectivamente. Estas cifras demuestran la alta reproducibilidad de la generación de un gradiente de pH en el sistema FIA, como así también la reproducibilidad de la resolución quimiométrica por MCR-ALS, utilizando las restricciones adecuadas.

Modelado quimiométrico por MCR-ALS y análisis predictivo. Los métodos basados en FIA se caracterizan por ser métodos que no incluyen la separación física de los componentes, por lo que fluyen juntos a través del sistema y llegan al detector al mismo momento. En consecuencia, para el análisis de muestras complejas o sistemas multi-componentes resulta necesario utilizar mecanismos alter-

nativos que aseguren y faciliten la resolución e identificación de cada compuesto. Una excelente alternativa para superar este inconveniente es la generación de datos de segundo orden, en donde la resolución de los componentes se logra en virtud de sus diferencias espectrales y temporales. En FIA, adicionalmente a las diferencias espectrales, la discriminación de los componentes se realiza a través de cambios en el perfil temporal obtenidos como resultado de la generación de un gradiente químico (Saurina, 2010). En este trabajo, para la obtención de datos de segundo orden, se procedió a la generación de un gradiente de pH, registrándose el espectro Vis a cada tiempo de elución. La determinación de DXR en muestras de plasma se llevó a cabo mediante modelado quimiométrico de arreglos de 2-vías aumentado por ext-MCR-ALS.

Para la descomposición bilineal, se construyeron matrices aumentadas en columna (\mathbf{D}_{aum}) que incluyeron las matrices de datos correspondientes a las muestras de validación o muestras de plasma con interferentes, según corresponda, y las matrices de datos de calibración. El número de especies espectralmente activas se determinó mediante SVD y las estimaciones iniciales espectrales se obtuvieron utilizando una metodología basada en SIMPLISMA (Windig *et al.*, 1991). En las muestras de validación el número de componentes obtenido mediante análisis por SVD fue igual a 3, correspondiéndose a cada especie de DXR presente en el equilibrio ácido-base, DXR^+ y DXR , y a la contribución del plasma. Por otro lado, en las muestras conteniendo DXR en presencia de RFP, el número de especies contribuyentes obtenido por SVD fue igual a 4, correspondiéndose, aquí, a las contribuciones de las especies de DXR presentes en el equilibrio, DXR^+ y DXR , al plasma y a RFP. Es importante destacar que si bien el valor de la constante de disociación de RFP se encuentra incluido en el rango de pH analizado, $\text{pK}_{a \text{ RFP}} = 7.9$ (ScienceDirect, 2008), y se esperarían contribuciones espectrales de las diferentes especies de RFP involucradas en el equilibrio, se observó que el espectro de RFP no depende del valor del pH, es decir, no se observaron diferencias espectrales para RFP en los distintos valores de pH estudiados. En consecuencia, solo un componente fue necesario para explicar la contribución de RFP en las muestras.

Resulta necesario mencionar que las muestras de calibración se construyeron adicionando los analitos a plasma humano debido, principalmente, a que las interacciones del plasma con los analitos modifican significativamente las propiedades de estos últimos, fundamentalmente, sus características espectrales.

Finalmente, la resolución por ALS para la matriz \mathbf{D}_{aum} se llevó a cabo aplicando solo la restricción de no negatividad. En la Figura 3.22 se muestran los perfiles temporales y espectrales obtenidos como resultado de la descomposición bilineal por MCR-

ALS para una de las muestras de plasma conteniendo DXR en presencia de RFP.

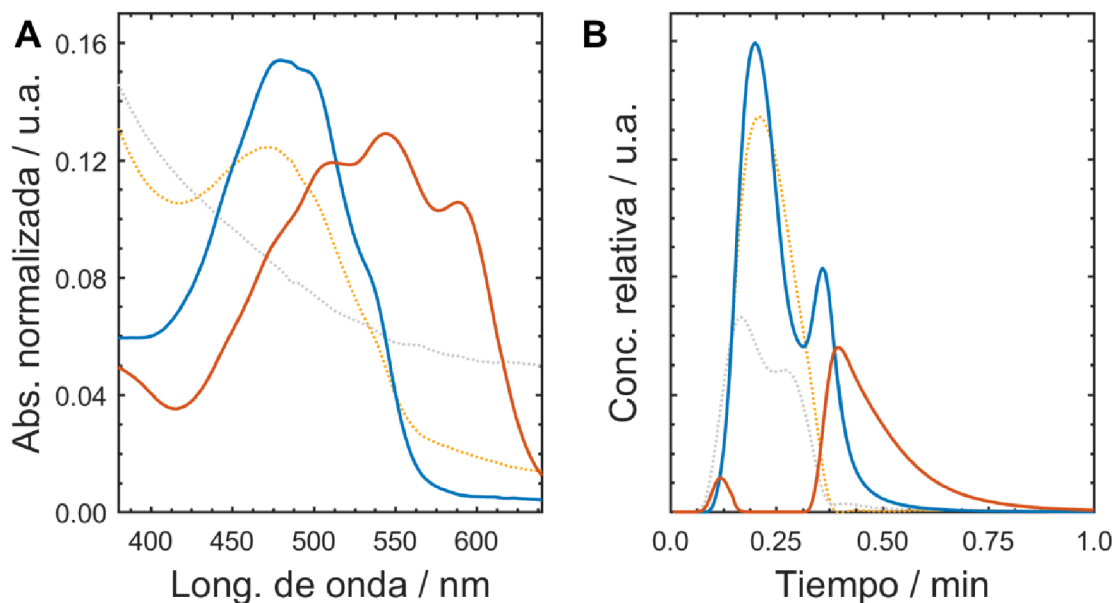


Figura 3.22. Perfiles espectrales (A) y temporales (B) obtenidos para una muestra de plasma conteniendo DXR en presencia de RFP, mostrándose las contribuciones de las especies involucradas en el equilibrio: DXR⁺ (azul), DXR (rojo), RFP (línea de puntos naranja) y plasma (línea de puntos gris)

Obtenidas las matrices **C** y **S** correspondientes al resultado de la descomposición bilineal, se utilizó la información contenida en **C** para la construcción de una curva de calibración tipo univariada, o pseudounivariada, a fin de predecir las concentraciones de los componentes presentes en muestras de validación y muestras de prueba. Además, para estimar las FOMs que representan al método analítico se siguió el procedimiento propuesto por Bauza *et al.* (Bauza *et al.*, 2012) para datos de segundo orden y modelado quimiométrico por MCR-ALS.

En casos donde la concentración del analito, por descomposición bilineal, se encuentra distribuida en dos factores o componentes ($C_{\text{DXR}} = [\text{DXR}^+] + [\text{DXR}]$ en este caso), es posible seguir 2 estrategias diferentes para la predicción de la concentración del compuesto y, por consiguiente, para la estimación de FOMs. La primera estrategia se basa en la selección del factor o componente que demuestre mejores parámetros analíticos, como ser mayor sensibilidad, mayor rango lineal y mejor ajuste de linealidad, mientras que la segunda utiliza la información obtenida para ambos factores, en otras palabras, se considera como respuesta analítica a la sumatoria de las contribuciones individuales de cada factor.

Es importante remarcar, además, que el cálculo de SEN para datos de segundo

orden requiere de la información espectral obtenida correspondiente al factor evaluado. Para la primera estrategia, esta información se puede extraer fácilmente de la matriz **S**, sin embargo, para la segunda estrategia, debido a que la información necesaria para el análisis es la correspondiente a ambos componentes, el cálculo de SEN se realiza utilizando la combinación lineal de los perfiles espectrales obtenidos para cada componente (Olivieri *et al.*, 2015). En la Figura 3.23 se representan las curvas de abundancia obtenidas por MCR-ALS respecto a la concentración establecida para las muestras de calibración para las especies DXR⁺ y DXR, como también, la curva correspondiente a C_{DXR} y las respectivas curvas de ajuste.

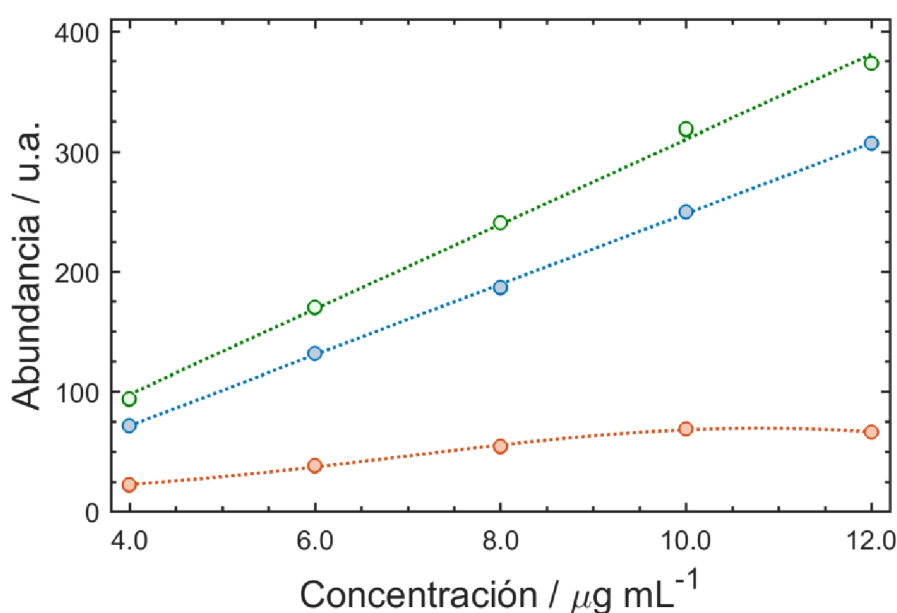


Figura 3.23. Curvas de abundancia con respecto a la concentración, obtenidas de la descomposición por MCR-ALS, para DXR⁺ (azul), DXR (rojo) y para la suma de contribuciones individuales (verde). En línea de puntos se muestran las curvas ajustadas de regresión

Puede observarse que el factor correspondiente a DXR⁺ presenta mayor rango lineal y sensibilidad en comparación con DXR. Además, en virtud de la simplicidad de cálculo y a que los resultados predictivos obtenidos no presentaron diferencias significativas entre las diferentes formas de cálculo, el análisis cuantitativo se llevó a cabo siguiendo la primera estrategia, es decir, utilizándose la información obtenida para el factor correspondiente a DXR⁺.

La Tabla 3.9 resume los valores de concentración obtenidos para las muestras de validación y las muestras de DXR en presencia de RFP. Para asegurar la exactitud del método desarrollado, se evaluaron las recuperaciones de manera estadística me-

diante un test de hipótesis t , donde la aceptación de la hipótesis nula H_0 (Olivieri, 2015) validó los valores de recuperación experimentales, demostrándose que el valor de recuperación media obtenido y el valor ideal 100 % presentó diferencias estadísticamente no significativas, para un nivel de confianza del 95 % ($\alpha = 0.05$) y $\nu = n-1$ grados de libertad ($\nu = 5$). Sobre la base de estos resultados, se aseguró la exactitud del método propuesto para la determinación DXR en plasma, aún en presencia de potenciales interferentes. Finalmente, los valores estimados para LOD y LOQ resultaron igual a $0.12 \mu\text{g mL}^{-1}$ y $0.38 \mu\text{g mL}^{-1}$ para DXR en plasma, respectivamente.

Tabla 3.9. Concentración de DXR predicha en muestras de validación y muestras de prueba en presencia de RFP ^a

Muestra	NOMINAL	PREDICHA^B	% RECUPERACIÓN
Val_1	6.0	6.4 (0.3)	106.7
Val_2	10.0	9.6 (0.7)	96.0
Val_3	8.0	7.9 (0.1)	98.8
Val_{RFP_4}^c	7.5	7.7 (0.1)	102.7
Val_{RFP_5}^c	6.0	6.3 (0.1)	105.0
Val_6	5.5	5.6 (0.1)	102.8
REP %^d		4.7	
\bar{R}_{exp}^e		101.8	

^a Las concentraciones están dadas en $\mu\text{g mL}^{-1}$;

^b Entre paréntesis se muestran las desviaciones estándares calculadas para 3 réplicas de preparación;

^c Muestras de prueba conteniendo DXR en plasma en presencia de $14.3 \mu\text{g mL}^{-1}$ y $12.5 \mu\text{g mL}^{-1}$ de RFP para Val_{RFP_4} y Val_{RFP_5}, respectivamente;

^d REP %: error relativo de predicción, expresado en %, $REP \% = 100 \times \frac{\sqrt{\frac{1}{l} \sum_1^l (c_{nom} - c_{pred})^2}}{c}$, para $l = 6$;

^e \bar{R}_{exp} , Valor de recuperación media, expresado en %.

Considerándose que para la terapia de pacientes masculinos adultos enfermos se utilizan dosis de entre 60 mg m^{-2} y 75 mg m^{-2} , lo que lleva a una concentración inicial de DXR de aproximadamente $20 \mu\text{g mL}^{-1}$, de la que solo un pequeño porcentaje se encuentra presente en plasma (Trevisan *et al.*, 2003), y con base en los valores obtenidos de LOD y LOQ, con el método propuesto es posible determinar concentraciones de DXR en plasma de pacientes tratados, en concentraciones inferiores a las utilizadas en el tratamiento, hasta un valor aproximado de 0.01 % de la concentración inicial.

La calidad y eficiencia del modelado quimiométrico se evaluaron cuantitativa-

mente calculando el grado de solapamiento espectral (s_{12}) entre los espectros de cada componente obtenido por PKFIT (\mathbf{s}_1) y los perfiles espectrales obtenidos por MCR-ALS (\mathbf{s}_2), siguiendo la expresión de la ecuación 3.1, obteniéndose valores de s_{12} iguales a 0.9997 y 0.9917 para la especie DXR⁺ y DXR, respectivamente. Estas cifras permitieron asegurar la eficiencia del modelado quimiométrico y la capacidad de obtener información espectral de alta calidad mediante modelado quimiométrico por MCR-ALS.

Posteriormente, se realizó un estudio de precisión evaluándose los parámetros de repetibilidad y precisión intermedia. El análisis de repetibilidad se llevó a cabo analizando la variabilidad estadística introducida en la determinación de DXR en plasma en tres niveles de concentración (inferior, medio y superior del rango de concentraciones utilizado en la calibración), por triplicado. Por otro lado, para el análisis de precisión intermedia, se evaluó la variabilidad estadística introducida en la determinación de DXR en plasma en tres niveles de concentración (inferior, medio y superior del rango de concentraciones utilizado en la calibración), por triplicado, en tres semanas diferentes. Así, se preparó y analizó un total de 27 muestras independientes de DXR, en ausencia de interferentes. En todos los casos, se determinó el CV % y se realizó un test de análisis de varianza, ANOVA, a fin de determinar la presencia de diferencias estadísticamente significativas entre muestras preparadas y analizadas en semanas diferentes. De los resultados obtenidos, se observa que todos los valores de CV % resultaron menores al valor recomendado por la Agencia Europea de Medicinas (EMA) para la determinación de drogas en muestras biológicas, que establece valores límites de CV % iguales a 15 % (EMA, 2011). Además, el test ANOVA arrojó valores de probabilidad p mayores al nivel de corte establecido $p = 0.05$, demostrándose que existen diferencias estadísticamente no significativas entre los análisis de predicción realizados en diferentes semanas. En la Tabla 3.10 se muestran los resultados obtenidos para el estudio de precisión, mediante el análisis de repetibilidad y precisión intermedia.

Adicionalmente, se evaluó la exactitud de la metodología analítica desarrollada mediante un gráfico de concentración predicha con respecto a concentración nominal y la estimación de la región elíptica de confianza conjunta. Esta última consiste en graficar la región de confianza mutua de la pendiente y la ordenada al origen para el gráfico de concentración predicha con respecto a concentración nominal, en el plano pendiente-ordenada al origen (Olivieri *et al.*, 2014b). Para este análisis se utilizaron las concentraciones predichas obtenidas en las muestras de validación, muestras de prueba y muestras de precisión, realizándose un ajuste de cuadrados mínimos ordinarios (OLS:

Ordinary Least-Squares). Como se puede observar en la Figura 3.24, el punto ideal 1 y 0 para pendiente y ordenada al origen, respectivamente, se encuentran incluidos dentro de la región de confianza, para un nivel del 95 %, si bien se observa una pequeña desviación de la idealidad en la pendiente de la curva ajustada de concentración predicha con respecto a concentración nominal.

Tabla 3.10. Resultados obtenidos para el análisis de repetibilidad y precisión intermedia ^a

<i>C_{DXR}</i>	<i>SEMANA 1</i>		<i>SEMANA 2</i>		<i>SEMANA 3</i>		<i>ANOVA</i> ^c
	PREDICHA	CV % ^b	PREDICHA	CV %	PREDICHA	CV %	
NOMINAL	4.2		4.3		4.4		
4.0	4.1	5.2	4.2	1.4	4.3	3.6	0.178
	3.8		4.3		4.1		
	8.3		8.4		8.4		
7.0	8.1	4.5	8.4	0.2	8.5	0.7	0.069
	7.6		8.4		8.5		
	13.4		12.3		13.8		
12.0	12.2	7.8	13.7	5.6	14.2	1.5	0.085
	11.5		13.4		13.9		

^a Las concentraciones están dadas en $\mu\text{g mL}^{-1}$;

^b Coeficiente de variación porcentual. $CV\% = (SD/\bar{c}) \times 100$, expresado en %, donde *SD* es el desvío estándar para $\nu = n-1$ grados de libertad y \bar{c} es la concentración media calculada;

^c Valor de probabilidad *p* obtenido del análisis de variancia, ANOVA, donde H_0 =todas las medias obtenidas en cada semanas son iguales para un nivel $\alpha = 0.025$ y $\nu = n-1$ grados de libertad.

Finalmente, con el objetivo de obtener una mayor comprensión sobre la exactitud del método, se analizaron las 2 muestras de prueba conteniendo DXR en presencia de RFP, Val_{RFP_4} y Val_{RFP_5}, por triplicado, mediante un método basado en HPLC acoplado a detección UV, siguiendo el procedimiento de Álvarez-Cedrón *et al.* (Álvarez-Cedrón *et al.*, 1999). En la Figura 3.25 se observa el perfil cromatográfico obtenido para la muestra Val_{RFP_4}.

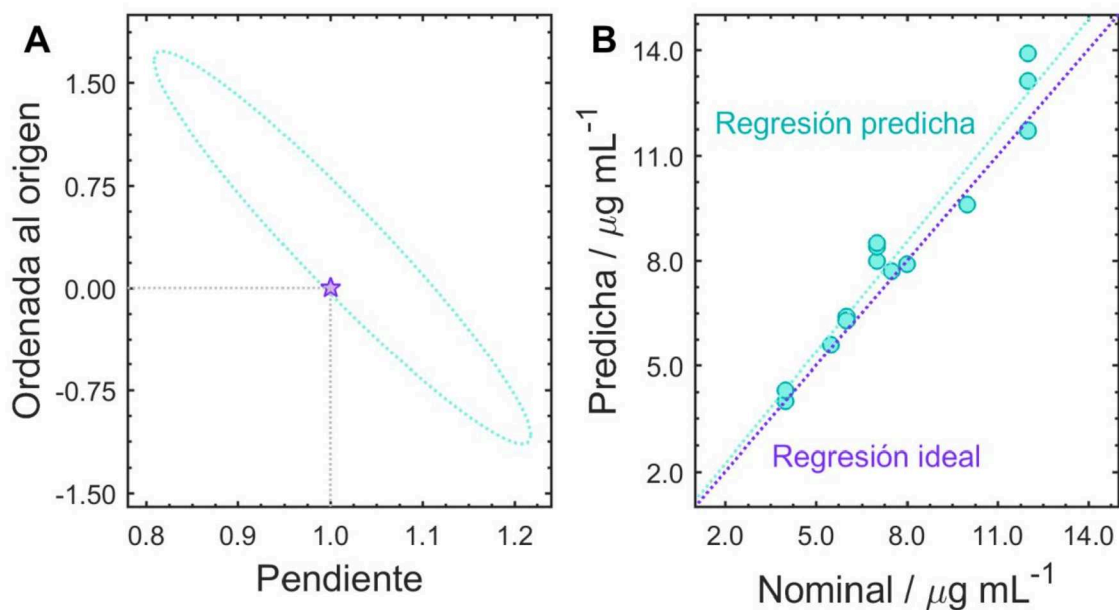


Figura 3.24. (A) Región elíptica de confianza conjunta en el plano pendiente-ordenada al origen, para un nivel de confianza del 95 %, indicándose el punto ideal [1, 0] (estrella violeta) para pendiente y ordenada al origen, respectivamente (B) Gráfico de concentración predicha con respecto a concentración nominal, mostrándose la línea ajustada para los valores experimentalmente (línea de puntos celeste) y la curva de regresión ideal (línea de puntos violeta)

Las concentraciones predichas obtenidas por la metodología analítica basada en HPLC-UV fueron igual a $7.5 \mu\text{g mL}^{-1}$ (SD = 0.1) y $6.08 \mu\text{g mL}^{-1}$ (SD = 0.04) para las muestras $\text{Val}_{\text{RFP}_4} = 7.5 \mu\text{g mL}^{-1}$ de DXR y $\text{Val}_{\text{RFP}_5} = 6.0 \mu\text{g mL}^{-1}$ de DXR.

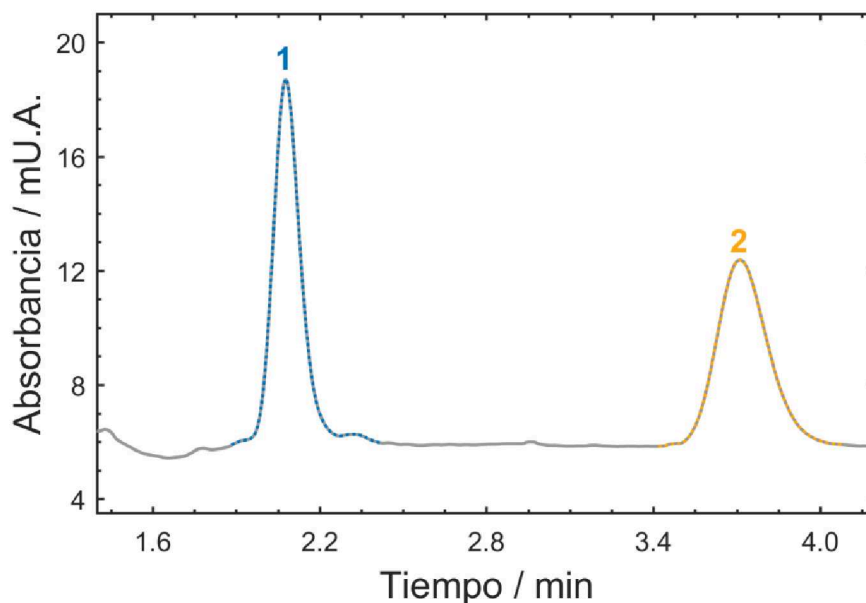


Figura 3.25. Cromatograma obtenido para la muestra $\text{Val}_{\text{RFP}_4}$, mostrándose las señales obtenidas para DXR (1, azul) y RFP (2, naranja), registrado a una longitud de onda de 254 nm

Un test estadístico t de comparación de medias demostró que existen diferencias

estadísticamente no significativas entre los valores de predicción y recuperación obtenidos utilizando el método basado en FIA y HPLC-UV.

CONCLUSIÓN

Se desarrolló un método para la determinación de DXR en plasma humano utilizando un sistema basado en la generación de un gradiente de pH por FIA acoplado a detección de arreglo de diodos Vis. La generación de datos de segundo orden demostró ser una excelente estrategia para lograr la resolución y cuantificación del compuesto de interés en muestras complejas o sistemas multi-componentes, aprovechándose de la ventaja de segundo orden que muestran las técnicas quimiométricas de calibración multivariada. Se seleccionó MCR-ALS como algoritmo para el análisis cuantitativo, permitiendo obtener, además, información cualitativa altamente confiable acerca del sistema en estudio.

Sobre la base de los resultados obtenidos, y considerándose los bajos límites de detección alcanzados, la metodología analítica propuesta resulta en una alternativa simple y conveniente para la determinación de DXR en plasma de pacientes bajo tratamiento. Se demostró, también, que con el método propuesto se alcanzan resultados comparables a los obtenidos por técnicas analíticas que involucran instrumental más sofisticado, como HPLC, en un tiempo de análisis notablemente menor. Visto desde esta perspectiva, este método podría ser utilizado, además, para la determinación de otros compuestos en muestras biológicas, extendiéndose a muestras complejas de diferente naturaleza.

MCR-ALS es una herramienta versátil que permitió realizar un análisis descriptivo minucioso del sistema, en referencia al gradiente de pH generado, sin requerir de mediciones experimentales de pH en línea. Esta metodología no solo permitió develar los valores de pH generados dentro del sistema, sino que resultaría en una excelente estrategia para el análisis y determinación de constantes de disociación de nuevos compuestos.

3. 4. DESARROLLO DE UNA MÉTODO PARA EL ANÁLISIS ESTRUCTURAL DE PROTEÍNAS EN MEDIO ACUOSO BASADO EN ESPECTROSCOPIA INFRARROJA LASER COMBINADA A HERRAMIENTAS QUIMIOMÉTRICAS PARA PROCESAMIENTO Y ANÁLISIS DE DATOS

External cavity-quantum cascade laser spectroscopy for mid-IR transmission measurements of proteins in aqueous solution – Alcaráz, M.R., et al., Anal Chem 2015, 87, 6980–6987

EC-QCL mid-IR transmission spectroscopy for monitoring dynamic changes of protein secondary structure in aqueous solution on the example of β -aggregation in alcohol-denatured α -chymotrypsin – Alcaráz, M.R., et al., Anal Bioanal Chem 2016, 408, 3933–3941

CONSIDERACIONES GENERALES

La espectroscopia infrarroja (*IR*) es una técnica versátil utilizada para el análisis cuantitativo y descriptivo de una amplia variedad de muestras. En la década del 60', el desarrollo de los algoritmos para el cálculo de series complejas de Fourier, o transformadas de Fourier (Cooley *et al.*, 1965), impulsó la creación de los interferómetros de Michelson-Morley utilizados en espectroscopia IR, que, hasta el día de hoy, están presentes en los espectrofotómetros IR acoplados a transformada de Fourier (*FT-IR: Fourier Transform-InfraRed*). Estos espectrofotómetros, generalmente, utilizan fuentes de luz termal, como por ejemplo, fuentes globares, que a pesar de emitir en un amplio rango de energía, presentan una potencia de radiación muy baja en la región del espectro IR, principalmente, del IR medio (*mid-IR*). La baja intensidad de emisión de este tipo de fuentes constituye una limitación para mediciones de muestras con matrices altamente absorbentes, como lo son las matrices acuosas. En estos casos, el agua, presente como solvente, absorbe la mayor parte de la intensidad de luz que queda disponible luego de la interacción con el compuesto de interés. Consecuentemente, los pasos ópticos utilizados para estas mediciones son restringidos a unos pocos μm , en virtud de la combinación de dos factores principales: **1-** *el alto coeficiente de absorción*

molar de una banda espectral a 1645 cm^{-1} correspondiente a las vibraciones por torsión del enlace HOH del agua; y 2- el bajo poder de emisión de las fuentes de luz termales usadas en los espectrofotómetros FT-IR. Los pasos ópticos generalmente utilizados para las mediciones de proteínas en solución acuosa se encuentran en un rango de aproximadamente $8\text{ }\mu\text{m}$ (Yang *et al.*, 2015), para evitar la absorción total de la energía IR en la región correspondiente a las vibraciones de torsión HOH. El uso de estos pasos ópticos para el análisis de proteínas en medio acuoso, pueden, por un lado, perjudicar la robustez instrumental del método, debido a obstrucciones que pudiesen ocurrir dentro de la celda, y además requieren de altas concentraciones de proteínas ($> 10\text{ mg mL}^{-1}$) (Fabian *et al.*, 2006).

Recientemente, se comenzó a comercializar una nueva fuente de energía lumínica IR, que incluye a los láseres de cascada cuántica con cavidad externa (*EC-QCL: External Cavity-Quantum Cascade Laser*), que operan a temperatura ambiente y presentan una alta densidad de potencia de emisión en un amplio rango espectral. Los QCL son láseres unipolares basados en transiciones inter sub-bandas de electrones dentro de varias bandas de un conductor de semi-conducción (Curl *et al.*, 2010; Yao *et al.*, 2012). En palabras simples, el principio de funcionamiento de un QCL se puede explicar de la siguiente manera: *un electrón que permanece en una banda de conducción luego de emitir un fotón laser, es reciclado por inyección, por aplicación de una diferencia de potencial adecuada, a una región activa adyacente idéntica a la anterior, desde donde emite otro fotón. Esto se repite sucesivamente hasta alcanzar la última banda o región activa disponible* (Curl *et al.*, 2010). Este efecto de cascada cuántica es el responsable del alto poder que pueden alcanzar las bandas de emisión de los QCL. La Figura 3.26 resume esquemáticamente el funcionamiento de este tipo de láseres.

Por otro lado, la cavidad externa de los EC-QCL permite la selección de diferentes números de onda mediante una red de difracción externa, como también, el registro en un amplio rango espectral. El alto poder de emisión de este tipo de fuente permite incrementar la longitud de los pasos ópticos para mediciones de transmisión, incluso en presencia de matrices fuertemente absorbentes. Recientemente se han reportado métodos analíticos utilizando EC-QCL como fuente de luz IR para el análisis de muestras complejas en soluciones acuosas para el monitoreo de procesos online (Siegmann-Hegerfeld *et al.*, 2013) y aplicaciones médicas (Brandstetter *et al.*, 2010; Brandstetter *et al.*, 2012; Brandstetter *et al.*, 2013b).

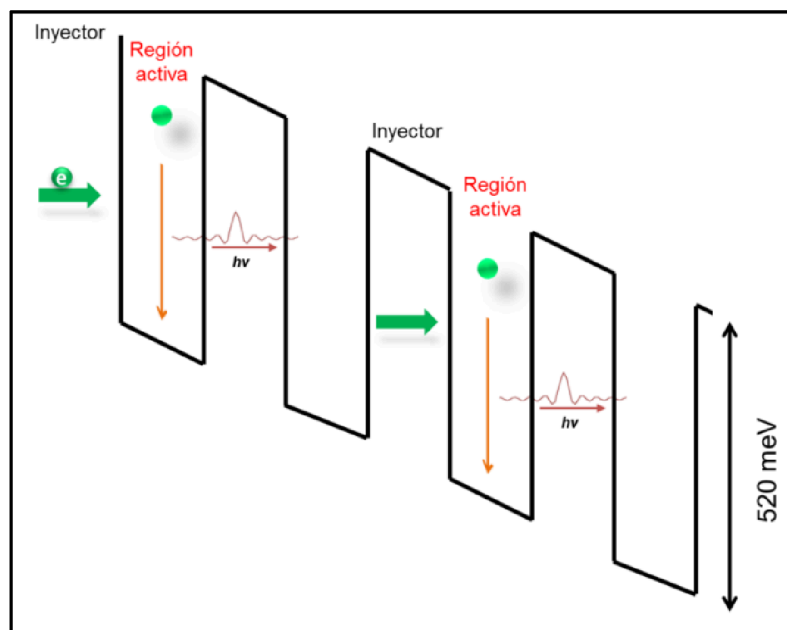


Figura 3.26. Representación esquemática del funcionamiento de un láser de cascada cuántica, QCL

La espectroscopia IR es un método analítico bien establecido para el análisis estructural de proteínas (Barth, 2007; Yang *et al.*, 2015) y el monitoreo de cambios dinámicos en sus estructuras secundarias. Las bandas de absorción más prominentes de las proteínas en el espectro IR son causadas por vibraciones del grupo peptídico y son referidas como bandas amida. De las bandas principales que se observan en la región IR, la banda amida I ($1600-1700\text{ cm}^{-1}$), originada principalmente por las vibraciones por estiramiento de los grupos carboxilos (C=O), es la utilizada con mayor frecuencia para el análisis de estructuras secundarias (Fabian *et al.*, 2006). Los diferentes patrones de los puentes de hidrógeno, las interacciones dipolo-dipolo, las orientaciones geométricas en hélices α , láminas y giros β o estructuras de configuración aleatoria inducen diferentes frecuencias en las vibraciones C=O que pueden ser correlacionadas con la respectivas estructuras secundarias (Bal Ram, 1999). Dentro de las técnicas utilizadas para mediciones IR de proteínas, las basadas en mediciones de transmisión son utilizadas, habitualmente, para muestras en solución, mientras que los métodos de reflectancia total atenuada (*ATR: Attenuated Total Reflectance*) son empleados para estudios de absorción o investigaciones en láminas delgadas.

Las alteraciones en la estructura secundaria de las proteínas pueden ser inducidas por cambios en las condiciones externas, tales como pH del medio, temperatura, presión, presencia de co-solventes, surfactantes o agentes crioscópicos, y son generalmente acompañadas de la desnaturalización de las proteínas. El cambio más pro-

minente de la estructura secundaria es la transición desde hélices α hacia láminas β , resultado de un fenómeno de agregación. Si bien los procesos de desnaturalización están acompañados de disrupciones de hélices α y formación de láminas β , giros β y hélices tipo poliprolina II, la exposición a alcohol, en particular, induce y estabiliza dominios de hélices α (Matsuo *et al.*, 2012). Pese a que el efecto de los alcoholes, especialmente los fluor-sustituídos, por ejemplo, 2,2,2-trifluoroetanol (TFE), sobre las proteínas ha sido ampliamente estudiado, el mecanismo físico por el que el alcohol afecta a la conformación de las proteínas es aún incierto.

La predisposición de proteínas con estructura secundaria predominantemente de láminas β , como la α -quimiotripsina (α -CT), hacia la agregación en fibrillas amiloideas ante la presencia de TFE ha sido extensamente demostrada por análisis turbidimétrico, de luz dinámica dispersada, termodinámico, de fluorescencia intrínseca, entre otros (Khodarahmi *et al.*, 2009; Pallarès *et al.*, 2004; Rezaei-Ghaleh *et al.*, 2007). Estos trabajos reportan que las interacciones TFE-proteína y proteína-proteína dependen fuertemente de condiciones externas, tales como el pH, la concentración de TFE y la temperatura. La formación de fibrillas amiloideas es promovida por concentraciones bajas o intermedias de TFE (> 35 %), pero a altas concentraciones (\geq 50 %) es suprimida debido a la pronunciada estabilización de estructuras helicoidales α no-nativas (Harrison *et al.*, 2007). Se ha demostrado, además, que la exposición de α -CT a altas concentraciones de TFE lleva consigo la instantánea formación de estructuras de hélices α no nativas, sucedida por la formación gradual de agregados de láminas β intermoleculares (Dong *et al.*, 1998).

En este trabajo se propone el desarrollo de una nueva metodología IR utilizando un láser EC-QCL como fuente lumínica para mediciones de transmisión mid-IR de proteínas en medio acuoso. Para el tratamiento de señales obtenidas, se propone una estrategia de procesamiento basado en algoritmos quimiométricos, a fin de disminuir el ruido instrumental introducido por el EC-QCL. Adicionalmente, se estudió el efecto de TFE sobre la proteína α -CT, monitoreándose los cambios dinámicos en la estructura secundaria de la proteína ante variaciones de pH y concentración proteica, registrándose espectros mid-IR en un determinado periodo de tiempo. Con el objetivo de obtener resultados con interpretación física y poder inferir acerca de las conformaciones estructurales involucradas en el proceso, se seleccionó MCR-ALS como algoritmo quimiométrico para el modelado de los datos multidimensionales obtenidos.

MATERIALES Y MÉTODOS

Reactivos. El fosfato de sodio monobásico dihidrato calidad p.a. ($\text{NaH}_2\text{PO}_4 \cdot 2\text{H}_2\text{O}$) fue adquirido de Fluka (Buchs, Suiza). El fosfato de sodio dibásico dihidrato ($\text{Na}_2\text{HPO}_4 \cdot 2\text{H}_2\text{O}$) BioUltra calidad biología molecular, la solución de hidróxido de sodio 50 % en agua (NaOH), el ácido clorhídrico 37 % (HCl) calidad ACS y 2,2,2-trifluoroetanol (TFE) calidad ReagentPlus ≥ 99 % fueron obtenidos de Sigma-Aldrich (Steinheim, Alemania). Las proteínas liofilizadas β -Lactoglobulina de leche bovina (β -LG, ≥ 85 %), hemoglobina de sangre bovina (HEMO), albúmina de suero bovino (BSA , ≥ 98 %), lisozima de clara de huevo (Lys , > 90 %) y α -quimotripsina de páncreas bovino (α -CT, ≥ 85 %) fueron adquiridas de Sigma-Aldrich (Steinheim, Alemania) y utilizadas sin tratamiento previo. El agua de laboratorio ultrapura fue obtenida mediante un sistema de purificación de agua Milli-Q MilliPore (Billerica, Estados Unidos).

Soluciones de trabajo. Se preparó una solución reguladora de fosfatos (PBS , $\text{NaH}_2\text{PO}_4/\text{Na}_2\text{HPO}_4$) disolviendo cantidades apropiadas de las sales correspondientes y ajustando el pH a 7.0 con HCl o NaOH , según corresponda, obteniéndose una concentración formal de 0.016 mol L^{-1} . Las soluciones madres de proteínas se prepararon diariamente por disolución de la cantidad apropiada del liofilizado proteico correspondiente en PBS . Se analizaron diferentes proteínas por técnicas de FT-IR y EC-QCL-IR en distintos niveles de concentración comprendidos entre 1.0 mg mL^{-1} y 20.0 mg mL^{-1} .

Para el análisis de desnaturalización se prepararon diferentes soluciones de α -CT disolviendo la cantidad apropiada de proteína liofilizada directamente en 1.0 mL de una solución mezcla de $\text{TFE}:\text{PBS}$ 50:50 (TFEPBS). Inmediatamente después, la solución resultante se transfirió a la celda de flujo del espectrofotómetro EC-QCL-IR para el subsecuente análisis. El pH de la solución PBS fue ajustado con NaOH o HCl , según corresponda previo a la preparación de la solución mezcla TFEPBS . Se preparó un set de 7 muestras conteniendo 20.0 mg mL^{-1} de α -CT a diferentes valores de pH en el rango de 5.8 y 8.2 y un set de 7 muestras a pH 6.6 conteniendo diferentes niveles de concentración de α -CT, en el rango de entre 5.0 mg mL^{-1} y 60.0 mg mL^{-1} .

Instrumentación y procedimiento. Las mediciones de pH se realizaron utilizando un potenciómetro pH 330i 410A (Wissenschaftlich-Technische Wer-

kstätten GmbH, Weilheim, Alemania) equipado con un electrodo de vidrio combinado Sentix® 61 (Wissenschaftlich-Technische Werkstätten GmbH, Weilheim, Alemania) acoplado a un sensor de temperatura.

Mediciones FT-IR. Las mediciones espectroscópicas de absorción infrarroja acoplada a transformada de Fourier (*FT-IR*) se realizaron utilizando un espectrofotómetro FT-IR Bruker 80v (Ettlingen, Alemania), equipado con un detector MCT (HgCdTe, MCT: Mercury-Cadmium-Telluride) operado con nitrógeno líquido (LN_2) ($D^* = 4 \times 10^{10} \text{ cm Hz}^{0.5} \text{ W}^{-1}$ a $9.2 \mu\text{m}$). Se utilizó una celda de flujo compuesta por dos ventanas de CaF_2 y un separador de politetrafluoroetileno (*PTFE*) de $8 \mu\text{m}$ de espesor. El espectrofotómetro se purgó con aire seco para disminuir el contenido de vapor de agua atmosférico. Los espectros se registraron en modo de adquisición de datos en doble canal en la región espectral comprendida entre 400.0 cm^{-1} y 4000.0 cm^{-1} con una resolución espectral de 2 cm^{-1} , a temperatura constante de $25 \text{ }^\circ\text{C}$. Para cada espectro se promediaron 32 mediciones consecutivas (tiempo de adquisición: 6.7 s) y para la conversión a transmitancia se utilizó una función de apodización Blackman-Harris de tercer término, con un factor de sensibilidad cero igual a 2. En caso de ser necesario, se substrajeron las bandas de absorción del vapor de agua atmosférico. El análisis y la adquisición de datos se realizaron utilizando el programa informático OPUS 7.2 (Bruker, Ettlingen, Alemania).

Mediciones EC-QCL-IR. Para las mediciones espectroscópicas de absorción laser infrarroja (*EC-QCL-IR*) se utilizó un dispositivo ensamblado en el laboratorio comprendido por un láser de cascada cuántica con cavidad externa (*EC-QCL*, Daylight Solutions Inc., San Diego, Estados Unidos) termoeléctricamente enfriado, operando a una velocidad de repetición de 100 kHz y un ancho de pulso de 500 ns.

Los espectros de transmisión IR se registraron en el rango espectral localizado entre 1729.30 cm^{-1} y 1565.06 cm^{-1} , cubriendo la región de la banda amida I de las proteínas. Se utilizó un espejo parabólico recubierto en oro para concentrar el haz de luz IR transmitida por la muestra al detector. Para la detección se utilizó un detector MCT operado a $-60 \text{ }^\circ\text{C}$ por enfriamiento termoeléctrico ($D^* = 4 \times 10^9 \text{ cm Hz}^{0.5} \text{ W}^{-1}$ a $9.2 \mu\text{m}$). El procesamiento y digitalización de la señal detectada se logró utilizando un integrador de dos canales y un conversor analógico-a-digital (*ADC: Analog-to-Digital Converter*) NI DAQ 9239 de 24 bits (National Instruments Corp., Austin, EE.UU). Cada espectro, comprendido por 24000 datos, se registró en un tiempo igual a 1.5 s. Para el

control del dispositivo se utilizó un controlador desarrollado en el laboratorio soportado en LabView GUI 11.0 (National Instruments Corp., Austin, EE.UU) con estructura servidor-cliente. El dispositivo utilizado se muestra esquemáticamente en la Figura 3.27.

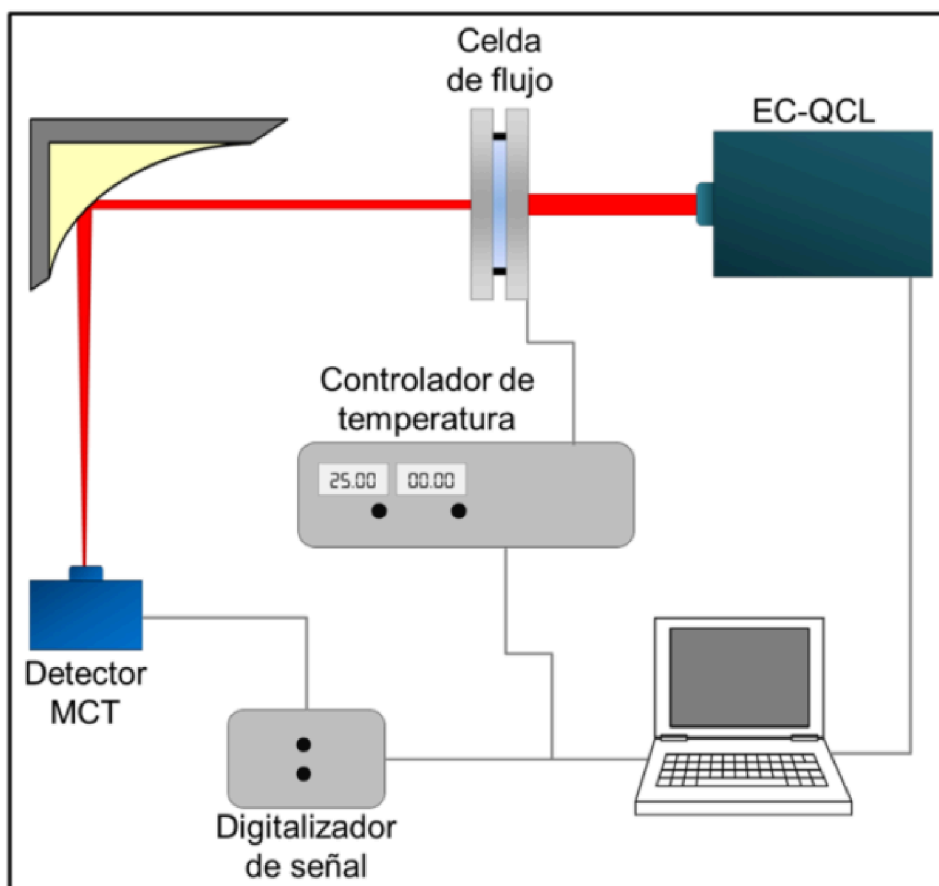


Figura 3.27. Representación esquemática del dispositivo utilizado para las mediciones EC-QCL-IR

Las mediciones EC-QCL-IR se llevaron a cabo a 25 °C utilizando una celda de flujo construida en el laboratorio con control de temperatura, equipada con dos ventanas de CaF₂ transparentes a la luz mid-IR y un separador de PTFE de 38 μm de espesor. Para reducir la influencia del vapor de agua, el instrumento se aisló en una caja de polietileno con flujo de aire seco constante. Para cada espectro, se promediaron 50 mediciones consecutivas de transmisión mid-IR (tiempo de adquisición: 100 s). La resolución espectral se calculó evaluándose el ancho de banda del espectro de vapor de agua y comparándose con los espectros adquiridos por FT-IR, utilizando distintos valores de resolución. Se obtuvo en una resolución para EC-QCL-IR igual a 0.2 cm⁻¹.

Para el análisis de desnaturalización de α-CT en presencia de TFE se obtuvieron matrices de datos tiempo-espectro de absorbancia IR (*TIR: Time-InfraRed spectra*)

registrados en el rango espectral comprendido entre 1710.0 cm^{-1} y 1585.0 cm^{-1} , en un periodo de tiempo de entre 2 min y 240 min. De esta manera, las matrices TIR consistieron en 38×4600 datos en la dimensión temporal y espectral, respectivamente, que luego se analizaron por MCR-ALS.

Programas informáticos y algoritmos para el análisis de datos. Para el procesamiento de las señales obtenidas por EC-QCL-IR se desarrolló un procedimiento basado en un algoritmo de corrección de deformaciones (*COW: Correlation Optimized Warping* (Nielsen *et al.*, 1998)), que se encuentra asequible en internet accediendo a la página web http://www.models.life.ku.dk/DTW_COW. Los algoritmos necesarios para operar MCR-ALS se encuentran asequibles en internet accediendo a la página web <http://www.mcrals.info/>. El procesamiento de datos y el análisis químico con MCR-ALS se realizó en MatLab 7.10 (MathWorks Inc., Natick, Estados Unidos).

RESULTADOS

Procesamiento de señales obtenidas por EC-QCL-IR. En espectroscopia EC-QCL-IR, el barrido espectral a través del rango de número de onda se logra por la rotación cuasi-continua de la red de difracción ubicada en la cavidad externa de la fuente lumínica EC-QCL. Debido a pequeños saltos ocurridos en la cascada cuántica y a la rotación de la red de difracción de la cavidad externa, las curvas de emisión resultantes presentan, en su distribución general, una estructura fina característica, comúnmente denominadas *mode-hops* (Wysocki *et al.*, 2005). Por otro lado, las imperfecciones mecánicas del láser conllevan a una desalineación de las estructuras finas entre mediciones consecutivas, como se puede observar en la Figura 3.28. Estas diferencias, si bien parecen despreciables en las curvas de transmisión, conllevan a la introducción de un ruido considerablemente alto en los espectros de absorbancia correspondientes (Brandstetter *et al.*, 2013a). Estas fluctuaciones no presentan un patrón aleatorio, por lo que promediar un número suficientemente elevado de mediciones consecutivas no resulta en una opción eficiente para la reducción de las perturbaciones resultantes. Además, debido a que el corrimiento o deformación de los *mode-hops* no es constante entre mediciones consecutivas, una simple corrección por desplazamiento

horizontal, es decir, a través del eje de número de onda, tampoco se presenta como una estrategia efectiva para la reducción de las perturbaciones en el espectro de absorbancia, si bien en la literatura se pueden encontrar diferentes trabajos utilizando esta metodología (Brandstetter *et al.*, 2013a). Asimismo, el uso de técnicas basadas en filtrado y/o suavizado de señales no es recomendable debido al alto ruido presente en el espectro de absorbancia, pudiéndose introducir distorsiones en la forma general del espectro de absorbancia final, llevando consigo una malinterpretación de la información contenida en el espectro IR evaluado.

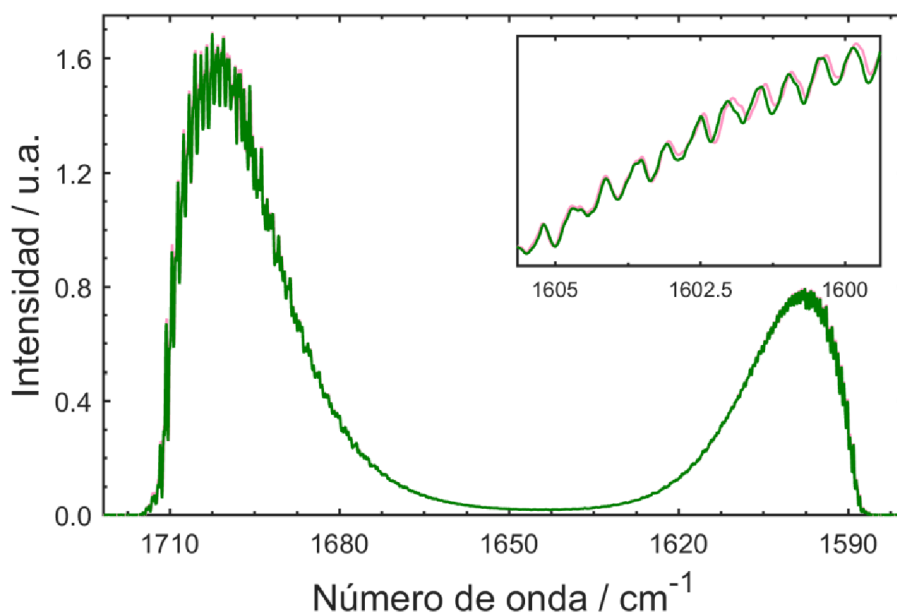


Figura 3.28. Curvas de transmisión consecutivas obtenidas por EC-QCL-IR para agua ultrapura. En el recuadro se amplifica la región espectral comprendida entre 1605 cm^{-1} y 1600 cm^{-1} , mostrándose las diferencias entre las fluctuaciones presentes de las estructuras finas de 2 líneas de transmisión consecutivas

En las fuentes EC-QCL, el recorrido completo de la red de difracción en movimiento continuo es referido como el registro espectral completo en un rango definido de número de onda (Brandstetter *et al.*, 2012). Las imperfecciones del láser y las inexactitudes ocurridas en la adquisición de los datos, debido a retardos en la red de difracción y en la detección de la señal, introducen desviaciones en la escala de número de onda, haciendo necesaria la calibración de instrumento.

Por otro lado, se observó que el tipo de ADC utilizado en la adquisición de los datos afecta notablemente al espectro de absorbancia resultante debido a los diferentes retardos generados por las inexactitudes del pulso de detección y el procesamiento de digitalización interno del ADC. Como se dijo, el eje temporal se corresponde direc-

tamente al eje de número de onda, por lo tanto, cualquier retardo introducido en la adquisición de datos conlleva imprecisiones en las posiciones espectrales.

Para corregir las desviaciones ocasionadas por los retardos en la adquisición, se propuso una estrategia de calibración del instrumento, utilizándose las líneas de absorción del vapor de agua. Se utilizaron los datos tabulados para las bandas de vapor de agua que se encuentran en la base de datos HITRAN (Rothman *et al.*, 2013) como referencia y se compararon con las bandas de vapor de agua obtenidas por EC-QCL-IR, obteniéndose el factor de corrección o factor de calibración.

A fin de superar los inconvenientes mencionados y obtener espectros de alta calidad que brinden información equivalente a la obtenida por técnicas infrarrojas de referencia, como FT-IR, se desarrolló una nueva estrategia quimiométrica para el procesamiento de señales EC-QCL-IR:

La primera etapa se basa en reducir el número de datos correspondientes a cada curva de transmisión obtenidas usando interpolación cúbica y una herramienta de suavizado Savitzky-Golay, con el objetivo de reducir el ruido instrumental sin alterar el patrón característico de las estructuras finas o mode-hops.

*Seguidamente, se utilizó el algoritmo COW para la corrección de las diferencias entre mediciones consecutivas. La utilización de este algoritmo es la clave elemental del procedimiento que permite corregir la desalineación de las fluctuaciones entre mediciones consecutivas antes de promediarlas (COW 1), como así también, la desalineación presente entre las curvas de transmisión obtenidas para la solución de referencia (I_0) y la muestra (I) (COW 2), antes del cálculo de absorbancia. Se utilizan las estructuras finas inherentes del EC-QCL, mode-hops, corrigiéndose las imperfecciones por alineación horizontal a través del eje de número de onda y por efecto elástico de las fluctuaciones. Para llevar a cabo la corrección se requiere, principalmente, de 2 parámetros: **1-** longitud de segmento, r ; y **2-** grado de elasticidad, l . Aquí, r se mantuvo constante en un valor de 20, correspondiéndose a un rango espectral de 0.45 cm^{-1} y considerándose la longitud media de los mode-hops, mientras que l se seleccionó y optimizó dentro de un rango de 2 a 5 puntos para cada muestra individualmente.*

El filtrado de las curvas de transmisión I_0 e I se realizó utilizando un procedimiento basado en transformada rápida de Fourier (FFT: Fast Fourier Transform) de paso corto con apodización Blackman-Harris de cuarto

término. Cabe mencionar que se evaluaron diferentes metodologías de suavizado y filtrado, por ejemplo, método de suavizado Savitzky-Golay, método de suavizado de media móvil, filtrado utilizando herramientas de wavelet y FFT, siendo esta última la seleccionada para el procedimiento final.

Por último, el espectro de absorbancia IR se obtuvo calculando la absorbancia correspondiente a cada número de onda siguiendo la expresión:

$$A_{\nu_i} = -\log \frac{I_{\nu_i}}{I_{0,\nu_i}} \quad 3.4$$

donde A_{ν_i} es la absorbancia calculada a i número de onda, I_{ν_i} e I_{0,ν_i} son las intensidades de transmisión obtenida a i número de onda para la muestra y la referencia, respectivamente.

En la Figura 3.29 se sintetiza la metodología desarrollada en este trabajo para el procesamiento de las señales obtenidas por EC-QCL-IR. La Figura 3.30 muestra una comparación entre los espectros de absorbancia de la proteína α -CT calculados sin procesamiento, con corrección por alineación horizontal según se reporta en la literatura y obtenido utilizando la metodología de procesamiento desarrollada en este trabajo, basada en el empleo del algoritmo COW.

Puede apreciarse claramente que la no corrección de las señales introduce un nivel de ruido considerable en el espectro de absorbancia. Asimismo, se puede observar que la corrección por alineación horizontal no resulta suficientemente eficiente para la reducción de ruido y que la estrategia de procesamiento aquí desarrollada es una valiosa herramienta a utilizar para obtener señales EC-QCL-IR con niveles de ruido altamente disminuidos. Queda así demostrado que el adecuado uso de una estrategia de procesamiento de señales es clave para la obtención de espectros de alta calidad.

Finalmente, el uso de técnicas de filtrado y/o suavizado pudieron ser utilizadas posterior al procedimiento desarrollado, sin riesgo de introducir deformaciones en los espectros de absorbancia resultantes.

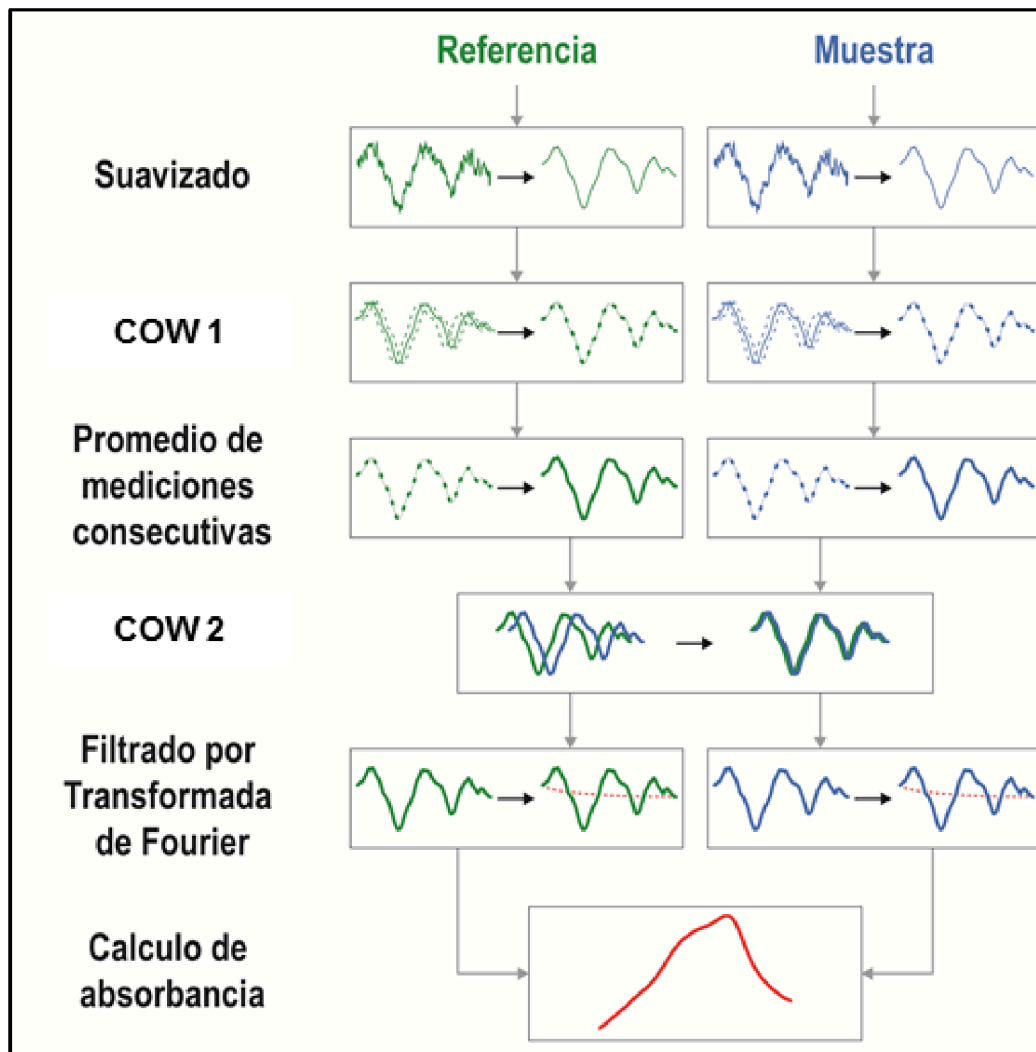


Figura 3.29. Metodología de análisis desarrollada para el tratamiento de señales obtenidas por EC-QCL-IR

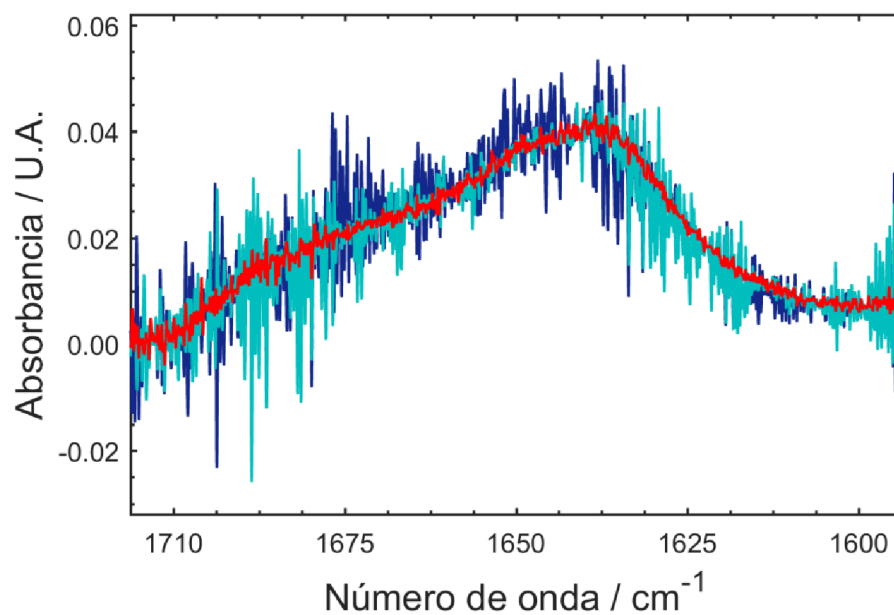


Figura 3.30. Espectros de absorbancia IR obtenidos por EC-QCL-IR sin procesamiento (azul), con corrección por alineación horizontal (celeste) y utilizando la metodología de procesamiento basada en COW (rojo) para una solución de 5.0 mg mL^{-1} de α -CT en PBS pH 7.0

Análisis comparativo entre FT-IR y EC-QCL-IR. Al momento de comparar dos instrumentos de similares características, además de los parámetros instrumentales, es importante tener en cuenta el nivel de ruido instrumental introducido. Una estrategia simple para cuantificar el ruido generado por los espectrofotómetros IR es calculando el error cuadrado medio (*RMS: Root Mean Square*) de las líneas 100 % de agua, registradas en idénticas condiciones. Estas líneas se obtienen calculándose la absorbancia resultante de dos espectros de transmisión subsecuentes de una misma solución. En situaciones ideales, el resultado se corresponde a una línea plana de absorbancia 0 en todo el rango espectral evaluado (Brandstetter *et al.*, 2012). Sin embargo, en condiciones reales, estas líneas no son planas y presentan intensidades que oscilan en valores cercanos a cero, siendo el valor RMS un dato representativo para su evaluación.

El valor RMS de las líneas 100 % obtenidas utilizando celdas con pasos ópticos de 38 μm y 8 μm para EC-QCL-IR y FT-IR, respectivamente, fueron 0.05 mA.U. para FT-IR y 0.37 mA.U. para EC-QCL-IR. Considerándose que el paso óptico para EC-QCL-IR es 4 veces mayor al utilizado para FT-IR, es de esperarse que el valor de absorbancia de una misma solución muestra se aproxime a una relación similar. Sin embargo, la SNR obtenida para los 2 espectrofotómetros resultó en el mismo orden de magnitud para una misma solución proteica conteniendo 20.0 mg mL^{-1} de α -CT: $SNR_{FT-IR} \sim 820$ y $SNR_{EC-QCL-IR} \sim 550$. En consecuencia, para evaluar correctamente el nivel de ruido instrumental es necesario considerar todos los factores involucrados en la medición.

Por un lado, el nivel de ruido introducido por las fuentes termales de baja intensidad de los espectrofotómetros FT-IR es, generalmente, considerada despreciable. En cambio, las fuentes laser de alta intensidad contribuyen significativamente al ruido global de las mediciones, particularmente cuando son operados en modo de pulso. Por otro lado, es importante tener en cuenta que la capacidad de detección específica (D^*) de los detectores MCT operados con LN_2 (~ -200 °C), utilizados en los espectrofotómetros FT-IR, es aproximadamente 10 veces mayor a la del detector MCT operado termoeléctricamente (-60 °C), utilizado en el espectrofotómetro EC-QCL-IR. Finalmente, un punto importante a ser considerado es el paso óptico utilizado para las mediciones. Cuando suficiente energía lumínica atraviesa la muestra y llega al detector, principalmente al usar pasos ópticos cortos, el nivel de ruido instrumental se mantiene prácticamente invariable. Sin embargo, al aumentar el paso óptico por encima de un valor al que prácticamente toda la intensidad de luz emitida por la fuente es absorbida

por la muestra, el ruido en las mediciones comienza a aumentar desproporcionadamente, por ejemplo, el valor de RMS_{FT-IR} obtenido para agua utilizando una celda de 38 μm resultó igual a 99.94 mA.U.. En consecuencia, el aumento en el paso óptico resulta exclusivo para los diseños instrumentales basados en EC-QCL-IR que, si bien se obtienen valores de SNR comparables a los espectrofotómetros FT-IR, lleva consigo un importante incremento de la robustez instrumental de los sistemas de análisis.

Para demostrar la calidad de los espectros IR obtenidos con EC-QCL-IR se registraron espectros por FT-IR, que se utilizaron luego como referencia, debido a que esta última es una técnica bien establecida para el análisis estructural de proteínas. Para este estudio, se seleccionaron adecuadamente 5 proteínas que presentan diferencias en su estructura secundaria y tienen espectros IR bien estudiados: **1-, 2- α -CT** y **β -LG**: proteínas compuestas mayoritariamente por estructuras de láminas β ; **3-, 4- BSA** y **HEMO**: proteínas compuestas predominantemente por estructuras de hélices α (Barth, 2007; Murayama *et al.*, 2004); y **5- Lys**: proteína que contiene hélices α y láminas β en iguales proporciones (Dong *et al.*, 2000; Oshokoya *et al.*, 2014).

En la Figura 3.31 se muestran los espectros de las 5 proteínas seleccionadas obtenidos por EC-QCL-IR, utilizando una celda de 38 μm de paso óptico, y por FT-IR, usando una celda de 8 μm de paso óptico.

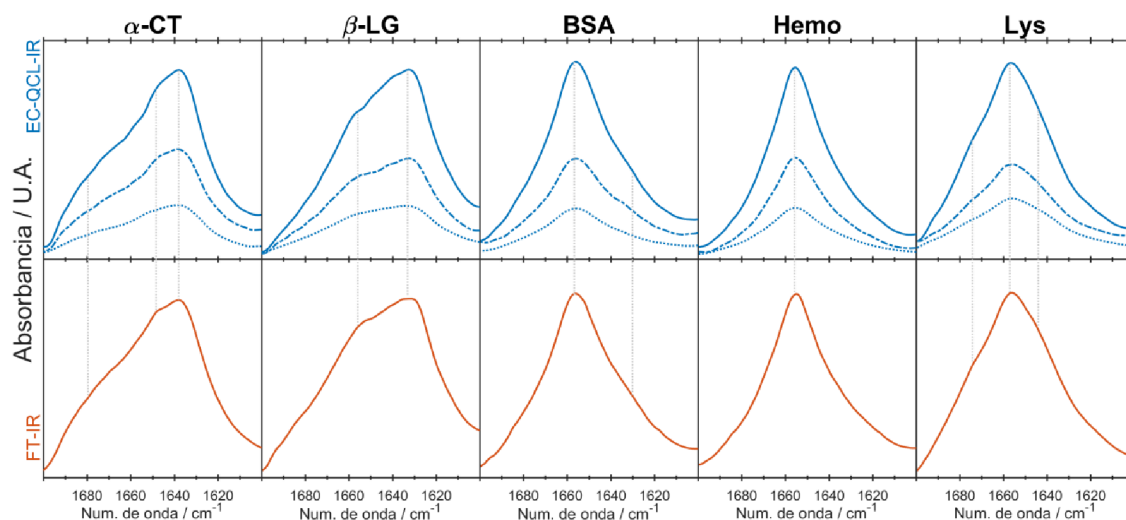


Figura 3.31. (Arriba) Espectros de absorbancia IR obtenidos por EC-QCL-IR de 5 soluciones proteicas conteniendo 10.0 mg mL^{-1} (línea continua azul), 5.0 mg mL^{-1} (línea de rayas y puntos azul) y 2.5 mg mL^{-1} (línea de puntos azul) y (abajo) obtenidos por FT-IR (línea continua roja) para soluciones proteicas conteniendo 20.0 mg mL^{-1} . Las líneas de punto grises remarcan la congruencia entre las formas espectrales de los espectros obtenidos con los diferentes instrumentos

Puede observarse que la posición de las bandas de absorción de los espectros

obtenidos por EC-QCL-IR se corresponde con la de los espectros obtenidos con el método convencional, FT-IR. Adicionalmente, a fin de cuantificar la congruencia entre los espectros, se evaluó el grado de solapamiento espectral (s_{12}) entre los espectros obtenidos por FT-IR (\mathbf{s}_1) y los obtenidos por EC-QCL-IR (\mathbf{s}_2), utilizando la ecuación 3.1. Este análisis arrojó valores de 0.99569, 0.99668, 0.99959, 0.99853 y 0.99783 para BSA, Hemo, α -CT, BLG y Lys, respectivamente, permitiendo inferir que los espectros adquiridos con el instrumento ensamblado en el laboratorio contienen información equivalente a los obtenidos por el método de referencia.

Para una evaluación más detallada y profunda, se calcularon y compararon los espectros de absorbancia derivados, como se muestran en la Figura 3.32. Esta metodología es comúnmente empleada en el análisis IR para la resolución de bandas solapadas en regiones espectrales complejas (Yang *et al.*, 2015).

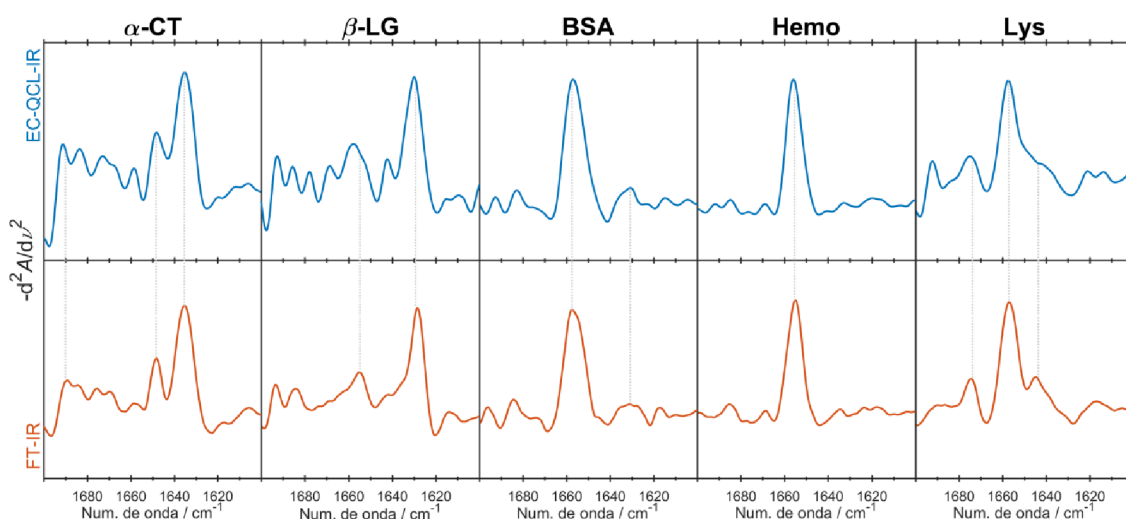


Figura 3.32. (Arriba) Espectros de segunda derivada invertidos obtenidos por EC-QCL-IR (líneas azules) y (abajo) FT-IR (líneas rojas) para las diferentes soluciones proteicas. Las líneas de punto grises remarcan la congruencia entre las formas espectrales de los espectros obtenidos con los diferentes instrumentos

La evaluación del grado de solapamiento espectral (s_{12}) entre los espectros de segunda derivada obtenidos por FT-IR (\mathbf{s}_1) y los obtenidos por EC-QCL-IR (\mathbf{s}_2) arrojó valores mayores a 0.9670, en todos los casos. Estas cifras resultaron altamente satisfactorias considerándose el aumento de sensibilidad en la diferenciación de bandas espectrales ocasionado por el uso de segundas derivadas. Así, se demostró que utilizando la metodología de procesamiento desarrollada para el análisis de señales EC-QCL-IR se pueden obtener espectros de alta calidad, conteniendo información equivalente a la obtenida por FT-IR.

Desnaturalización química de α -CT con monitoreo EC-QCL-IR. Con el propósito de evaluar la capacidad del instrumento EC-QCL-IR para realizar análisis dependientes del tiempo, como así también de la metodología de procesamiento desarrollada para el tratamiento de los datos, se utilizó un modelo cinético de desnaturalización química que consistió en monitorear el cambio dinámico de la estructura secundaria de la proteína α -CT en distintas condiciones desnaturalizantes. El cambio estructural analizado consiste en la transición de las estructuras hélices α inducidas por TFE, hacia la formación de agregados intermoleculares con estructuras de láminas β (β -agregación). Se evaluaron 2 condiciones experimentales diferentes, cambios de pH del medio y concentración proteica, analizándose las velocidades iniciales de la reacción de β -agregación.

En primer lugar, se registraron espectros EC-QCL-IR de α -CT en condiciones nativas no desnaturalizantes y, posteriormente, se monitoreó el efecto desnaturalizante ocasionado por la exposición de la proteína a TFE. El espectro de la α -CT nativa presenta un máximo de banda espectral amida I ubicado a 1638 cm^{-1} , la que es característica para los componentes de baja y alta frecuencia en estructuras secundarias de láminas β (Dong *et al.*, 1990). Ante la exposición a 50 % TFE en agua, se produce un cambio en el máximo de la banda amida I hacia una posición ubicada a 1654 cm^{-1} , valor representativo de las estructuras proteicas helicoidales α (Dong *et al.*, 1990). Este fenómeno observado concuerda con lo descrito en trabajos reportados en referencia a la generación de estructuras secundarias helicoidales α ante la exposición a TFE de proteínas nativas compuestas mayoritariamente por estructuras de láminas β (Dong *et al.*, 1998; Jasanoff *et al.*, 1994). La transición inducida por el alcohol desde la estructura secundaria β nativa hacia la estructura helicoidal α inducida se lleva a cabo en un periodo de tiempo en el rango de milisegundos (Lu *et al.*, 1997), dificultándose el monitoreo por EC-QCL-IR de la transición, debido a que la obtención de cada espectro IR demanda un tiempo no menor a 1.5 s. Las estructuras helicoidales α inducidas por TFE de proteínas ricas en lámina β en su configuración nativa, como α -CT, no son estables (Dong *et al.*, 1998) y dependen de la proporción de TFE presente en el medio. Consecuentemente, las estructuras de hélices α inducidas, al no ser estables, comienzan a plegarse formando una nueva conformación rica en láminas β .

En la Figura 3.33 se muestra el cambio gradual de los espectros IR de la proteína α -CT desde una conformación de hélices α inducidas hacia una de láminas β , luego de la exposición a TFE, en el rango de tiempo de entre 2 min y 240 min y el espectro IR de la proteína nativa en condiciones no desnaturalizantes.

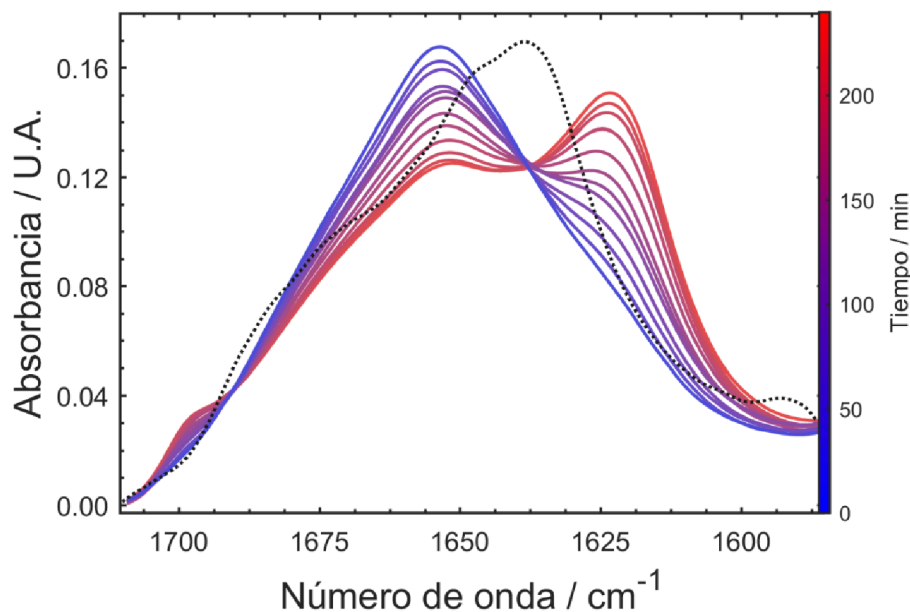


Figura 3.33. Espectros de absorbancia IR obtenidos por EC-QCL-IR para una solución conteniendo 20.0 mg mL⁻¹ de α -CT en TFE/PBS (50:50) pH 7.0 a 25 °C. Los espectros se registraron en un periodo de tiempo comprendido entre 2 min y 240 min. Las líneas azules muestran los espectros de α -CT conformada mayoritariamente por hélices α inducidas, mientras que las líneas rojas muestran los espectros de la proteína luego de la formación gradual de láminas β intermoleculares. El espectro correspondiente a α -CT nativa se muestra en línea de puntos negra. A la derecha se indica la transición conformacional dependiente del tiempo de reacción en escala de colores

Es notable la diferencia que se observa entre los espectros correspondientes a la proteína nativa y a la proteína en condiciones desnaturalizantes, mostrándose un importante cambio en la posición del máximo de banda desde ~ 1640 cm⁻¹ a ~ 1654 cm⁻¹, lo que indica la formación de estructuras α inducidas por TFE. A medida que transcurre la reacción, la intensidad de la banda ubicada a 1654 cm⁻¹, correspondiente a la estructura de hélices α inducidas, disminuye mientras que aparecen 2 bandas ubicadas a 1623 cm⁻¹ y 1697 cm⁻¹. Este incremento de las bandas espectrales es atribuido, generalmente, a los agregados de estructuras láminas β anti-paralelas (Filosa *et al.*, 2001; Murayama *et al.*, 2004).

Experimentalmente se comprobó, además, que las proteínas que presentan estructura secundaria nativa predominantemente con configuración helicoidal α , por ejemplo, BSA o HEMO, no revelan cambios espectrales en el periodo de tiempo evaluado en idénticas condiciones, lo que indicaría la no formación de láminas β intermolecular en virtud de la alta estabilidad de la conformación de las hélices α inducidas.

Modelado de datos por MCR-ALS. El análisis quimiométrico se realizó aplicando descomposición bilineal de los datos cinéticos con MCR-ALS en su versión extendida, aprovechando la capacidad de obtener perfiles físicamente interpretables de los componentes involucrados en el sistema de análisis. Para el análisis con MCR-

ALS se utilizaron los espectros no filtrados, es decir, los espectros de absorbancia obtenidos posterior a la etapa COW 2, evitando la etapa de filtrado por FFT, debido a que la presencia de vapor agua y el ruido residual, principalmente en niveles bajos de concentración proteica, introducen desviaciones en los espectros que, si bien no son visiblemente significativos, dificultarían la correcta resolución del problema. Además, para favorecer el modelado quimiométrico, se aplicó un método de corrección de línea de base multidimensional basado en el método de cuadrados mínimos asimétricos propuesto por Eilers (Eilers, 2005).

Se realizaron 2 análisis independientes con el objetivo de evaluar los efectos ocasionados sobre la velocidad de agregación ante diferentes condiciones experimentales. Por un lado, se obtuvieron 7 matrices de datos correspondientes al estudio del efecto del pH con las que se construyó la matriz aumentada en columnas \mathbf{D}_{pH} . Por otro lado, las matrices de datos obtenidas para el análisis del efecto de concentración proteica se agruparon construyéndose la matriz \mathbf{D}_{con} , también aumentada en columnas. Posteriormente, la descomposición bilineal se llevó a cabo para cada para cada matriz extendida, individualmente.

El número de componentes presentes en el sistema se determinó mediante SVD y la obtención de las estimaciones iniciales de los perfiles temporales para la inicialización de MCR-ALS se realizó utilizando una metodología basada en SIMPLISMA (Windig *et al.*, 1991). En la etapa de la optimización iterativa por ALS de la descomposición bilineal se aplicaron las restricciones de unimodalidad en modo temporal y no-negatividad en modo espectral y temporal. Así, se obtuvieron los perfiles temporales (\mathbf{C}) y espectrales (\mathbf{S}) conteniendo la evolución estructural $\alpha \rightarrow \beta$ dependiente del tiempo, a cada valor de pH o nivel de concentración, y los perfiles espectrales correspondientes a las distintas conformaciones de la proteína, respectivamente. Es importante destacar que, debido a la falta de información relacionada a la cinética, no fue posible obtener *a priori* los parámetros cinéticos relacionados a la reacción química involucrada y, en consecuencia, no se pudieron utilizar herramientas quimiométricas de modelado rígido, las que son sugeridas en la literatura para el estudio de sistemas cinéticos (Bezemer *et al.*, 2001; de Juan *et al.*, 2000; Sánchez-Ponce *et al.*, 2005).

Efecto del pH sobre la velocidad de agregación. Se evaluó el efecto del pH del medio sobre la formación de láminas β intermoleculares inducidas por TFE para una solución conteniendo 20.0 mg mL⁻¹ de α -CT en un rango de pH de entre 5.8 y 8.2. En primer lugar, se realizó una inspección visual de los cambios físicos de las solucio-

nes en todas las condiciones experimentales estudiadas, no observándose fenómenos de precipitación y/o gelación de la proteína, como tampoco incremento en la turbidez de la solución, en un periodo de hasta 24 h posteriores al inicio de la reacción. Esta observación sugiere que los efectos mostrados para las condiciones experimentales difieren de los efectos de agregación para fibrillas amiloideas (Khodarahmi *et al.*, 2009; Pallarès *et al.*, 2004; Rezaei-Ghaleh *et al.*, 2007). Los resultados coinciden, además, con trabajos recientes que demuestran que a altas concentraciones de TFE (> 50 %) la proteína se encuentra en un estado no-nativo deficiente de agregación y no es propensa a la formación de fibrillas tipo amiloideas (Rezaei-Ghaleh *et al.*, 2008; Rezaei-Ghaleh *et al.*, 2007).

Los perfiles temporales y espectrales obtenidos por MCR-ALS para la matriz D_{pH} se muestran en la Figura 3.34.

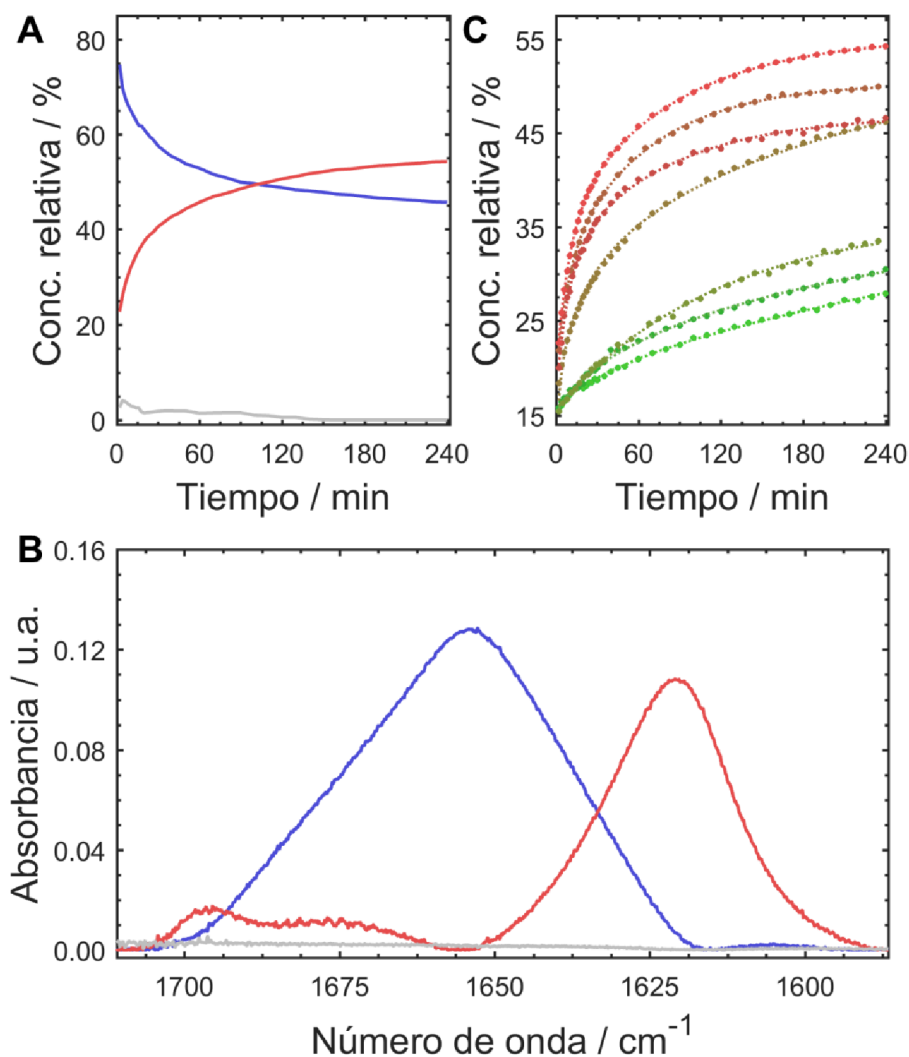


Figura 3.34. Perfiles temporales relativos (A) y espectrales (B) obtenidos por MCR-ALS del monitoreo de la reacción de β -agregación de 20.0 mg mL^{-1} de α -CT en TFE/PBS pH 8.2, mostrándose la evolución y los perfiles espectrales de las estructuras secundarias correspondientes a láminas β (rojo) y hélices α (azul). En gris se muestra el perfil obtenido para el ruido instrumental. Los perfiles temporales obtenidos para las láminas β a los diferentes valores de pH analizados (C) se muestran en escala de colores de verde a rojo para valores desde pH de 5.8 a 8.2, respectivamente

Se puede apreciar, claramente, la fuerte dependencia que presenta la formación de láminas β con el valor de pH del medio (Figura 3.34.C), mostrándose que a pH elevados se presenta la mayor tendencia a la formación de agregados. Es importante remarcar, además, que para valores altos de pH, la reacción parece concluir dentro de las 4 h, mientras que la agregación a bajos valores de pH no parece ser completada dentro del periodo de tiempo observado.

El comportamiento de la reacción de β -agregación en dependencia con el pH se evaluó cuantitativamente a partir de la velocidad de agregación inicial (v_0), la que se consideró como la pendiente obtenida del ajuste lineal de la concentración de láminas β respecto al tiempo de los primeros 4 tiempos registrados. En la Figura 3.35 se observa el cambio ocurrido en la v_0 en dependencia a los valores de pH del medio, mostrándose una progresión sigmoidea con valores bajos a pH ácidos y valores altos a pH alcalinos, alcanzándose el punto de transición a un valor de pH cercano a 7.0.

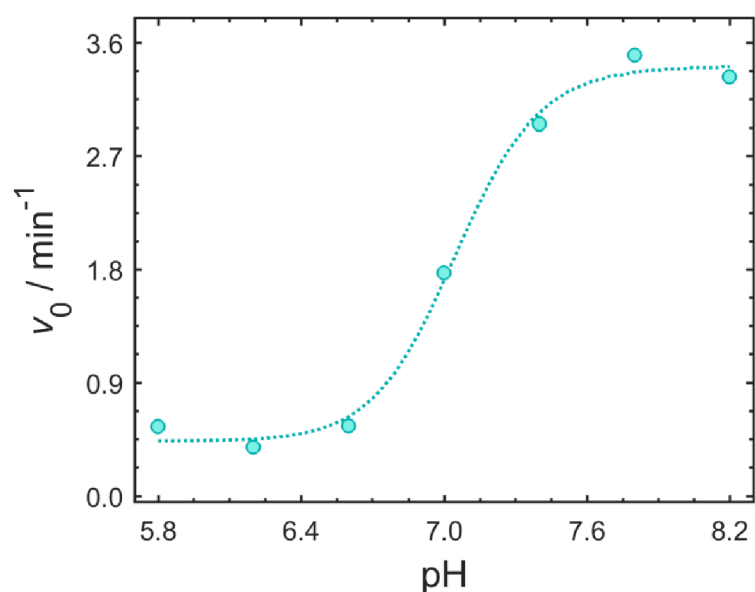


Figura 3.35. Efecto del pH del medio sobre la velocidad inicial v_0 de la reacción de agregación de 20.0 mg mL^{-1} de α -CT en TFE/PBS. El valor de v_0 se calculó utilizando los perfiles temporales obtenidos por MCR-ALS

En general, las proteínas en soluciones acuosas presentan baja solubilidad a valores de pH cercanos a su punto isoelectrico (pI). En las condiciones experimentales aquí evaluadas, la proteína presenta grupos cargados positiva y negativamente, lo que genera una distribución anisotrópica de cargas sobre su superficie, dando origen a posibles dipolos. En consecuencia, las interacciones proteína-proteína, o intermoleculares, se encuentran altamente favorecidas, traducándose en procesos de agregación

energéticamente favorables (Chi *et al.*, 2003; Wang *et al.*, 2010). Es posible, además, que el efecto electrostático desestabilizante de la conformación proteica a valores de pH cercanos al pI se intensifique por la adición de TFE. Este fenómeno se refleja en el incremento de la v_0 de la formación de láminas β a valores de pH > 7.0 ($pI_{\alpha\text{-CT}} = 8.4$), resultando comparable a los resultados reportados para la formación de fibrillas amiloides a concentraciones de TFE intermedias (Rezaei-Ghaleh *et al.*, 2007). Es importante mencionar que experimentalmente no se observaron agregados β en condiciones no desnaturizantes dentro del rango de pH evaluado, es decir, en ausencia de TFE.

Finalmente, se evaluó la reproducibilidad del sistema monitoreando la formación de láminas β de una solución de 20.0 mg mL⁻¹ de α -CT a pH 7.0, por triplicado. Estas condiciones se seleccionaron debido a que el punto de inflexión correspondiente a la curva de v_0 respecto a pH se encuentra cercano a este valor, esperándose, en consecuencia, la máxima variabilidad del sistema. El coeficiente de variación de la v_0 obtenido resultó igual a 4.2 %, certificándose la excelente precisión del método para el monitoreo de agregación proteica utilizando EC-QCL-IR.

Efecto de la concentración proteica sobre la velocidad de agregación. Se monitoreó la formación de láminas β para el rango de niveles de concentración proteica comprendido entre 5.0 mg mL⁻¹ y 60.0 mg mL⁻¹ en 50 % TFE/PBS a pH 6.6. Previamente, se observó que en concentraciones menores a 5.0 mg mL⁻¹ de proteína se produce un efecto de gelación con consecuente precipitación en un periodo menor a 10 s, posterior a la disolución con 50 % TFE/PBS. En consecuencia, se seleccionaron aquellas concentraciones que no mostraron precipitación o formación visible de agregados dentro de las 4 h posteriores a la disolución.

Se estudió la evolución temporal del contenido de láminas β a diferentes concentraciones de proteína utilizando MCR-ALS y evaluándose la v_0 en dependencia al valor de concentración. Los perfiles temporales y espectrales obtenidos por MCR-ALS se muestran en la Figura 3.36. En la Figura 3.36.C es posible notar la dependencia que presenta la formación de láminas β intermoleculares con la concentración de proteína.

Para el estudio de la dependencia de la v_0 con la concentración proteica, fue necesario realizar dos análisis separadamente: **1-** *considerándose valores de v_0 absolutos*; y **2-** *considerándose valores de v_0 relativos a la concentración*. La Figura 3.37 muestra el comportamiento de la v_0 con respecto a la concentración de α -CT en términos absolutos (Figura 3.37.A) y relativos a la concentración (Figura 3.37.B).

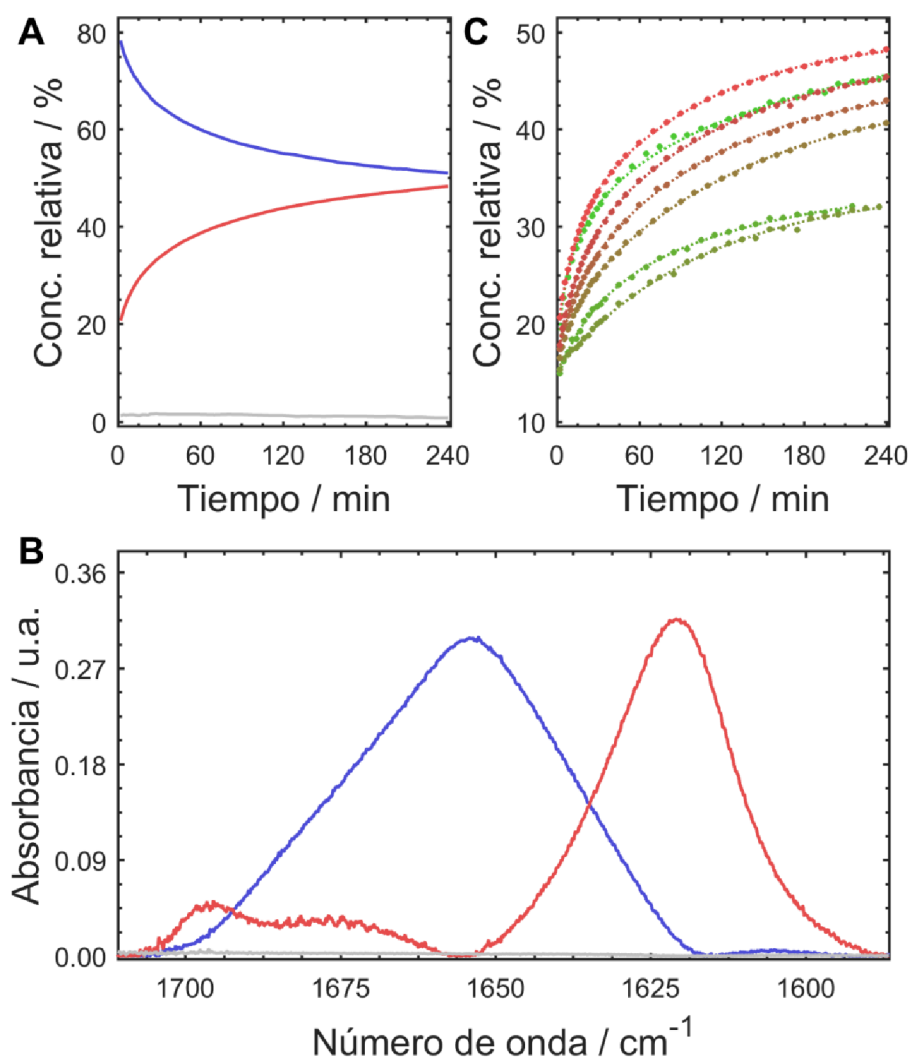


Figura 3.36. Perfiles temporales relativos (A) y espectrales (B) obtenidos por MCR-ALS del monitoreo de la reacción de β -agregación de 60.0 mg mL^{-1} de α -CT en TFE/PBS pH 6.6, mostrándose la evolución y el perfil espectral de las estructuras secundarias correspondientes a láminas β (rojo) y hélices α (azul). En gris se muestra el perfil obtenido para el ruido instrumental. Los perfiles temporales obtenidos para las láminas β en los diferentes niveles de concentración proteica (C) se muestran en escala de colores desde verde a rojo para valores de concentración desde 5.0 mg mL^{-1} a 60.0 mg mL^{-1} , respectivamente

A partir del análisis de valores de v_0 absolutos, calculadas con la abundancia absoluta obtenida de MCR-ALS, los resultados sugieren que a bajas concentraciones de proteína la velocidad de agregación es independiente de la concentración. Sin embargo, se observó que la v_0 relativa presenta valores considerablemente elevados a niveles bajos de concentración proteica, lo que explicaría las observaciones evidenciadas a concentraciones menores a 5.0 mg mL^{-1} : *gelación y precipitación inmediatamente después a la disolución*. Por otro lado, el cambio característico del comportamiento de v_0 para concentraciones mayores a 20.0 mg mL^{-1} indicaría que este valor de concentración representa la concentración crítica de la reacción en estudio, comportamiento que también ha sido demostrado para la agregación de numerosas proteínas en con-

diciones desnaturalizantes (Dong *et al.*, 1998; Wang *et al.*, 2010). Generalmente, el incremento en la formación de agregados a altos niveles de concentración proteica se debe, en parte, a la mayor probabilidad de asociaciones proteína-proteína, es decir, al mayor número de interacciones inter-moleculares.

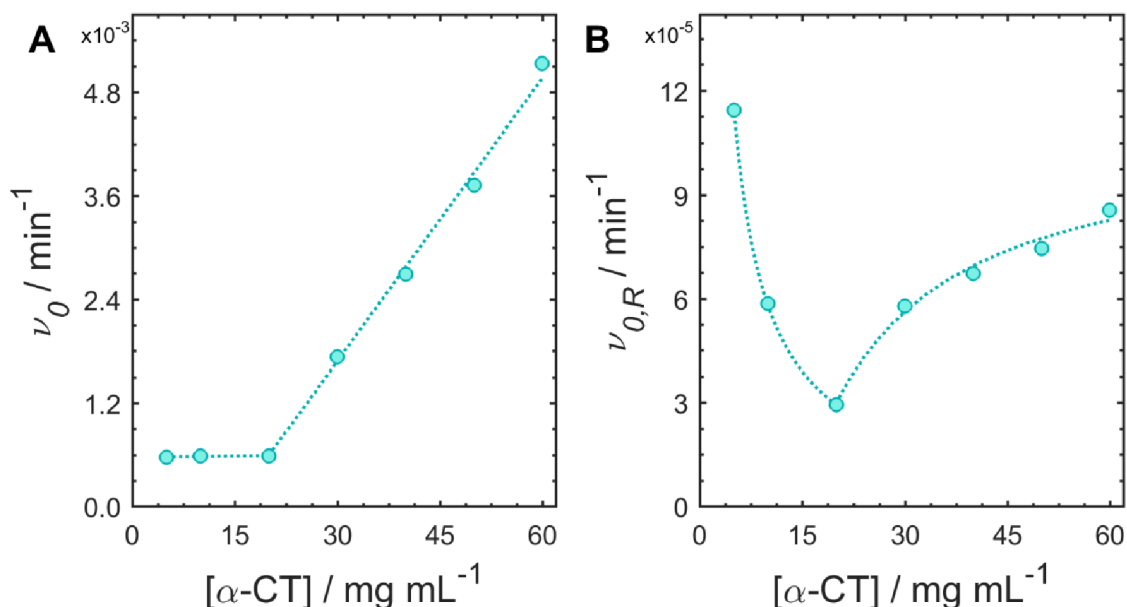


Figura 3.37. Velocidad inicial absoluta v_0 (A) y relativa $v_{0,R}$ (B) de la reacción de β -agregación para diferentes niveles de concentración de α -CT en TFEPBS pH 6.6, calculadas a partir de los perfiles temporales obtenidos de MCR-ALS

Análisis espectral del proceso de β -agregación. Mediante el análisis por MCR-ALS se pudieron identificar 3 componentes principales: 2 involucrados en el proceso de agregación y 1 referido al ruido instrumental residual del procesamiento de señales basado en COW. Es posible notar que los perfiles espectrales obtenidos presentan los máximos de banda y la forma característica de los componentes asociados a las distintas conformaciones estructurales de la proteína. Uno de los componentes muestra un máximo de banda a 1654 cm^{-1} , asignado a las estructuras secundarias proteicas helicoidales α (Barth, 2007), cuyo decaimiento a lo largo del tiempo (Figura 3.34.A y Figura 3.36.A) concuerda con el decaimiento del contenido de hélices α durante la β -agregación. El segundo componente involucrado en el proceso de agregación muestra una banda pronunciada a 1621 cm^{-1} con una banda más débil a 1695 cm^{-1} , típicas de los patrones espectrales asociados a láminas β anti-paralelas intermoleculares (Chirgadze *et al.*, 1976; Yang *et al.*, 2015); además, se observa una banda secundaria ubicada a 1685 cm^{-1} , la que puede ser asociada a conformaciones de giros β (Yang *et al.*, 2015). El componente restante, referido al ruido instrumental, no presenta nin-

guna forma característica en los perfiles obtenidos, como tampoco ningún cambio representativo en su intensidad a lo largo del tiempo.

Los resultados espectrales obtenidos por el análisis de MCR-ALS para los dos estudios realizados independientemente, pH y concentración proteica, demuestran la robustez de la descomposición bilineal para sistemas de los que no se tiene información preliminar.

CONCLUSIÓN

Por primera vez, se utilizó un instrumento para mediciones de transmisión IR usando un láser EC-QCL como fuente lumínica para el análisis de estructuras secundarias de proteínas en solución acuosa. Además, se desarrolló una rutina de procesamiento de señales que resultó crucial para la disminución de las variaciones inherentes de los espectros de transmisión obtenidos con EC-QCL-IR. La utilización de una nueva tecnología láser combinada a un correcto procesamiento de los datos permitió evaluar características espectrales de proteínas con distintas estructuras secundarias en medio acuoso en concentraciones tan bajas como 2.5 mg mL^{-1} . La evaluación matemática del solapamiento espectral y la comparación de los espectros derivados confirman la excelente comparabilidad entre las mediciones EC-QCL-IR de transmisión con los espectros de proteína adquiridos por FT-IR convencional de transmisión. Asimismo, se demostró la validez de la ley de Lambert-Beer para un rango de concentración proteica de entre 1.0 mg mL^{-1} y 20.0 mg mL^{-1} .

El alto poder óptico de la fuente de luz láser permitió realizar mediciones de transmisión IR de proteínas utilizando pasos ópticos mayores ($\leq 38 \text{ }\mu\text{m}$) a los utilizados en espectrofotómetros FT-IR convencionales ($\sim 8 \text{ }\mu\text{m}$). El aumento en la longitud del paso óptico accesible abre un amplio rango de nuevas posibilidades experimentales para mediciones IR en una región espectral que ha sido limitada debido a la gran absorbancia del agua, cuando esta es utilizada como solvente. Esto significa un gran paso para las investigaciones de proteínas por IR en medio acuoso, las que han sido llevadas a cabo mayoritariamente en medio deuterado acompañadas de complicados procedimientos de preparación de muestras y ambigüedades respecto a su ambiente no-nativo.

Además, se monitoreó la formación gradual de agregados de láminas β en α -CT posterior a la inducción de estructuras de hélices α no-nativas por exposición a TFE,

para distintos valores de pH y en distintos niveles de concentración proteica. Se comprobó que la v_0 de la β -agregación aumenta cuando el pH del medio se acerca al valor correspondiente al pI de la proteína, debido, en parte, a las grandes interacciones electroestáticas atractivas inter-moleculares presentes. Se demostró la dependencia de la velocidad de β -agregación con la concentración proteica, obteniéndose un valor de concentración peptídica crítica igual a 20.0 mg mL^{-1} . Los resultados obtenidos sugieren que el TFE maximiza los puentes de hidrógenos intermoleculares y reduce los puentes hidrógenos formados entre las moléculas de proteínas y el agua circundante, induciéndose conformaciones proteicas compactas. La formación gradual de agregados de láminas β desde estructuras helicoidales α inducidas por TFE concuerda con estudios anteriores que demuestran que las estructuras de hélices α representan un estado cinéticamente favorecido, mientras que las láminas β constituyen al estado preferiblemente termodinámico de la α -CT (Dong *et al.*, 1998).

Por otro lado, los resultados espectrales obtenidos por MCR-ALS de los experimentos dependientes de pH y concentración proteica muestran un alto grado de similitud, demostrándose la alta calidad de los espectros IR obtenidos por EC-QCL-IR y la potencialidad del instrumento para el monitoreo de reacciones cinéticas a condiciones experimentales límites, como baja concentración proteica o en reacciones que muestran una elevada velocidad de reacción.

3. 5. MONITOREO Y ANÁLISIS ESTRUCTURAL DE CAMBIOS CONFORMACIONALES INDUCIDOS POR TEMPERATURA DE UN HOMO-POLIPÉPTIDO MEDIANTE DICROÍSMO CIRCULAR, FT-IR Y EC-QCL-IR CON RESOLUCIÓN QUIMIOMÉTRICA DE DATOS

External cavity-quantum cascade laser infrared spectroscopy for secondary structure analysis of proteins at low concentration – Schwaighofer A. & Alcaráz M.R., et al., Sci Rep, 2016, 6, 33556-33566

Application of MCR-ALS to reveal intermediate conformations in the thermally induced α - β transition of poly-L-lysine monitored by FT-IR spectroscopy– Alcaráz, et al., Enviado para revisión y posterior publicación a Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy

CONSIDERACIONES GENERALES

El análisis estructural de biomoléculas es de gran importancia en la biología y la bioquímica, principalmente, para la caracterización de las propiedades de plegamiento de proteínas y polipéptidos, como también para el monitoreo de los cambios dinámicos ocurridos ante perturbación. Los métodos analíticos usados para la investigación de estructuras secundarias de biomoléculas incluyen cristalografía por rayos X (*XRC: X-Ray Crystallography*), resonancia magnética nuclear (*NMR: Nuclear Magnetic Resonance*), espectroscopia de dicroísmo circular (*CD: Circular Dichroism*), espectroscopia Raman y espectroscopia infrarroja acoplada a transformada de Fourier (*FT-IR*) (Kelly et al., 2005; Pelton et al., 2000; Yang et al., 2015). Si bien NMR y XRC son capaces de proveer información estructural a niveles atómicos de resolución, NMR se encuentra restringida al estudio de biomoléculas relativamente pequeñas (≤ 40 kDa) en altos niveles de concentración y XRC requiere de cristales proteicos de alta pureza, que no siempre es posible obtener, particularmente, cuando se analizan proteínas integrales de membrana (Moraes et al., 2014).

Por el contrario, la espectroscopia de CD, Raman o FT-IR son consideradas técnicas de baja resolución que proveen información estructural general de las macromo-

lécúlas. Sin embargo, estas técnicas son rutinariamente utilizadas para la determinación de estructuras secundarias de proteínas y para el monitoreo de cambios dinámicos en las estructuras proteicas, debido a que no requieren de preparaciones complejas de muestra y demandan tiempos de análisis relativamente cortos. Debido a sus respectivas características, la espectroscopia IR provee estimaciones dependientes de láminas β anti-paralelas, mientras que la espectroscopia de CD brinda predicciones más precisas acerca de las estructuras de hélices α (Pribic *et al.*, 1993). Considerándose la información complementaria que proveen estas técnicas, su aplicación conjunta llevaría consigo la obtención de resultados más confiables (Bagińska *et al.*, 2008; Czarnik-Matusiewicz *et al.*, 2006).

La espectroscopia de CD es una técnica basada en las diferencias en la absorción de luz circularmente polarizada hacia la izquierda y hacia la derecha al interactuar con compuestos ópticamente activos, es decir, con especímenes quirales que no presenten plano de simetría (Parson, 2015a; b). En proteínas, el cromóforo más relevante es el grupo amida del enlace peptídico que presenta absorbanza en la región UV lejana del espectro electromagnético. El enlace peptídico, si bien, debido a su carácter parcial de doble enlace, no es ópticamente activo por sí mismo, su entorno quiral hace que sea posible su detección por CD. Adicionalmente, en virtud de que las transiciones electrónicas del enlace peptídico presentan una fuerte dependencia de las posiciones relativas y las orientaciones del grupo peptídico, los espectros CD proveen información sensible acerca de la estructura secundaria de proteínas y polipéptidos (Greenfield, 2007a; b). De las 5 transiciones electrónicas posibles del grupo peptídico, la transición $\pi \rightarrow \pi^*$ presenta la mayor energía y fuerza de dipolo eléctrico, visualizándose como una banda positiva a 190 nm del espectro CD, mientras que la transición $n \rightarrow \pi^*$, de muy baja energía, se identifica como una banda negativa centrada a 210 nm. Las transiciones $n \rightarrow \pi^*$ no presentan fuerzas de dipolo eléctrico, pero pueden ganar fuerza dipolar a través de las interacciones con los excitones de las transiciones $\pi \rightarrow \pi^*$ de los residuos cercanos, provocando la aparición de bandas negativas a longitudes de onda diferentes. Así, las distintas configuraciones estructurales adoptadas por las proteínas o los polipéptidos presentan bandas características en la región UV lejano permitiendo su identificación a partir de los espectros CD, por ejemplo, un modelo simple de hélices α se caracteriza por presentar un espectro CD con 2 bandas negativas ubicadas a 205 nm y 220 nm, mientras que para un modelo de láminas β se obtienen espectros con una única banda negativa centrada en 215 nm (Greenfield, 2007b; Parson, 2015a).

Por otro lado, las mediciones de transmisión IR son empleadas con frecuencia

para estudios estructurales de biomoléculas en solución. Para proteínas, las bandas de absorción más prominentes observadas en el espectro IR son las bandas amida causadas por las vibraciones del grupo peptídico, siendo la banda amida I y amida II las utilizadas comúnmente para el análisis de estructuras secundarias en proteínas y péptidos (Barth, 2007; Oberg *et al.*, 2004). Para las mediciones de proteínas en solución acuosa, la fuerte absorbancia de la banda correspondiente a la torsión del enlace HOH del agua, ubicada a 1645 cm^{-1} , se solapa con la banda amida I de las proteínas. Para superar este inconveniente se utilizan, alternativamente, soluciones deuteradas como solvente, donde la banda correspondiente a la torsión del enlace DOD se localiza a $\sim 1200\text{ cm}^{-1}$, creándose una región con absorbancia relativamente baja debida al solvente en la región espectral de entre 1500 cm^{-1} y 1800 cm^{-1} .

En soluciones deuteradas, la banda amida II, originada predominantemente por las vibraciones ocasionadas por estiramientos del grupo N-H, se centra en 1550 cm^{-1} , mientras que en soluciones acuosas se encuentra en 1450 cm^{-1} , denominándose, en consecuencia, banda amida II'. Por otro lado, en virtud de que la banda amida I es compuesta mayoritariamente por vibraciones de los grupos C=O, el corrimiento espectral observado en soluciones deuteradas es relativamente pequeño comparado con la banda amida II (entre 5 cm^{-1} y 10 cm^{-1}) (Dong *et al.*, 1995; Fabian *et al.*, 2006).

La α -poli-L-lisina (PLL) es un homo-polipéptido ampliamente utilizado como sustancia modelo para las investigaciones de estructura secundaria. Es una biomolécula que puede adoptar diferentes conformaciones estructurales dependiendo de un amplio rango de condiciones físicas y químicas. En la literatura se pueden encontrar investigaciones realizadas utilizando FT-IR acerca de los cambios estructurales inducidos por perturbaciones externas, por ejemplo, valor de pH (Dzwolak *et al.*, 2005), presión (Carrier *et al.*, 1990) y temperatura (Dzwolak *et al.*, 2004; Szyk *et al.*, 2008). Adicionalmente, se ha demostrado que su estado conformacional depende de la longitud de la cadena peptídica (Dzwolak *et al.*, 2004), la concentración peptídica (Yu *et al.*, 1973) y el grado de hidratación, principalmente, cuando se encuentra en forma de película fina (Ramer *et al.*, 2015). Se observó que en solución acuosa a valores de pH mayores a su pK_a ($pK_a = 10.5$) y a presión atmosférica, las cadenas laterales de lisina se encuentran desprotonadas y la PLL adopta una conformación de hélices α que, ante el incremento de la temperatura, se transforma gradualmente hacia una conformación de láminas β (Dzwolak *et al.*, 2004; Szyk *et al.*, 2008).

En este trabajo, se realizó un análisis conformacional de proteínas en medio deuterado utilizando la metodología EC-QCL-IR desarrollada en el capítulo anterior, a fin de evaluar la capacidad del instrumento para realizar mediciones de transmisión IR

en medio deuterado. Adicionalmente, se utilizó espectroscopia de CD, FT-IR y EC-QCL-IR para evaluar la dependencia de la temperatura de transición del cambio conformacional $\alpha \rightarrow \beta$ inducido del homo-polipéptido PLL con la concentración en un amplio rango de concentraciones. Finalmente, se realizó un análisis en detalle de los datos obtenidos por espectroscopia de CD y FT-IR para el monitoreo del cambio conformacional inducido por temperatura de PLL utilizando MCR-ALS como herramienta quimiométrica.

MATERIALES Y MÉTODOS

Reactivos. El fosfato de sodio tribásico anhidro (Na_3PO_4 , tech) fue adquirido de Alfa-Aesar (Karlsruhe, Alemania). El óxido de deuterio (D_2O , 99.9 atom % D), la solución de cloruro de deuterio 35 % en D_2O (DCI, 99 atom % D), el deuteróxido de sodio 30 % en D_2O (NaOD, 99 atom % D), los liofilizados de albúmina de clara de huevo de gallina (OVA), concanavalina A de *Canavalia ensiformis* (Jack bean) tipo IV (ConA), albúmina de suero bobino (BSA, $\geq 98.0\%$) e hidrobromuro de α -poli-L-lisina (PLL, PM: 15000–30000 Da por viscosidad) fueron obtenidos de Sigma-Aldrich (Steinheim, Alemania) y usados sin tratamiento previo.

Soluciones de trabajo. Se prepararon 2 soluciones reguladoras deuterada de fosfatos (PBSD) de concentración formal igual a 0.050 mol L^{-1} y 0.010 mol L^{-1} , disolviendo cantidades apropiadas de Na_3PO_4 anhidro en D_2O y ajustando el pH (correspondiente a $pD = pH + 0.4$ (Glasoe *et al.*, 1960)) al valor correspondiente con DCI o NaOD.

Las soluciones madres de proteína se prepararon diariamente por disolución de 10.0 mg del polvo liofilizado directamente en 1.0 mL de PBSD 0.050 mol L^{-1} pD 7.0 para BSA y OVA y pD 6.2 para ConA. Se preparó un set de 4 soluciones diluidas para cada proteína individualmente transfiriendo la alícuota apropiada de la solución madre y completándose el volumen a 1.0 mL con la respectiva solución PBSD. Para las mediciones por espectroscopia EC-QCL-IR se prepararon soluciones en un rango de concentraciones de entre 0.25 mg mL^{-1} y 2.00 mg mL^{-1} y para las mediciones por espectroscopia FT-IR se prepararon soluciones a una concentración de 5.00 mg mL^{-1} .

Las soluciones madres de PLL se prepararon diariamente por disolución de

10.0 mg de PLL liofilizada directamente en 1.0 mL de solución PBS-D pH 12.4. Se preparó 1 set de soluciones diluidas de PLL transfiriendo el volumen adecuado de la solución madre de PLL y completándose el volumen de 1.0 mL con solución PBS-D pH 12.4, controlándose y ajustándose el pH de cada solución preparada, en caso de ser necesario. Para las mediciones por espectroscopia de CD, se preparó 1 set de soluciones en un rango de concentración de PLL comprendido entre 0.25 mg mL^{-1} y 1.50 mg mL^{-1} en solución PBS-D 0.010 mol L^{-1} pH 12.4. Para estas mediciones se utilizó la PBS-D de menor concentración formal debido a que los iones fosfatos muestran alta absorción CD a longitudes de onda menores a 210 nm (Kelly *et al.*, 2005). Experimentos preliminares demostraron que la concentración de la solución reguladora no tiene efecto sobre la temperatura de transición del homo-polipéptido. Para las mediciones por espectroscopia EC-QCL-IR se preparó 1 set de soluciones de PLL en concentraciones finales de entre 0.25 mg mL^{-1} y 6.00 mg mL^{-1} en solución PBS-D 0.050 mol L^{-1} . Para las mediciones por FT-IR, se prepararon soluciones de PLL en un rango de concentraciones comprendido entre 5.00 mg mL^{-1} y 18.00 mg mL^{-1} en solución PBS-D 0.050 mol L^{-1} pH 12.4. En todos los casos, la solución PBS-D correspondiente se registró como referencia.

Instrumentación y procedimiento. Todas las soluciones se prepararon dentro de una caja de guantes de polietileno para preparación bajo corriente continua de N_2 seco.

Las mediciones de pH se realizaron utilizando un potenciómetro pH 330i 410A (Wissenschaftlich-Technische Werkstätten GmbH, Weilheim, Alemania) equipado con un micro-electrodo de vidrio combinado Sentix® MicD (Wissenschaftlich-Technische Werkstätten GmbH, Weilheim, Alemania), acoplado a un sensor de temperatura.

Mediciones por espectroscopia EC-QCL-IR. Las mediciones espectroscópicas de absorción laser infrarroja (*EC-QCL-IR*) se llevaron a cabo utilizando un dispositivo ensamblado en el laboratorio comprendido por un EC-QCL (Daylight Solutions Inc., San Diego, Estados Unidos) termoeléctricamente enfriado, operando a una velocidad de repetición de 100 kHz y un ancho de pulso de 500 ns, y un detector MCT con enfriamiento termoeléctrico ($-60 \text{ }^\circ\text{C}$, $D^* = 4 \times 10^9 \text{ cm Hz}^{0.5} \text{ W}^{-1}$ a $9.2 \text{ } \mu\text{m}$). Para el procesamiento de la señal detectada se usó un integrador de dos canales y para la digitalización de la señal se usó un controlador analógico-digital (ADC) NI DAQ 9239 de 24 bits (National Instruments Corp., Austin, EE.UU). Para el control del dispositivo se utili-

zó un controlador desarrollado en el laboratorio soportado en LabView GUI 11.0 (National Instruments Corp., Austin, EE.UU) con estructura servidor-cliente. Se utilizó una celda de flujo construida en el laboratorio, equipada con dos ventanas de CaF_2 y un espaciador de PTFE de 478 μm de espesor. Para reducir la influencia del vapor de agua, se aisló el instrumento en una caja de polietileno con flujo constante de aire seco.

Para los espectros de proteínas (BSA, OVA y ConA) se registraron 100 mediciones de transmisión consecutivas a una temperatura constante de 25 °C.

Para el análisis de cambio conformacional en PLL inducido por temperatura se registró un total de 20 mediciones de transmisión consecutivas (tiempo total de adquisición = 100 s) a cada nivel de temperatura, en un rango de temperatura comprendido entre 20 °C y 50 °C ($\Delta T = 2$ °C). La celda de medición se estabilizó durante 240 s a cada valor de temperatura previo a cada medición. A fin de minimizar el ruido espectral resultante de las imperfecciones mecánicas de la fuente de luz EC-QCL, se utilizó la rutina de procesamiento de datos basada en COW, explicada en el apartado **Procesamiento de señales obtenidas por EC-QCL-IR** de la sección 3.4.

Mediciones por espectroscopia FT-IR. Las mediciones espectroscópicas de absorción infrarroja acoplada a transformada de Fourier (FT-IR) se realizaron utilizando un espectrofotómetro FT-IR Bruker 80v (Ettlingen, Alemania), equipado con un detector MCT operado con LN_2 ($D^* = 4 \times 10^{10} \text{ cm Hz}^{0.5} \text{ W}^{-1}$ a 9.2 μm). Se utilizó una celda de flujo construida en el laboratorio, equipada con dos ventanas de CaF_2 , un separador de PTFE de 53 μm de espesor y control de temperatura. El espectrofotómetro se purgó con aire seco para disminuir el contenido de vapor de agua atmosférico. Los espectros se registraron en modo de adquisición de datos en doble canal en la región espectral comprendida entre 400 cm^{-1} y 4000 cm^{-1} con una resolución espectral de 2 cm^{-1} ; para su conversión se usó una función de apodización Blackman-Harris de tercer término, con un factor de sensibilidad cero igual a 4. Para el control del instrumento y registro de datos se utilizó el programa informático OPUS 7.2 (Bruker, Ettlingen, Alemania).

Las mediciones de las proteínas BSA, OVA y ConA se llevaron a cabo a temperatura constante de 25° C, promediándose un total de 64 mediciones consecutivas por espectro.

Para el análisis de cambio estructural de PLL inducido por temperatura, se promedió un total de 450 mediciones de transmisión consecutivas (tiempo de adquisición:

100 s) a cada nivel de temperatura, en un rango de temperaturas de entre 20 °C y 50 °C ($\Delta T = 2$ °C). Los espectros se registraron 240 s posteriores al cambio de temperatura, a fin de estabilizar la celda de medición, respetándose, así, la velocidad de calentamiento aplicada para las mediciones EC-QCL-IR. En caso de ser necesario, se substrajeron las bandas de absorción del vapor de agua atmosférico.

Para el análisis por MCR-ALS se registraron espectros IR por espectroscopia FT-IR en la región espectral comprendida entre 1710 cm^{-1} y 1395 cm^{-1} en el rango de temperatura de entre 10 °C y 70 °C, cada 2 °C. De esta manera, se obtuvieron matrices de datos temperatura-espectro IR (*TeIR: Temperature-InfraRed spectra*) que consistieron en 32 × 749 datos en la dimensión temporal y espectral, respectivamente.

Mediciones por espectroscopia de CD. El análisis de cambios conformacionales de PLL inducidos por temperatura se realizó por espectroscopia de dicroísmo circular (*CD*) en la región UV lejano, utilizando un espectropolarímetro Chirascan-plus CD (Applied Photophysics Ltd., Leatherhead, Surrey, Reino Unido) equipado con una unidad Peltier de control de temperatura (TC125; Quantum Northwest Inc., Liberty Lake, WA, Estados Unidos) y una cubeta de cuarzo de 1 mm de paso óptico. Se registraron los espectros cada 1.0 nm en el rango espectral comprendido entre 190.0 nm y 290.0 nm (tiempo de adquisición total = 100 s), con un velocidad de adquisición de 1 nm s^{-1} , en un rango de temperatura de entre 20 °C y 50 °C ($\Delta T = 2$ °C). Se respetó un tiempo de estabilización de 240 s para cada paso de temperatura, considerándose la velocidad de calentamiento utilizada para las mediciones FT-IR y EC-QCL-IR. De esta manera, se obtuvieron matrices de datos temperatura-espectro CD (*TeCD: Temperature-Circular Dichroism spectra*) que consistieron en 19 × 100 datos en la dimensión temporal y espectral, respectivamente, que luego se analizaron por MCR-ALS.

El control del instrumento y la adquisición de datos se realizaron utilizando el programa informático Pro-Data Chirascan-plus (Applied Photophysics Ltd., Leatherhead, Surrey, Reino Unido).

Programas informáticos y algoritmos para el análisis de datos. Para el procesamiento de las señales obtenidas por EC-QCL-IR se utilizó la rutina de procesamiento basada en el algoritmo de corrección de deformaciones COW, según se explicó en el apartado **Procesamiento de señales obtenidas por EC-QCL-IR** de la sección 3.4. Los algoritmos necesarios para operar MCR-ALS se encuentran asequibles en internet accediendo a la página web <http://www.mcrals.info/>. El proce-

samiento de datos y el análisis quimiométrico con MCR-ALS se realizó en MatLab 7.10 (MathWorks Inc., Natick, Estados Unidos).

RESULTADOS

Procesamiento de señales EC-QCL-IR. Debido a que los perfiles espectrales obtenidos para D₂O difieren a los obtenidos para agua ultrapura, fue necesario realizar un control de la rutina de procesamiento desarrollada haciendo un estudio comparativo con una técnica de referencia, como FT-IR. Es de destacar que el perfil espectral de D₂O difiere del espectro obtenido para el agua ultrapura haciéndose necesaria la evaluación en detalle de todos los parámetros utilizados para el procesamiento de las señales obtenidas. En la Figura 3.38 se muestran los espectros de transmisión EC-QCL-IR de dos mediciones consecutivas de D₂O.

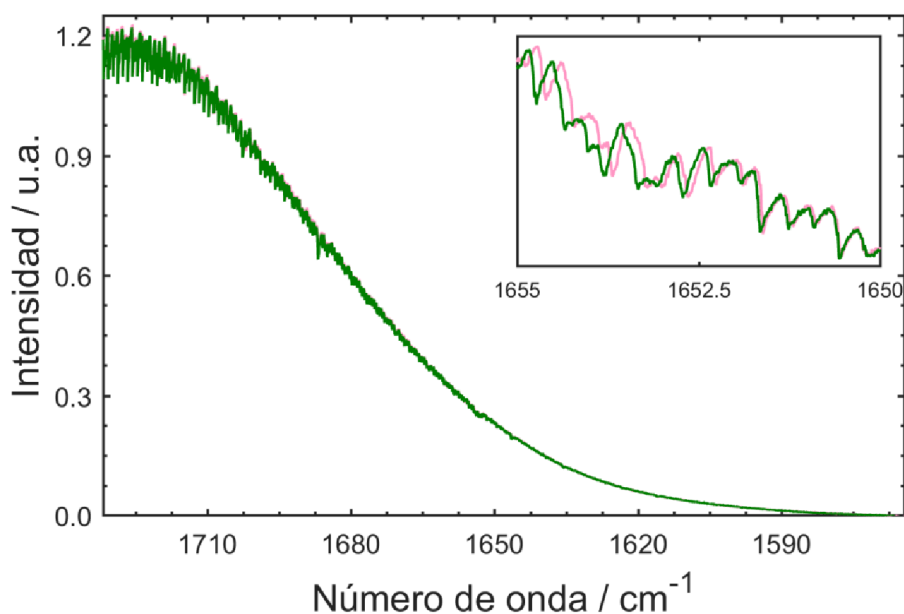


Figura 3.38. Curvas de transmisión consecutivas obtenidas por EC-QCL-IR para D₂O. En el recuadro se amplifica la región espectral comprendida entre 1655 cm⁻¹ y 1650 cm⁻¹, mostrándose las fluctuaciones presentes entre las estructuras finas de 2 líneas de transmisión consecutivas

A diferencia de las mediciones en agua, la etapa crucial para la obtención de los espectros en D₂O, además de la correspondiente a COW, fue la sustracción de la línea de base. Se observó que diferentes líneas de base en las líneas de transmisión llevan consigo distorsiones en la forma del espectro de absorbancia resultante, princi-

palmente en las regiones donde se obtienen valores mínimos de transmitancia ($< 1600 \text{ cm}^{-1}$). Para evaluar la capacidad del instrumento EC-QCL para realizar mediciones en medios distintos de agua, como D_2O , y de la rutina de procesamiento de datos desarrollada, se analizaron los espectros de 3 proteínas con estructura secundaria diferentes por espectroscopia EC-QCL-IR y FT-IR, en idénticas condiciones. Se seleccionaron proteínas modelos que exhiben diferencias en sus estructuras secundarias y, por consiguiente, que muestran diferentes patrones espectrales: **1- OVA**: proteína que contiene hélices α y láminas β en iguales proporciones (Dong *et al.*, 2000; Oshokoya *et al.*, 2014); **2- ConA**: proteína conformada mayoritariamente por láminas β ; y **3- BSA**: proteína compuesta predominantemente por estructuras de hélices α (Barth, 2007; Murayama *et al.*, 2004). En la Figura 3.39 se muestran los espectros de absorbancia IR de OVA, ConA y BSA en D_2O usando EC-QCL-IR y FT-IR.

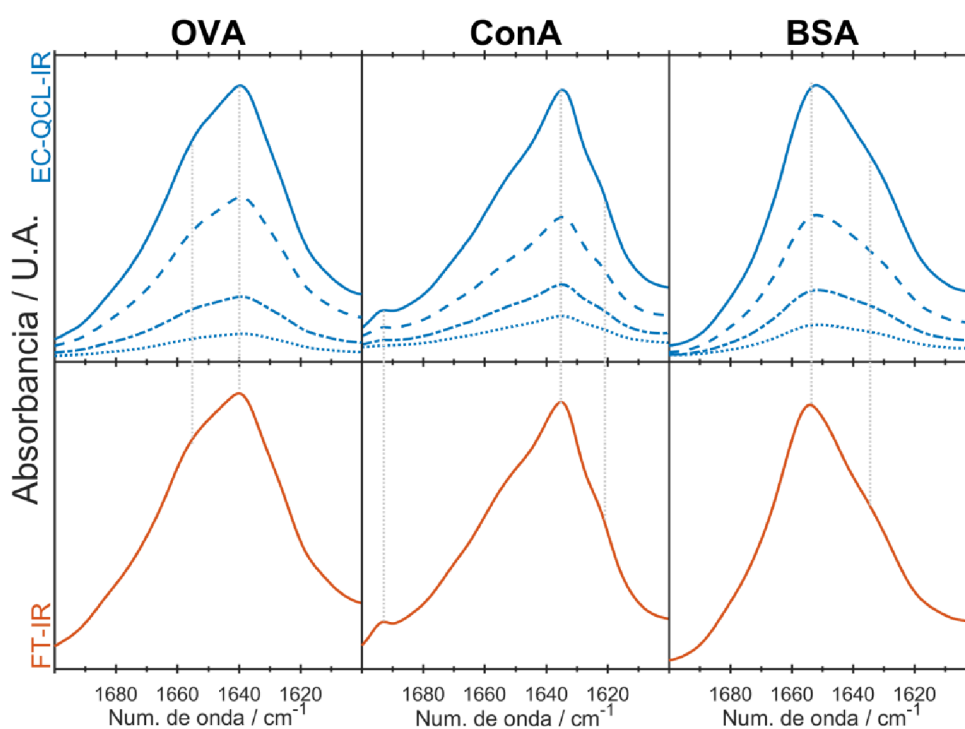


Figura 3.39. (Arriba) Espectros de absorbancia IR obtenidos por EC-QCL-IR de soluciones proteicas conteniendo 0.25 mg mL^{-1} (línea de puntos azul), 0.50 mg mL^{-1} (línea de rayas y puntos azul), 1.00 mg mL^{-1} (línea de rayas azul) y 2.50 mg mL^{-1} (línea continua azul) y (abajo) obtenidos por FT-IR (línea continua roja) para soluciones proteicas conteniendo 5.00 mg mL^{-1} . Las líneas de punto grises remarcan la congruencia entre las formas espectrales de los espectros obtenidos con los diferentes instrumentos

En todos los casos, se pudieron identificar características espectrales de las estructuras secundarias correspondientes a cada proteína en concentraciones proteicas tan bajas como 0.25 mg mL^{-1} . Para cuantificar el grado de concordancia entre espec-

tros se calculó el grado de solapamiento espectral (s_{12}) entre los espectros de FT-IR (s_1) y EC-QCL-IR (s_2) utilizando la ecuación 3.1. Los valores s_{12} obtenidos resultaron mayores a 0.9930, quedándose demostrado que los espectros adquiridos por EC-QCL-IR contienen información equivalente a los espectros registrados por FT-IR.

Transición conformacional $\alpha \rightarrow \beta$ de PLL inducida por temperatura. A presión atmosférica, bajas temperaturas ($> 25^\circ\text{C}$) y pH menores a 9.0, la repulsión entre cargas de los grupos lisina de PLL, que se encuentran protonados, inhibe la formación de estructuras ordenadas, como hélices α o láminas β , favoreciéndose la estructura polipeptídica de configuración desordenada (Carrier *et al.*, 1990). Por el contrario, a valores de pH superiores a su pK_a ($pK_a = 10.5$), la cadena lateral de lisina se desprotona y el péptido adopta la conformación ordenada de hélices α a bajas temperaturas. En idénticas condiciones, ante un incremento de temperatura ($> 30^\circ\text{C}$), las estructuras helicoidales α de PLL comienzan a desordenarse hasta adoptar una conformación ordenada de láminas β (Dzwolak *et al.*, 2004; Dzwolak *et al.*, 2005; Grigsby *et al.*, 2002; Szyc *et al.*, 2008). Adicionalmente, se ha demostrado que la velocidad de formación de láminas β inducidas por calentamiento aumenta con el incremento de pH y el aumento de la concentración peptídica (Davidson *et al.*, 1967).

En este trabajo se emplearon tres técnicas analíticas diferentes, CD, EC-QCL-IR y FT-IR, para evaluar la dependencia de la transición $\alpha \rightarrow \beta$ de PLL con la concentración en un rango amplio de concentraciones, comprendido entre 0.25 mg mL^{-1} y 10.00 mg mL^{-1} . Para analizar el comportamiento de la transición estructural de PLL ante el incremento de temperatura en todo rango de concentraciones se requirieron 3 técnicas espectroscópicas diferentes: **1-** *dicroísmo circular (CD) para evaluar el rango comprendido entre 0.25 mg mL^{-1} y 2.00 mg mL^{-1} ; 2-* *EC-QCL-IR para evaluar el rango comprendido entre 0.25 mg mL^{-1} y 6.00 mg mL^{-1} ; y 3-* *FT-IR para evaluar el rango comprendido entre 5.00 mg mL^{-1} y 10.00 mg mL^{-1} . De esta manera, no solo se cubrió un rango amplio de concentraciones, sino también se pudo evaluar la congruencia de los resultados obtenidos mediante diferentes técnicas espectroscópicas.*

La Figura 3.40 muestra los espectros de CD obtenidos para la transición conformacional inducida por temperatura de una solución de PLL a pD 12.4. A bajas temperaturas, la señal CD muestra dos picos con elipticidad ($[\theta]$) negativa centrados a 208 nm y 222 nm, representativos para macromoléculas con estructura secundarias de hélices α . Ante el incremento de la temperatura, se observa una pérdida de $[\theta]$ negati-

va de los espectros a las longitudes de onda mencionadas, mostrándose, a altas temperaturas, espectros de CD con una sola banda negativa centrada en 215 nm y con la forma espectral característica para las estructuras laminares β (Greenfield, 2007b). La evolución del cambio conformacional inducido por temperatura del homo-polipéptido se evaluó mediante el cambio de $[\theta]$ a 222 nm, longitud de onda que es particularmente sensible a los cambios en la estructura secundaria (Kelly *et al.*, 2005). Los puntos obtenidos del gráfico de $[\theta]$ respecto a la temperatura se ajustaron a una función sigmoidea de Boltzmann, estimándose el punto de inflexión, que es considerado como la temperatura de transición $\alpha \rightarrow \beta$ (recuadro de la Figura 3.40) del homo-polipéptido.

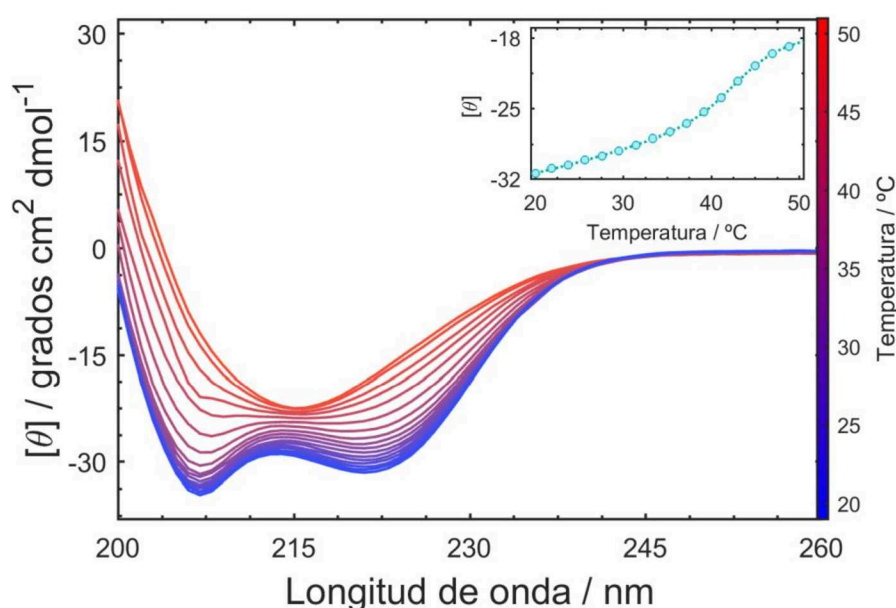


Figura 3.40. Espectros CD obtenidos en el rango UV lejano para una solución de 0.25 mg mL⁻¹ de PLL en PBSD pD 12.4, registrados en el rango de temperaturas de entre 20 °C y 50 °C, cada 2 °C. A la derecha se indica en escala de colores la transición conformacional α (azul) \rightarrow β (rojo) inducida por temperatura. En el recuadro se muestra la progresión de la elipticidad registrada a 222 nm en dependencia de la temperatura. La línea de puntos muestra la curva sigmoidea ajustada usando una función de Boltzmann.

En la Figura 3.41 se muestran los espectros IR obtenidos para la transición $\alpha \rightarrow \beta$ de PLL inducida por temperatura adquiridos por EC-QCL-IR. El espectro registrado a 20 °C muestra una máxima de banda a 1635 cm⁻¹, característico para las estructuras helicoidales α adoptadas por PLL en D₂O (Dzwolak *et al.*, 2005). En comparación con las bandas de absorción típicas obtenidas para proteínas globulares con configuración de hélices α (aproximadamente a 1652 cm⁻¹ en D₂O), como por ejemplo BSA, la posición espectral de la conformación helicoidal α de PLL se encuentra desplazada hacia números de onda menores. Este cambio en el máximo de banda ha sido explicado mediante el acoplamiento intra- e inter-molecular no perturbado causado por el alto grado

de regularidad de las estructuras helicoidales de un homo-polipéptido (Jackson *et al.*, 1989). Adicionalmente, la posición real de la banda de absorción IR de PLL depende de la longitud de la cadena peptídica (Dzwolak *et al.*, 2004; Jackson *et al.*, 1989).

Ante el aumento de la temperatura se observa que la banda atribuida a la configuración de hélices α decrece al mismo tiempo que aparecen 2 bandas posicionadas a 1611 cm^{-1} y 1681 cm^{-1} . Este patrón espectral emergente que contiene bandas espectrales de alta y baja frecuencia ha sido atribuido a conformaciones de láminas β anti-paralelas adoptadas por PLL en D_2O (Carrier *et al.*, 1990; Dzwolak *et al.*, 2005; Zhang *et al.*, 2011). Este mismo comportamiento se observó para las mediciones por FT-IR de soluciones con alta concentración de PLL.

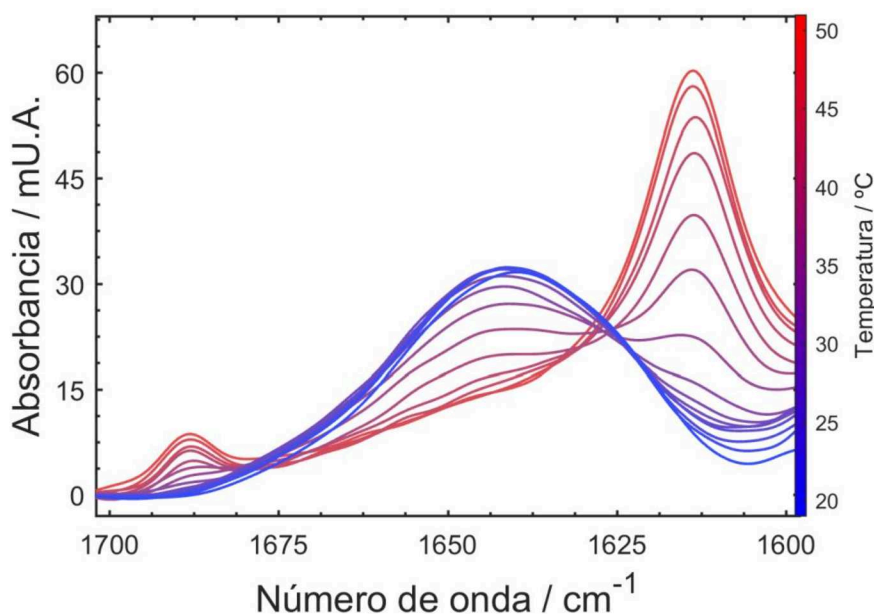


Figura 3.41. Espectros IR obtenidos para una solución conteniendo 0.50 mg mL^{-1} de PLL en PBSD pD 12.4, registrados en función de la temperatura por espectroscopia EC-QCL-IR. A la derecha se indica en escala de colores la transición conformacional α (azul) $\rightarrow\beta$ (rojo) inducida por temperatura

Finalmente, con el objetivo de analizar el progreso de la transición conformacional a distintas temperaturas en diferentes niveles de concentración, se evaluó la absorbancia de la banda espectral de baja frecuencia, localizada a 1611 cm^{-1} , correspondiente a la conformación de láminas β . Es importante mencionar que, debido a que la velocidad de calentamiento afecta la transición $\alpha\rightarrow\beta$ en función de la temperatura (Dzwolak *et al.*, 2005), se respetó la misma velocidad de calentamiento en todas las mediciones. Los puntos obtenidos del gráfico de absorbancia respecto a la temperatura se ajustaron a una función sigmoidea de Boltzmann, cuyo punto de inflexión representa la

temperatura de transición $\alpha \rightarrow \beta$ del homo-polipéptido. La Figura 3.42 muestra las curvas de transición obtenidas por espectroscopia EC-QCL-IR y FT-IR para diferentes concentraciones de PLL respecto a la temperatura. Se puede observar que las temperaturas de transición mayores corresponden a los menores niveles de concentración peptídica, indicándose la baja propensión a la formación de láminas β ante el aumento de la temperatura para bajas concentraciones de homo-polipéptido.

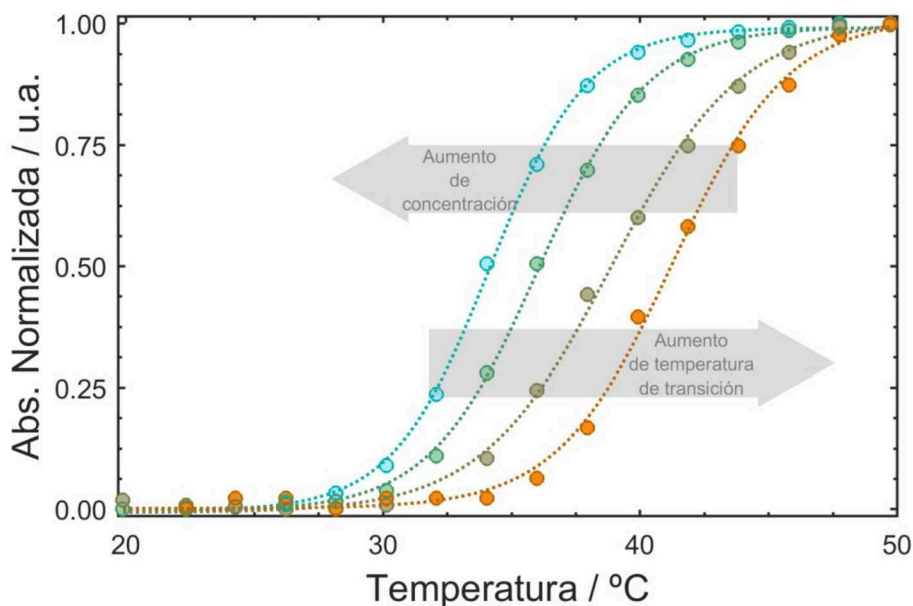


Figura 3.42. Progresión del cambio de absorbancia IR relativa obtenida por evaluación de la altura de la banda localizada a 1611 cm^{-1} para distintos niveles de concentración. Para soluciones conteniendo 0.50 mg mL^{-1} (círculos anaranjados) y 3.00 mg mL^{-1} (círculos verde oscuros) se analizaron los espectros obtenidos por EC-QCL-IR, mientras que para las soluciones conteniendo 6.50 mg mL^{-1} (círculos verde claros) y 9.00 mg mL^{-1} (círculos celestes) se utilizaron los espectros obtenidos por FT-IR. Las líneas de puntos representan la curva sigmoidea ajustada a una función de Boltzmann

Finalmente, se analizó el comportamiento de la temperatura de transición $\alpha \rightarrow \beta$ con respecto a la concentración de PLL en todo el rango de concentraciones, evaluándose los datos obtenidos por las 3 técnicas utilizadas. En la Figura 3.43 se muestran las temperaturas de transición $\alpha \rightarrow \beta$ obtenidas por CD, EC-QCL-IR y FT-IR para cada nivel de concentración evaluado. El comportamiento observado asegura la consistencia de los resultados obtenidos utilizando técnicas espectroscópicas diferentes y revela la dependencia de la temperatura de transición $\alpha \rightarrow \beta$ con la concentración peptídica, mostrándose un decaimiento exponencial en función de la concentración de homo-péptido. Esta dependencia de la temperatura de transición con la concentración también ha sido demostrada en numerosos sistemas proteicos reportados anteriormente (Qi *et al.*, 1995; Santiago *et al.*, 2010; Shi *et al.*, 2013), si bien para PLL, hasta el momento, no se han

reportado investigaciones en una forma sistemática en tan amplio rango de concentraciones.

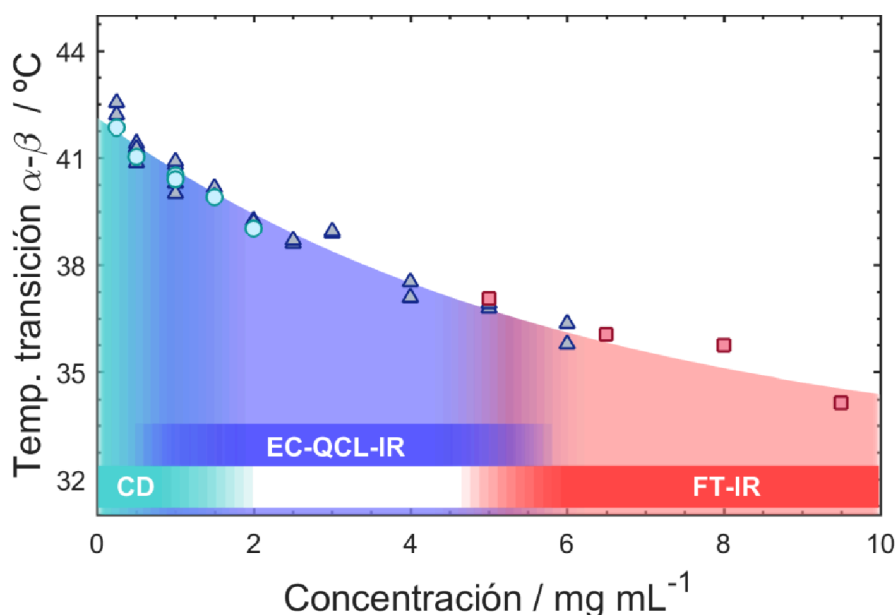


Figura 3.43. Temperaturas de transición correspondientes al cambio conformacional $\alpha \rightarrow \beta$ inducido por temperatura determinados por espectroscopia CD (círculos celestes), EC-QCL-IR (triángulos azules) y FT-IR (cuadrados rojos). Las bandas coloreadas representan la región de concentración cubierta por CD (celeste), EC-QCL-IR (azul) y FT-IR (rojo)

Así, con el sistema modelo propuesto, queda demostrada la necesidad de desarrollar técnicas experimentales que permitan el estudio conformacional de proteínas en bajos niveles de concentración, pero que, además, permitan el análisis en un amplio rango de concentraciones, a fin de asegurar la comparabilidad de los resultados obtenidos por diferentes técnicas y reducir ambigüedades o imprecisiones en su interpretación.

Modelado de datos por MCR-ALS. En primer lugar se planteó la posibilidad de utilizar un modelado de datos conjunto por MCR-ALS para un mismo experimento monitoreado mediante diferentes técnicas espectroscópicas, es decir, utilizar MCR-ALS en su versión extendida, construyendo matrices aumentadas en filas, según se reporta en la literatura para el análisis de intermediarios proteicos (Navea *et al.*, 2002; Navea *et al.*, 2006). Sin embargo, no fue posible llevar a cabo este modelado debido a incompatibilidades encontradas en los resultados obtenidos por las diferentes técnicas. Por un lado, la dependencia de la transición conformacional con la concentración demanda obtener experimentos por IR y CD al mismo nivel de concentración, lo cual no

resultó posible para FT-IR debido a la baja sensibilidad de la técnica. Si bien mediante EC-QCL-IR es posible llegar a niveles de concentración tan bajos como los necesarios para CD, los espectros IR obtenidos a altas temperaturas indican que la configuración laminar β presenta dos perfiles espectrales diferentes, a diferencia de CD que sólo necesita un componente para explicar la misma configuración estructural. Por lo tanto, un modelado conjunto entre CD e IR no favorecería al modelado quimiométrico, sino que llevaría a obtener resultados erróneos. En consecuencia, se llevó a cabo el modelado por MCR-ALS para datos de CD y FT-IR, individualmente. El modelado de los datos por MCR-ALS se realizó a fin de obtener los perfiles espectrales de los componentes estructurales puros involucrados en el proceso de transición térmica para cada técnica espectroscópica y poder combinar los resultados a fin de concluir correctamente acerca de las configuraciones presentes en el proceso térmico.

Modelado de datos de CD. Dicroísmo circular es una técnica espectroscópica que permite, al igual que IR, realizar análisis estructurales de macromoléculas a partir de la evaluación de los espectros. Adicionalmente, la obtención de perfiles espectrales por MCR-ALS permite obtener información relevante referida a los procesos configuraciones que ocurren en la transición $\alpha \rightarrow \beta$ inducida por temperatura de PLL.

Para el modelado de los datos obtenidos de los experimentos de CD, se utilizó la matriz de datos TeCD obtenida para una solución conteniendo 0.25 mg mL^{-1} de PLL a pD 12.4 en el rango espectral de entre 200 nm y 260 nm. El número de componentes espectroscópicamente activos se determinó mediante SVD, indicando que fueron necesarios 3 factores para explicar la varianza del sistema. Para la resolución por MCR-ALS se utilizaron estimaciones iniciales obtenidas por EFA (Maeder *et al.*, 2009). Es importante mencionar que para este tipo de datos no fue posible utilizar metodologías basadas en SIMPLISMA debido a que estos algoritmos no permiten el análisis de datos que presenten valores negativos (Bogomolov *et al.*, 2005; Windig, 2009), como es el caso de los espectros obtenidos por CD. Posteriormente, para la optimización iterativa por ALS se aplicaron las restricciones de unimodalidad, no-negatividad y condición de sistema cerrado solo al modo correspondiente a la progresión conformacional en función de la temperatura, no aplicándose restricciones al modo espectral. De la descomposición bilineal, se obtuvieron los perfiles espectrales y de progresión conformacional correspondientes a las distintas estructuras secundarias adoptadas por el homo-polipéptido en la transición estructural $\alpha \rightarrow \beta$ inducida por temperatura.

Los resultados obtenidos por MCR-ALS se muestran en la Figura 3.44, confir-

mándose la existencia de 3 configuraciones estructurales principales adoptadas por el homo-péptido durante la transición $\alpha \rightarrow \beta$ inducida por temperatura. El perfil de progresión conformacional muestra la evolución de las distintas estructuras secundarias adoptadas por PLL en el proceso térmico.

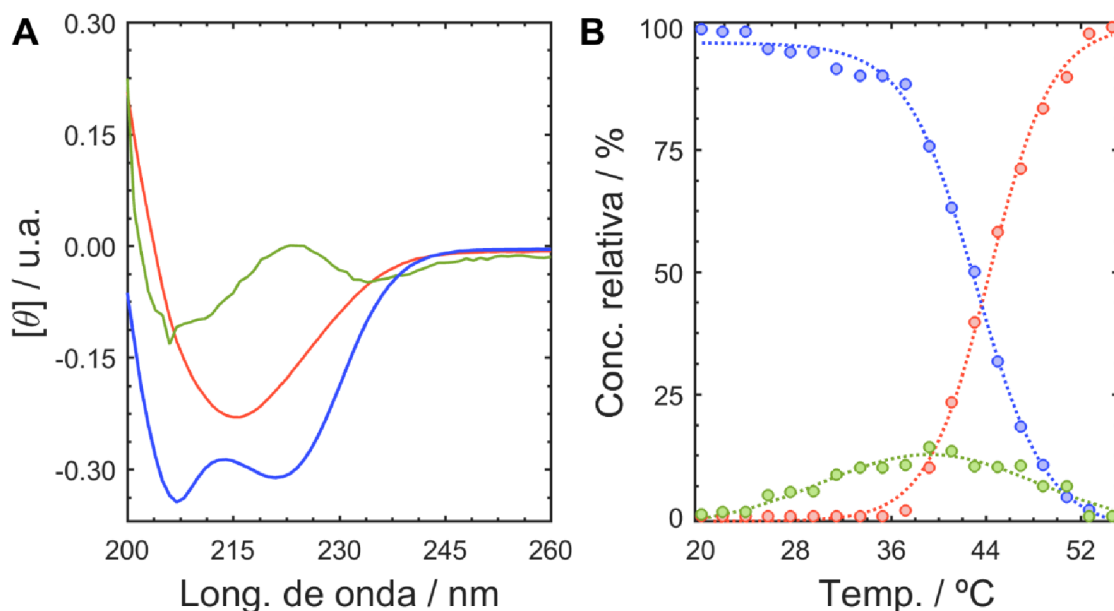


Figura 3.44. Perfiles espectrales (A) y de progresión conformacional dependientes de la temperatura (B) correspondientes a la estructura secundaria de hélices α (azul), láminas β (rojo) y estructuras desordenadas (verde)

A bajas temperaturas, la conformación adoptada exhibe bandas características de péptidos o proteínas con configuración predominantemente compuesta por hélices α , demostrada a través de las bandas negativas localizadas a 208 nm y 222 nm (Czarnik-Matusewicz *et al.*, 2006; Greenfield, 2007a; b). A medida que la temperatura aumenta por encima de los 30 °C se presenta una nueva configuración estructural que muestra una banda negativa a aproximadamente 215 nm, que puede ser asociada, principalmente, a estructuras secundarias ricas en láminas β (Greenfield, 2007b), evolucionando hasta el final del experimento, siendo la estructura predominante a temperaturas por encima de 44 °C. Adicionalmente, es posible identificar un componente intermedio que evoluciona positivamente hasta alcanzar un valor máximo a aproximadamente 40 °C y que luego decrece hasta alcanzar un valor mínimo al final del experimento. Este componente, que muestra elipticidad positiva a 220 nm, puede ser asociado a conformaciones extendidas, estructuras aleatorias o desordenadas (Czarnik-Matusewicz *et al.*, 2006; Greenfield, 2007b). Estas observaciones explicarían que el re-arreglo del péptido desde una estructura helicoidal α hacia una laminar β no es de forma directa, sino que se

produce a través de un estado intermediario de configuración extendida, el que facilitaría la transición estructural por minimización energética (Czarnik-Matusewicz *et al.*, 2006). Cabe mencionar que experimentos realizados para distintas concentraciones de PLL en un rango de temperatura de hasta 90 °C arrojaron idénticos resultados a los explicados anteriormente, demostrándose que el comportamiento mostrado por PLL por CD ante el aumento de temperatura solo depende de la concentración y no de la temperatura alcanzada al final del experimento.

Modelado de datos de FT-IR. Con el objetivo de proveer mayor información al sistema y considerando que FT-IR permite registrar también cambios espectrales ocasionados por el incremento de la temperatura en la región correspondiente a la banda amida II', el análisis por MCR-ALS se realizó utilizando los datos obtenidos en la región espectral de la banda amida I' y amida II'. En la Figura 3.45 se muestran los espectros registrados en la región correspondiente a la banda amida I' y amida II' para el cambio estructural de PLL inducido por temperatura.

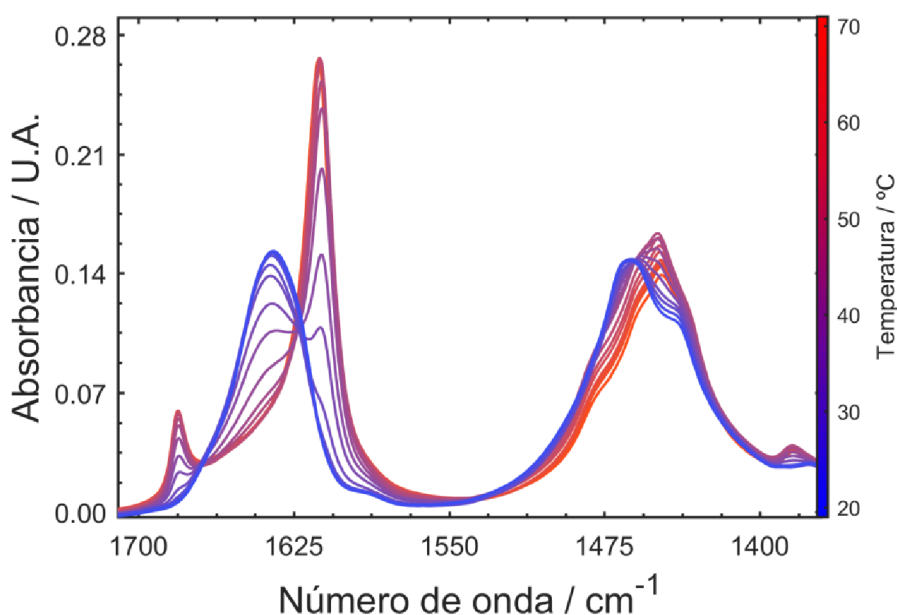


Figura 3.45. Espectros IR obtenidos para una solución conteniendo 18.00 mg mL⁻¹ de PLL en PBSD pD 12.4 registrados en función de la temperatura por espectroscopia FT-IR, mostrándose la banda amida I' y amida II'. A la derecha se indica en escala de colores la transición conformacional α (azul) \rightarrow β (rojo) inducida por temperatura

Se observó que el incremento de la temperatura por encima de la temperatura de transición tiene un efecto notable sobre la formación de nuevas configuraciones de láminas β . Por lo tanto, teniendo en cuenta que al aumentar la concentración de PLL la temperatura de transición disminuye, se modelaron los datos obtenidos para una alta

concentración de PLL, 18.00 mg mL⁻¹. Adicionalmente, a fin de obtener la mayor contribución de las diferentes configuraciones estructurales presentes a temperaturas menores y mayores a la temperatura de transición (< 35 °C, para concentraciones mayores a 10.00 mg mL⁻¹) se evaluó el rango de temperaturas comprendido entre 10 °C y 70 °C.

Previamente a la descomposición bilineal, se determinó el número de componentes espectroscópicamente activos utilizando SVD, necesitándose 4 componentes para explicar la varianza del sistema. Para la inicialización de MCR-ALS se utilizaron estimaciones iniciales espectrales obtenidas por una metodología basada en SIMPLISMA (Windig *et al.*, 1991). El análisis quimiométrico se realizó mediante MCR-ALS aplicándose restricciones asociadas a la naturaleza del sistema en estudio, es decir, unimodalidad, no-negatividad y condición de sistema cerrado en el modo correspondiente a la progresión de la concentración en función de la temperatura y no-negatividad al modo espectral. La evolución de las estructuras secundarias de PLL ante el incremento de temperatura y los perfiles espectrales asociados a dichas estructuras, ambos obtenidos de la descomposición bilineal, se muestran en la Figura 3.46.

Al inicio de la etapa de calentamiento, la conformación adoptada por el péptido exhibe una banda amida I' asimétrica con un máximo en 1635 cm⁻¹ y una banda amida II' que muestra una estructura con 2 máximos bien definidos (1461 cm⁻¹ y 1439 cm⁻¹), correspondiéndose a una configuración secundaria de hélices α para polipéptidos en D₂O (DeFlores *et al.*, 2009; Jackson *et al.*, 1989). El máximo de intensidad de la banda amida I' encontrada para PLL con configuración de hélice α (1635 cm⁻¹) se encuentra ubicada a frecuencias menores a los correspondientes para proteínas globulares, sugiriendo la presencia de estructuras helicoidales atípicas, como por ejemplo, hélices tipo 3₁₀ o hélices α_{II} (Jackson *et al.*, 1989). Ante el incremento de temperatura, el surgimiento de una banda amida I' ubicada a 1645 cm⁻¹, acompañada de un hombro localizado a ~1665 cm⁻¹ y una pérdida de estructura de la banda amida II', sugiere que PLL adopta una configuración aleatoria que permite la transición desde una estructura ordenada de hélices α hacia una estructura ordenada de láminas β (DeFlores *et al.*, 2009). Cuando la temperatura supera los 30 °C, la conformación predominante del homo-polipéptido presenta una banda amida I' dividida en modos de alta (1610 cm⁻¹) y baja (1680 cm⁻¹) intensidad representativos de láminas β (Jackson *et al.*, 1989; Yang *et al.*, 2015), mientras que la banda amida II', que presenta un máximo a 1450 cm⁻¹, también muestra una drástica modificación ante el cambio estructural del homo-polipéptido.

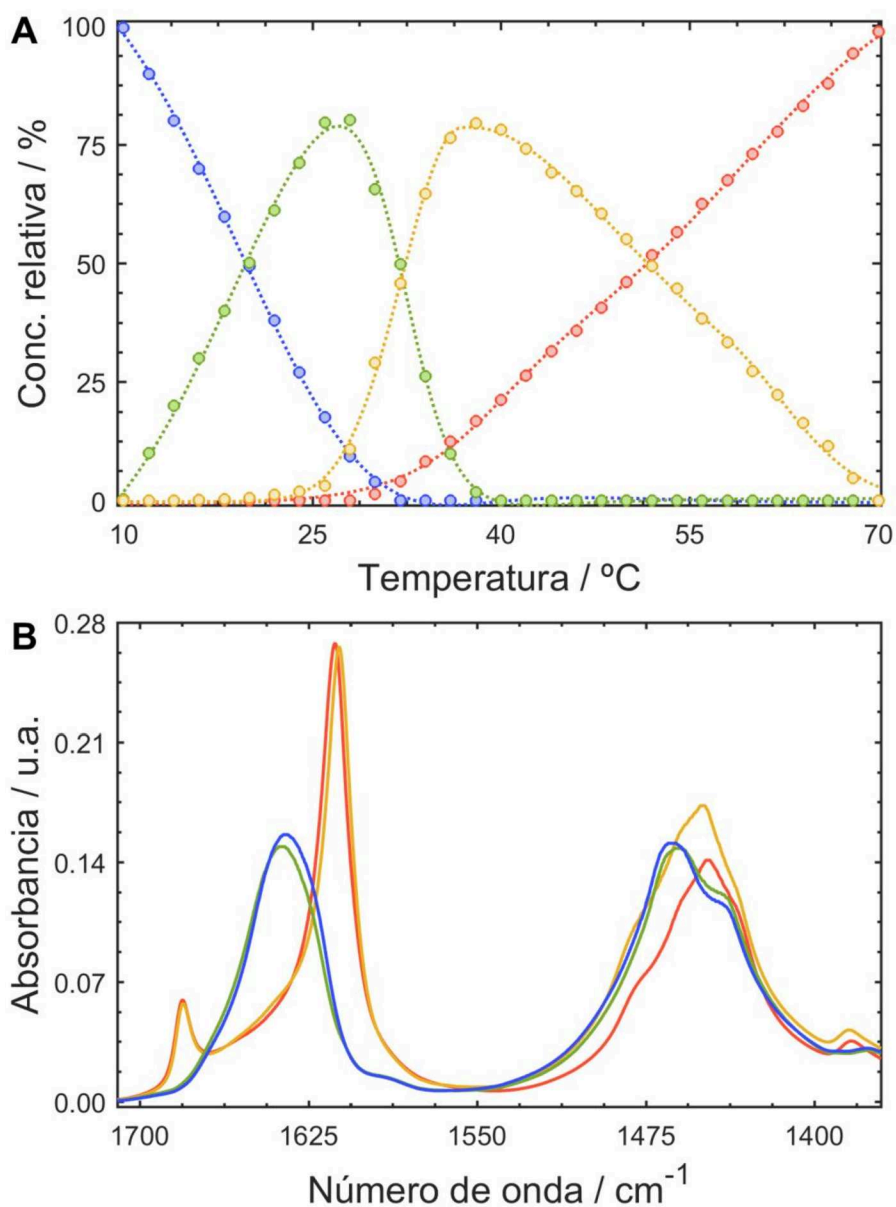


Figura 3.46. Perfiles espectrales (A) y de progresión conformacional dependiente de la temperatura (B) correspondientes a la estructura secundaria de helices α (azul), estructuras desordenadas (verde), láminas β_1 (marrón) y láminas β_2 (rojo)

Estas observaciones muestran una alta concordancia con los resultados obtenidos por CD, si bien se muestran variaciones en las temperaturas a las que se producen los cambios estructurales, lo que es de esperarse considerando la dependencia de la temperatura de transición con la concentración de PLL. Sin embargo, la mayor diferencia encontrada entre CD y FT-IR es la aparición de una nueva estructura de láminas β a altas temperaturas que es visualizada en los espectros IR. A temperaturas superiores a los 40 °C se observa un corrimiento de los máximos de intensidad de la banda amida I' hacia número de ondas mayores, siendo la banda de mayor intensidad la más afectada. La banda amida II', además de mostrar un decaimiento de la señal,

sufre un corrimiento del máximo de intensidad hacia número de ondas menores, mostrando un hombro bien definido a 1480 cm^{-1} . Estos cambios en las bandas amida I' y amida II' no han sido estudiados hasta el momento, si bien un estudio realizado para PLL utilizando espectroscopia Raman de resonancia ultravioleta (UVRM) demuestra la presencia de intermediarios de láminas β durante la transición conformacional estimulada por cambios en el contenido de agua de un film de PLL deshidratado (Müller *et al.*, 1996). Se observó, además, que los corrimientos en los máximos de intensidad IR son invariables en el tiempo y solo dependen de la temperatura.

Los corrimientos de las bandas correspondientes a las estructuras secundarias de láminas β han sido, generalmente, atribuidos a re-arreglos de su organización interna, es decir, desde una estructura menos densa a una estructura de láminas β más compacta (Ioannou *et al.*, 2015; Polzi *et al.*, 2012; Sarroukh *et al.*, 2013). Adicionalmente, se ha comprobado que ante el incremento de la temperatura, los puentes de hidrógeno del solvente de las cadenas β son reemplazados por puentes de hidrógeno inter-cadena, conllevando a la formación de las estructuras de lámina β final (Jiji *et al.*, 2006; Szyk *et al.*, 2008).

La transición conformacional inducida térmicamente para PLL a valores elevados de pH (> 10.5) involucra una conformación intermedia atribuida a estructuras de hélices extendidas (Dzwolak *et al.*, 2004; Szyk *et al.*, 2008) o conformaciones aleatorias (Davidson *et al.*, 1967) que tienen lugar a temperaturas inferiores a $40\text{ }^{\circ}\text{C}$. Este comportamiento ha sido también demostrado para los datos obtenidos por CD e indicarían que la transición $\alpha \rightarrow \beta$ se lleva a cabo mediante pérdida de configuraciones ordenadas. En una primera etapa, ante el aumento de la temperatura, se rompen los puentes hidrógenos intra-moleculares que estabilizan la configuración de hélices α adoptándose irreversiblemente una estructura de configuración extendida o aleatoria. Adicionalmente, el aumento de la energía térmica facilita las interacciones hidrofóbicas entre los grupos lisina laterales estimulándose la formación irreversible de láminas β (Dzwolak *et al.*, 2005; Immaneni *et al.*, 1998). Este aumento de energía provee la entalpía positiva necesaria para romper los puentes hidrógenos y favorecer a la asociación de los residuos lisil que forman la estructura de láminas β (Davidson *et al.*, 1967). La gran intensidad de interacciones hidrofóbicas presentes a elevadas temperaturas es la base principal del incremento en la propensión de la formación de láminas β .

CONCLUSIÓN

Se desarrolló una metodología analítica basada en mediciones de transmisión IR laser para investigaciones de estructuras secundarias de macromoléculas en soluciones deuteradas. Para demostrar el potencial de la metodología desarrollada, se monitorearon los cambios conformacionales en la estructura secundaria de un homopolipéptido en diferentes niveles de concentración inducidos por temperatura.

Las mediciones EC-QCL-IR a muy bajas concentraciones acopladas a un correcto procesamiento de datos, permitieron la combinación de múltiples técnicas espectroscópicas, utilizadas, generalmente, para el análisis conformacional de macromoléculas en solución. Esto representa una gran ventaja para la correcta interpretación del comportamiento de los sistemas biológicos, siendo particularmente representativa en sistemas que muestran efectos dependientes de la concentración, donde el estudio en un amplio rango de concentraciones resulta de gran interés. Se analizaron los cambios conformacionales de PLL inducidos por temperatura en un rango de concentraciones amplio utilizando CD, EC-QCL-IR y FT-IR. Los resultados muestran gran concordancia entre las técnicas usadas, revelándose la presencia de una dependencia entre la temperatura de transición $\alpha \rightarrow \beta$ con la concentración del homopolipéptido.

Aprovechando la ventaja de MCR-ALS de obtener perfiles con interpretación física, se realizó un análisis detallado de la transición conformacional ocurrida en PLL inducida por temperatura, brindando información acerca de las conformaciones estructurales adoptadas por el péptido durante todo el proceso de calentamiento. El análisis previo del comportamiento del sistema evitó realizar un modelado quimiométrico combinado que hubiese resultado incorrecto frente a la subestimación de los efectos de la concentración del péptido sobre la temperatura de transición. El análisis individual por MCR-ALS de los experimentos de CD y FT-IR permitió identificar, fehacientemente, las configuraciones principales adoptadas por el péptido en todo el proceso térmico. Asimismo, la combinación de FT-IR con MCR-ALS permitió resolver una nueva configuración estructural de láminas β presente en PLL a altas temperaturas, la que, hasta el momento, no ha sido identificada o investigada en detalle.

Estos resultados muestran un gran avance para estudios que involucran muestras conteniendo bajas concentraciones de macromoléculas, por ejemplo, fluidos biológicos, o investigaciones conformacionales donde bajas concentraciones proteicas son de gran relevancia, por ejemplo, estudios en vacunas o proteínas terapéuticas de bajas solubilidad (Dong *et al.*, 2006).

3. 6. OBTENCIÓN DE DATOS DE TERCER ORDEN UTILIZANDO CROMATOGRAFÍA LÍQUIDA DE ALTO RENDIMIENTO ACOPLADA A DETECCIÓN CON MATRICES EXCITACIÓN-EMISIÓN DE FLUORESCENCIA. ANÁLISIS Y RESOLUCIÓN DE DATOS

Modelling four and three-way fast high-performance liquid chromatography with fluorescence detection data for quantitation of fluoroquinolones in water samples – Alcaráz, M.R., et al., Anal Chim Acta 2014, 809, 37–46

A new modelling strategy for third-order fast high-performance liquid chromatographic data with fluorescence detection. Quantitation of fluoroquinolones in water samples – Alcaráz, M.R., et al., Anal Bioanal Chem 2015, 407, 1999–2011

CONSIDERACIONES GENERALES

En los últimos años, se ha observado un notable aumento en el número de aplicaciones quimiométricas en el área de la química analítica. La potencialidad demostrada por la combinación de estas áreas ha llevado consigo un incansable interés por el desarrollo de nuevas metodologías analíticas, el desarrollo de algoritmos quimiométricos versátiles, pero o su vez eficientes para el análisis de datos, como también, del interés por investigar las ventajas y bondades de los datos multidimensionales. En trabajos publicados recientemente, se ha demostrado que el incremento en el número de dimensiones en los datos multi-vía impacta positivamente en las propiedades analíticas del método, traduciéndose en una mejora de las cifras de mérito, en virtud de, principalmente, el aumento de sensibilidad y la mejora en la selectividad en un sistema multi-componente (Bauza *et al.*, 2012; Olivieri, 2014).

En lo que respecta a calibración multivariada, los datos de primer y segundo orden han sido extensamente estudiados, siendo posible encontrar un sinnúmero de aplicaciones analíticas para la determinación de una amplia variedad de compuestos utilizando diferentes equipamientos o metodologías analíticas. Las aplicaciones basadas en cromatografía líquida de alto rendimiento acoplada a detección espectral de

absorción o de emisión de fluorescencia con modelado de datos de segundo orden han demostrado ser altamente eficientes para el análisis de muestras complejas en presencia de múltiples analitos, evitándose etapas de pre-procesamiento o tratamiento de muestras, aprovechándose y explotándose satisfactoriamente la ventaja de segundo orden (Muñoz de la Peña *et al.*, 2015; Olivieri *et al.*, 2014b).

Por otro lado, si bien aún se continúan investigando las ventajas analíticas adicionales a la ventaja de segundo orden, el análisis de datos de orden superior para aplicaciones analíticas constituye un campo que merece el esfuerzo de ser explorado. Algunos autores plantean las posibles ventajas adicionales que se presentarían en la calibración multivariada de orden superior, que, a pesar de que no se ha logrado un consenso en la comunidad científica, las consideran como ventaja de tercer orden (Wu *et al.*, 2015). Estas propiedades se caracterizan por: **1- aumento de la sensibilidad y la selectividad;** **2- resolución de problemas de co-linealidad;** y **3- posibilidad de resolver un problema complejo en un arreglo de datos correspondiente a una sola muestra, en independencia a muestras adicionales.**

Actualmente, y con el avance de la tecnología en calidad de instrumentación analítica, se ha simplificado la obtención de datos multidimensionales. Sin embargo, es importante tener en cuenta las características de los datos generados en dependencia al instrumental utilizado. Bajo este punto de vista, la estrategia experimental seleccionada para la generación de los datos multidimensionales o datos multi-vía podría tener un efecto significativo en los resultados finales y, principalmente, sobre la resolución quimiométrica del problema.

En este capítulo se analizan 3 estrategias o metodologías analíticas utilizando diferentes arreglos instrumentales para la generación de datos de tercer orden, basados en un procedimiento cromatográfico acoplado a la detección de matrices de excitación-emisión de fluorescencia. La primera metodología propuesta, *Metodología I*, se basa en la recolección de fracciones al final del procedimiento cromatográfico con posterior detección de matrices de excitación-emisión para cada fracción recolectada. Esta metodología ha sido propuesta con anterioridad por Bro *R.* en 1997 en su trabajo de tesis (Bro, 1997), la que solo ha sido desarrollada solo de manera teórica. La *Metodología II* se propone siguiendo el procedimiento reportado por Lozano *et al.* (Lozano *et al.*, 2013), registrándose matrices tiempo de retención-espectro de emisión de fluorescencia para varias inyecciones cromatográficas, utilizando longitudes de onda de excitación diferente para cada inyección de una misma muestra. Finalmente, se expone una estrategia novedosa, *Metodología III*, que permite el registro de matrices de excitación-emisión

de fluorescencia en línea al procedimiento cromatográfico.

Se analizan y comparan las características instrumentales, como también las características y propiedades matemáticas de los datos generados. Adicionalmente, se propone un método analítico modelo con modelado de datos con propósitos cuantitativos llevado a cabo con una de las metodologías propuestas.

MATERIALES Y MÉTODOS

Reactivos. Todos los estándares utilizados fueron de calidad grado analítico. Las drogas sólidas de enoxacina (*ENO*) y ofloxacina (*OFL*) fueron provistas por Sigma-Aldrich (Steinheim, Alemania). Ciprofloxacina (*CPF*) y danofloxacina (*DNF*) fueron adquiridas de Fluka (Buchs, Suiza) y marbofloxacina (*MRF*) fue obtenida de Molekula (Gillingham, Reino Unido). El acetonitrilo (*ACN*) y el metanol (*MeOH*), ambos calidad cromatografía líquida, fueron obtenidos de LiChrosolv (Merk Millipore Co., Darmstadt, Alemania). El ácido acético glacial (*HAc*) fue adquirido de Merk (Darmstadt, Alemania) y el acetato de sodio trihidrato (*NaAc*) calidad p.a. fue provisto por Anedra (La Plata, Argentina). El agua de laboratorio ultrapura fue obtenida por un sistema de purificación de agua Milli-Q MilliPore (Billerica, Estados Unidos).

Soluciones de trabajo. Se preparó una solución madre de cada f-QUI mediante disolución de la cantidad apropiada de la droga sólida correspondiente en metanol alcalinizado (pH ~9.0) obteniéndose una concentración final de 200.0 $\mu\text{g mL}^{-1}$. Para su conservación, todas las soluciones se almacenaron a 4 °C protegidas de la luz por un periodo no mayor a 2 meses.

Para la determinación cromatográfica se preparó una solución reguladora de HAc/NaAc, disolviendo una cantidad apropiada de NaAc en agua ultrapura, ajustando el pH a un valor de 4.0 con HAc glacial y completándose el volumen de 1000.0 mL con agua ultrapura, lográndose una concentración formal igual a 0.010 mol L⁻¹. Previamente a ser utilizada como fase móvil cromatográfica, la solución se filtró a través de una membrana de nylon con tamaño de poro de 0.45 μm (Sartorius, Goettingen, Alemania).

Muestras de calibración y validación. Para la evaluación y aná-

lisis de los datos generados utilizando las diferentes metodologías propuestas, se prepararon distintas soluciones muestra transfiriéndose una alícuota apropiada de cada solución madre de f-QUI a matraces aforados de 5.0 mL y completándose el volumen con agua ultrapura. Las soluciones se prepararon según la metodología a evaluar: **1- Metodología I:** solución ternaria de analitos conteniendo $4.0 \mu\text{g mL}^{-1}$ de OFL, $15.0 \mu\text{g mL}^{-1}$ de CPF y $5.0 \mu\text{g mL}^{-1}$ de DNF; **2- Metodología II:** solución binaria de analitos conteniendo 15.0 ng mL^{-1} de OFL y CPF; y **3- Metodología III:** solución conteniendo $7.0 \mu\text{g mL}^{-1}$ de CPF.

Para el desarrollo del método analítico modelo se preparó 1 set de calibración compuesto de 12 muestras estándares conteniendo OFL, CPF y DNF, utilizándose un diseño central compuesto fraccionado con 1 réplica en el punto central. Las concentraciones finales estuvieron comprendidas en el rango de entre $0.0 \mu\text{g mL}^{-1}$ y $100.0 \mu\text{g mL}^{-1}$ para OFL, $0.0 \mu\text{g mL}^{-1}$ y $150.0 \mu\text{g mL}^{-1}$ para CPF y $0.0 \mu\text{g mL}^{-1}$ y $25.0 \mu\text{g mL}^{-1}$ para DNF. Se construyó, además, 1 set de validación comprendido por 10 muestras estándares conteniendo OFL, CPF y DNF en niveles de concentración diferentes a los utilizados en la calibración, si bien comprendidas en el mismo rango; adicionalmente, se incorporaron MRF y ENO en todas las muestras de validación, las que se consideraron como componentes esperados no calibrados. Las soluciones de calibración y validación se prepararon transfiriéndose alícuotas apropiadas de las soluciones madres correspondientes a cada analito a matraces aforados de 25.0 mL y completándose el volumen con agua ultrapura.

Adicionalmente, se preparó un set de 6 muestras de agua de consumo fortificadas con cada f-QUI, adicionándose, además, MRF y ENO como interferentes. Las muestras de agua de consumo se prepararon transfiriéndose las alícuotas apropiadas de las soluciones madres de cada analito a matraces aforados de 25.0 mL y completándose el volumen con agua de red, agua mineral comercial o agua de napa subterránea apta para consumo, según corresponda. Finalmente, todas las soluciones muestras preparadas se sometieron a un proceso de pre-concentración basado en el procedimiento descrito en el apartado **Extracción en fase sólida** de la sección 3.2.

Instrumentación y procedimiento. Todos los experimentos cromatográficos se llevaron a cabo usando un instrumento Agilent 1100 LC (Agilent Technologies, Waldbronn, Alemania) equipado con desgasificador, bomba cuaternaria, inyector automático, compartimento termostatzado para columna, DAD UV-Vis, FSFD

y el programa informático ChemStation (Agilent Technologies, Waldbronn, Alemania) requerido para el control del instrumento y la adquisición de los datos. La separación cromatográfica se realizó utilizando una columna analítica 3.5 μm Zorbax Eclipse XDB-C18 (75 mm \times 4.6 mm) (Agilent Technologies, Waldbronn, Alemania) como fase estacionaria, con una temperatura de compartimento para columna de 35 $^{\circ}\text{C}$. La fase móvil consistió en una mezcla de solución reguladora 0.010 mol L⁻¹ de HAc/NaAc pH 4.0:MeOH:ACN en una proporción 71:20:9.

Las mediciones de pH se realizaron utilizando un potenciómetro Orion 410A (Massachusetts, Estados Unidos) equipado con un electrodo de vidrio combinado Boeco BA17 (Hamburgo, Alemania).

Metodología I – Recolección de fracciones. La separación cromatográfica se realizó en modo isocrático, con una velocidad de flujo de 1.8 mL min⁻¹, durante 2 min. A la salida del cromatógrafo se conectó un dispositivo automatizado, construido en nuestro laboratorio, que permitió la recolección de fracciones en placas de ELISA (*ELISA: Enzyme-Linked ImmunoSorbent Assay*) de 96 pocillos. Las fracciones se recolectaron en placas de ELISA de 96 pocillos comenzándose luego de 47 s de realizada la inyección cromatográfica, durante 34 s. Cada fracción se recolectó en un periodo de 2 s, implicando un volumen de 60 μL de muestra por pocillo, considerando la velocidad de flujo utilizada en el procedimiento cromatográfico. De esta manera, se obtuvo un total de 17 fracciones para cada muestra analizada. Al finalizar la recolección, la placa de ELISA se colocó en un espectrofotómetro de fluorescencia Perkin-Elmer LS-55 (Waltham, Estados Unidos) equipado con un accesorio para lectura de placas de ELISA acoplado a una fibra óptica y conectado a una computadora mediante un conector tipo RS232C. Las matrices de excitación-emisión de fluorescencia (*EEM*) se registraron variando la longitud de onda de emisión en el rango espectral comprendido entre 380.0 nm y 500.0 nm, en pasos de 5 nm, y registrándose el espectro de excitación en el rango comprendido entre 260.0 nm y 340.0 nm, cada 5 nm, a cada longitud de onda de emisión. De esta manera, se obtuvieron hiper-matrices de tiempo de retención-espectro de excitación-espectro de emisión de fluorescencia (*TEEM: Time-Excitation-Emission Matrix*) de 17 \times 17 \times 25 datos, respectivamente. La detección se realizó utilizando aberturas de rendija de excitación y emisión de 10 nm, operando el detector a un voltaje de 600 V y utilizando una velocidad de barrido espectral de 800 nm min⁻¹. Para el control del instrumento y la adquisición de los datos se utilizó una interfaz gráfica desarrollada en el laboratorio, escrita y operada en Processing

2.0b6 (Processing Foundation, Boston, EE.UU).

Metodología II – Múltiples corridas cromatográficas. La separación cromatográfica se realizó en modo isocrático, con una velocidad de flujo de 2.0 mL min^{-1} , durante 2 min. A fin de construir hiper-matrices TEEM, se realizaron 10 corridas cromatográficas, registrándose el espectro de emisión de fluorescencia en el rango espectral comprendido entre 380.0 nm y 500.0 nm a cada tiempo de retención, en el rango temporal de entre 0.0 min y 2.0 min, utilizando longitudes de onda de excitación diferentes para cada corrida cromatográfica. Las longitudes de onda de excitación utilizadas estuvieron comprendidas en el rango espectral de entre 260.0 nm y 305.0 nm, con $\Delta\lambda_{\text{ex}} = 5 \text{ nm}$. De esta manera, se obtuvieron hiper-matrices TEEM de $150 \times 10 \times 45$ datos para los modos temporal, espectral de excitación y espectral de emisión, respectivamente. El detector se operó con una ganancia del fotomultiplicador igual a 18, utilizando una velocidad de barrido igual a $10800 \text{ nm min}^{-1}$.

Metodología III – Detección en línea de EEM de fluorescencia. La separación cromatográfica se realizó en modo isocrático, con una velocidad de flujo de 0.5 mL min^{-1} , durante 5 min. Para el registro de las EEM se utilizó un espectrofotómetro de fluorescencia de barrido rápido Agilent Cary Eclipse (Agilent Technologies, Waldbronn, Alemania) y una celda de flujo de cuarzo de $10 \mu\text{L}$ de volumen ($1 \text{ mm} \times 1 \text{ mm}$ de paso óptico) conectada a la salida del cromatógrafo. La obtención de cada EEM demandó un tiempo de aproximadamente 12 s, permitiendo registrar múltiples EEM consecutivas para una misma corrida cromatográfica, sin requerir de la detención del flujo o la recolección de fracciones al final del procedimiento cromatográfico. Se registró un total de 15 EEM consecutivas para cada muestra analizada, variándose la longitud de onda de excitación en el rango espectral comprendido entre 260.0 nm y 300.0 nm, en pasos de 2.5 nm, y registrándose el espectro de emisión en el rango comprendido entre 390.0 nm y 490.0 nm, cada 3.75 nm, para cada longitud de onda de excitación. De esta manera, se obtuvieron hiper-matrices TEEM de $15 \times 15 \times 28$ datos para el modo temporal, espectral de excitación y espectral de emisión de fluorescencia, respectivamente. La detección se realizó utilizando aberturas de rendija de excitación y emisión de 10 nm, operando el detector a un voltaje de 850 V y utilizando una velocidad de barrido espectral de $18000 \text{ nm min}^{-1}$. Para el control del espectrofotómetro y la adquisición de los datos se utilizó el programa informático Cary Eclipse (Agilent Technologies, Waldbronn, Alemania). La automatización y sincroniza-

ción de los instrumentos se realizaron utilizando un controlador desarrollado en el laboratorio soportado en Python 2.7.11 (Python Software Foundation, <http://www.python.org/>) con estructura servidor-cliente. Los instrumentos, o clientes, se conectaron a través de una red LAN y sus interfaces gráficas, o GUI, necesarias para el control de cada instrumento fueron automatizadas utilizando rutinas de entrenamiento y operadas automáticamente mediante SikuliX 1.0.1 (RaiMan, <http://sikulix.com>).

Programas informáticos y algoritmos para el análisis de datos. El procesamiento de datos y el análisis quimiométrico se realizaron en MatLab 7.10 (MathWorks Inc., Natick, Estados Unidos). Los algoritmos necesarios para operar MCR-ALS se encuentran asequibles en internet accediendo a la página web <http://www.mcrals.info/>. El paquete MVC3, que incluye a los algoritmos quimiométricos para la resolución de datos de tercer orden, se encuentra asequible en internet accediendo a la página web www.iquir-conicet.gov.ar/descargas/mvc3.rar.

RESULTADOS

Requerimientos instrumentales. A fin de evaluar diferentes estrategias de generación de datos multidimensionales y analizar las propiedades de los datos obtenidos, se propusieron 3 arreglos instrumentales posibles que permitiesen obtener datos de tercer orden basados en separación cromatográfica y detección de matrices de excitación-emisión de fluorescencia para un sistema analítico modelo. Las tres metodologías propuestas tuvieron como base la separación cromatográfica de un sistema analítico de iguales características, utilizándose similares condiciones separativas. Cabe mencionar que no se realizará un análisis comparativo respecto a las intensidades absolutas obtenidas, en términos de sensibilidad o SNR, debido a que se desconocen los parámetros técnicos característicos correspondientes a los detectores de cada instrumento utilizado.

Metodología I. Para llevar a cabo la *Metodología I*, basada en la recolección de fracciones al final del procedimiento cromatográfico, además de requerir de 2 instrumentos analíticos diferentes, un cromatógrafo de alto rendimiento y un espectrofotómetro de fluorescencia, fue necesario desarrollar un dispositivo automatizado que

permitiese la recolección de fracciones en placas de ELISA de 96 pocillos. Si bien esto podría referirse a un impedimento instrumental, el dispositivo de recolección se construyó con material reciclado de impresoras fuera de uso y material disponible en el laboratorio, resolviéndose, así, el inconveniente de una manera sencilla, mediante una alternativa de fácil acceso y bajo costo (Siano, 2013). Adicionalmente, se debe considerar que para realizar las lecturas de fluorescencia se debe disponer de un espectrofotómetro de fluorescencia equipado con un accesorio para lectura de placas de ELISA que permita el registro de EEM de cada pocillo. Aquí, las lecturas de fluorescencia se realizaron con un espectrofotómetro de fluorescencia convencional, equipado con el accesorio para lectura de placas, controlado con un programa informático escrito por un integrante del laboratorio que permitió el registro de EEM de cada pocillo (Siano, 2013), debido a que el programa informático original para control del equipo y adquisición de datos no dispone de esta opción.

El análisis experimental completo de una muestra conteniendo 3 analitos se realizó en un tiempo total de 42 min, contemplándose el tiempo requerido para el procedimiento cromatográfico (2 min) y el demandado para la detección y registro de las 17 EEM (40 min). Como puede apreciarse, la etapa de detección es la responsable del elevado tiempo de análisis requerido para cada muestra, convirtiendo al método en una alternativa ineficiente para la determinación y análisis de muestras inestables o de muestras conteniendo soluciones volátiles. Sin embargo, el tiempo demandado en esta etapa puede ser drásticamente reducido utilizándose un espectrofotómetro de fluorescencia de barrido rápido.

Metodología II. Para proceder con la *Metodología II*, basada en múltiples inyecciones cromatográficas, se requiere solamente de un cromatógrafo de alto rendimiento equipado con un detector de fluorescencia de barrido rápido. Además, si bien no resultaría un requisito esencial para llevar a cabo los experimentos, la presencia de un inyector automático que permita realizar inyecciones repetidas de manera automatizada y precisa, minimiza las variaciones que pudiesen ocurrir, principalmente, en los tiempos de retención y las áreas cromatográficas correspondientes a los diferentes analitos, facilitando el procesamiento del dato multidimensional obtenido. Por lo tanto, a pesar de requerir de un único instrumento analítico, los módulos necesarios para llevar a cabo los experimentos, detector de fluorescencia de barrido rápido e inyector automático, no siempre se encuentran disponibles en los equipos utilizados en laboratorios de rutina, incrementándose, así, el costo y la complejidad del instrumental re-

querido para el desarrollo del método analítico.

En este trabajo, el análisis de una muestra binaria de analitos demandó un tiempo total de 40 min, considerándose que el tiempo consumido por el procedimiento cromatográfico fue de 2 min por cada corrida cromatográfica y que se realizaron 10 repeticiones de inyección de la misma muestra, incluyendo el tiempo demandado por el instrumento para realizar las diferentes inyecciones. Una de las principales características de esta metodología es que la resolución del espectro de excitación se encuentra ligada directamente al número de inyecciones cromatográficas y, por consiguiente, el tiempo de análisis total dependerá del tiempo consumido por cada corrida cromatográfica y de la calidad espectral de excitación deseada o requerida por el analista. Por lo tanto, el elevado tiempo de análisis solo podrá ser reducido a expensas de una pérdida de resolución o información espectral. En consecuencia, esta metodología se limita al análisis de muestras estables y muestras multi-componentes de características espectrales similares. Adicionalmente, en virtud del alto número de inyecciones necesarias, el elevado consumo de solventes y el requerimiento de volumen de muestra convierten a esta metodología en una alternativa cuestionable bajo el marco de la química analítica verde, como también, en una alternativa con un costo económico considerablemente elevado.

Metodología III. La metodología aquí propuesta como estrategia para lecturas en línea de EEM requiere de 2 instrumentos analíticos diferentes. Es necesario disponer de un cromatógrafo líquido de alto rendimiento convencional, de un espectrofotómetro de fluorescencia de barrido rápido y una celda de flujo de volumen reducido. En este caso, no resultaría necesario disponer de un cromatógrafo equipado con detectores sofisticados o de alta sensibilidad, debido a que la detección se realiza de manera externa, si bien la utilización de un inyector automático favorecería a la generación del dato multidimensional. Adicionalmente, en virtud de que el registro de las EEM se realiza en línea al procedimiento cromatográfico, la necesidad de un espectrofotómetro de fluorescencia de barrido rápido es el principal requerimiento instrumental, fundamentalmente, para corridas cromatográficas cortas, a fin de minimizar el efecto de elución de los analitos, es decir, de la distribución de la concentración del analito en el tiempo. Asimismo, los parámetros instrumentales seleccionados para el registro de las EEM, como rango espectral, resolución y velocidad de barrido espectral, deben permitir la obtención de matrices que contengan información y resolución espectral aceptable en un tiempo considerablemente corto.

Debido a que la *Metodología III* se propone como nueva alternativa para la obtención de datos de tercer orden, resultó necesario analizar detalladamente la estructura de los datos obtenidos utilizándose un sistema analítico simple. En consecuencia, se analizó una solución conteniendo solo un analito, realizándose réplicas de inyección y evaluando las variaciones que pudiesen presentarse a causa del arreglo instrumental, independientes a la composición de la muestra. El tiempo experimental total requerido para el análisis de una muestra conteniendo un analito fue de 5 min, obteniéndose 15 EEM consecutivas por muestra. Teniendo en cuenta que el registro de las EEM se realiza paralelamente al procedimiento cromatográfico, el tiempo de análisis total por muestra se define por las condiciones utilizadas en el método cromatográfico. Estas características convierten a la metodología propuesta en una excelente alternativa para la determinación de muestras multi-componentes y consecuente generación de datos de tercer orden desde el punto de vista del tiempo de análisis, automatización, consumo de reactivos y disponibilidad de muestra.

Propiedades de los datos multidimensionales. Al desarrollar un método basado en el procesamiento de datos multidimensionales, es importante analizar detalladamente las propiedades y características de los datos obtenidos a fin de seleccionar el pre-procesamiento correcto de los datos, pero, fundamentalmente, el algoritmo más adecuado para el procesamiento y el análisis de los datos.

En análisis multidimensional y/o multi-vía resulta crucial reconocer el tipo de datos que se manipulan en virtud de sus propiedades matemáticas, para poder establecer el correcto procedimiento de análisis y así obtener resultados inequívocos. Para datos de tercer orden, es importante reconocer, en primer lugar, si los datos cumplen el concepto de trilinealidad, en caso contrario, es necesario reconocer el modo en el que se pierde la trilinealidad, evaluando las posibles alternativas que pudiesen ser utilizadas para resolver el problema correctamente. Cabe destacar que, a diferencia de un set de datos de segundo orden dispuesto en un arreglo de tres vías, la trilinealidad de los datos de tercer orden es una propiedad que debe ser evaluada para cada muestra individualmente y es un fenómeno que ocurre en los tres modos instrumentales que son independientes entre sí. Así, un dato no trilineal de tercer orden se obtiene en presencia de fenómenos mutuamente dependientes en al menos 2 de los 3 modos instrumentales (Olivieri *et al.*, 2014b).

En este trabajo, se realizó un análisis cualitativo sobre los datos generados con

cada metodología propuesta con el objetivo de evaluar la propiedad de trilinealidad y, en caso de presentar desvíos en alguno de sus modos, analizar las diferentes estrategias posibles que pudiesen ser utilizadas para superar el inconveniente.

Metodología I. Se debe considerar que las fracciones recolectadas en los pocillos de las placas de ELISA a la salida del cromatógrafo representan los tiempos de retención correspondientes a los diferentes analitos presentes en la muestra. Por ello, es importante que los tiempos establecidos para la recolección, es decir, tiempo de inicio de recolección y tiempo de espera en cada pocillo, sean conocidos y constantes en todo el procedimiento experimental, a fin de poder reconstruir el perfil temporal correspondiente.

Una particularidad que presenta esta metodología es que las EEM registradas para cada pocillo muestran independencia entre sí, es decir, para una sustancia pura los espectros de excitación y emisión solo dependerán del analito y del medio en el que se encuentra, cuyas intensidades serán en virtud de la abundancia del analito presente en el pocillo correspondiente. Considerándose un sistema cromatográfico operando en modo isocrático, en el que la composición del medio se mantiene invariable, los espectros de excitación y emisión serán idénticos para todos los pocillos, difiriendo en su intensidad debido a la dispersión del analito resultado de la retención cromatográfica. De esta manera, y teniendo en cuenta que las EEM, por su naturaleza, son datos de segundo orden perfectamente bilineales, el dato de tercer orden generado se presenta como un dato trilineal, es decir, los modos instrumentales correspondientes a excitación, emisión y tiempo son independientes entre sí. En la Figura 3.47 se muestran las EEM registradas para cada pocillo, reconstruyéndose el perfil temporal obtenido para una solución muestra conteniendo tres analitos.

Un aspecto a tener en cuenta en la generación de datos multidimensionales es, además, el número de datos obtenidos en cada modo instrumental. Idealmente, se debería obtener un número de datos razonable y balanceado en cada uno de los modos, sin perder información relevante en ninguno de ellos. En este caso, la hipermatriz TEEM generada para cada muestra presentó un tamaño de $25 \times 17 \times 25$ datos para las dimensiones temporales, espectrales de excitación y emisión de fluorescencia, respectivamente. Si bien se podría considerar como una matriz de proporciones balanceadas, en el modo temporal se obtuvieron 25 fracciones discretas para una muestra proveniente de un procedimiento cromatográfico, lo que lleva consigo una baja resolución en este modo. Sin embargo, así como el tiempo de análisis puede ser

reducido utilizándose un espectrofotómetro de fluorescencia de barrido rápido, el número de fracciones pueden ser incrementado utilizando microplacas de ELISA de 384 pocillos, reduciéndose, así, el volumen necesario para la lectura en cada pocillo y, por lo tanto, el tiempo de recolección de cada fracción.

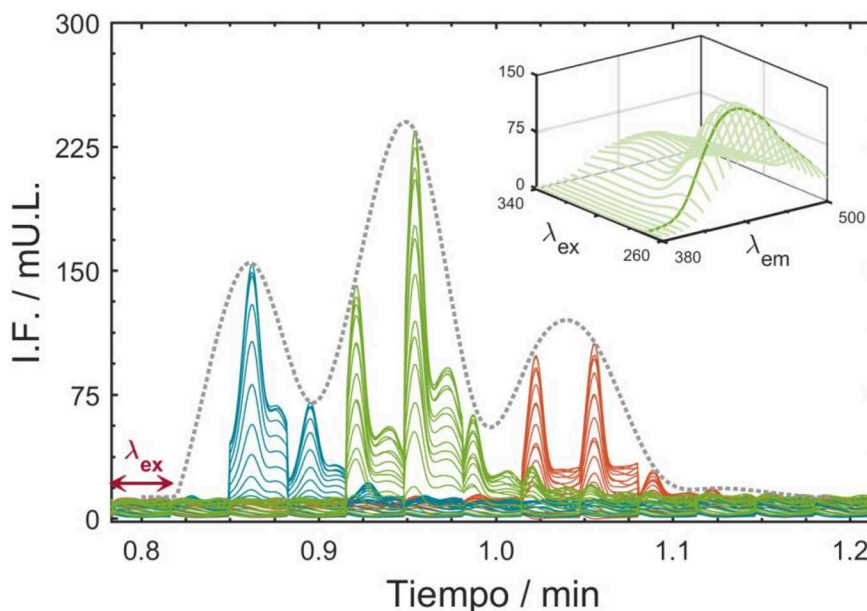


Figura 3.47. Matrices de excitación-emisión, vistas desde el modo de excitación, registradas para las fracciones recolectadas en pocillos de ELISA de muestras conteniendo los analitos puros de OFL (azul), CPF (verde) y DNF (rojo). En línea de puntos gris se representa el perfil cromatográfico obtenido a partir de los máximos de intensidad registrados en cada pocillo para una muestra conteniendo los tres analitos. En el recuadro se muestra una matriz de excitación-emisión registrada para una de las fracciones recolectadas de CPF, indicándose el espectro de emisión al máximo de excitación correspondiente.

Metodología II. El aspecto más importante a destacar para esta metodología es que los espectros de excitación son el resultado de diferentes inyecciones o corridas cromatográficas realizadas para una misma muestra. Por lo tanto, la región espectral de excitación cubierta y la resolución espectral dependerán directamente del número de inyecciones o corridas cromatográficas que se realicen. Además, es de notar que, a una misma longitud de onda de emisión, se podrán obtener espectros de excitación correspondientes a cada tiempo de retención, es decir, es posible reconstruir una matriz tiempo de retención-espectro de excitación a partir de los perfiles cromatográficos extraídos a una misma longitud de onda de emisión. Sin embargo, para reconstruir la matriz de tiempo de retención-espectro de excitación es muy importante tener en cuenta que el tiempo de retención de cada analito deber presentar reproducibilidad en cada corrida cromatográfica. Por lo tanto, un corrimiento mínimo observado en los tiempos de retención entre diferentes inyecciones para una misma muestra lle-

vará consigo una pérdida de trilinealidad para los datos de tercer orden correspondientes a las hiper-matrices TEEM. Esto se traduce a que el modo instrumental correspondiente a los espectros de excitación presenta dependencia con el modo correspondiente a los tiempos de retención. Este fenómeno puede representarse análogamente con datos cromatográficos de segundo orden dispuestos en un arreglo de tres vías para el análisis de múltiples muestras, en donde la trilinealidad se pierde en virtud de los corrimientos de picos cromatográficos presentes entre las diferentes muestras. Por consiguiente, y sobre la base de estas observaciones, los datos de tercer orden generados utilizando esta metodología solo serán trilineales si los tiempos de retención para cada analito en cada corrida cromatográfica, para una misma muestra, se mantienen, de otra manera, los datos se presentaran como datos no trilineales. La Figura 3.48 muestra una matriz de tiempo de retención-espectro de excitación reconstruida con los cromatogramas obtenidos a la misma longitud de onda de emisión, mostrándose que a causa de pequeñas diferencias en los tiempos de retención para distintas inyecciones se obtienen diferentes espectros de excitación para un mismo analito.

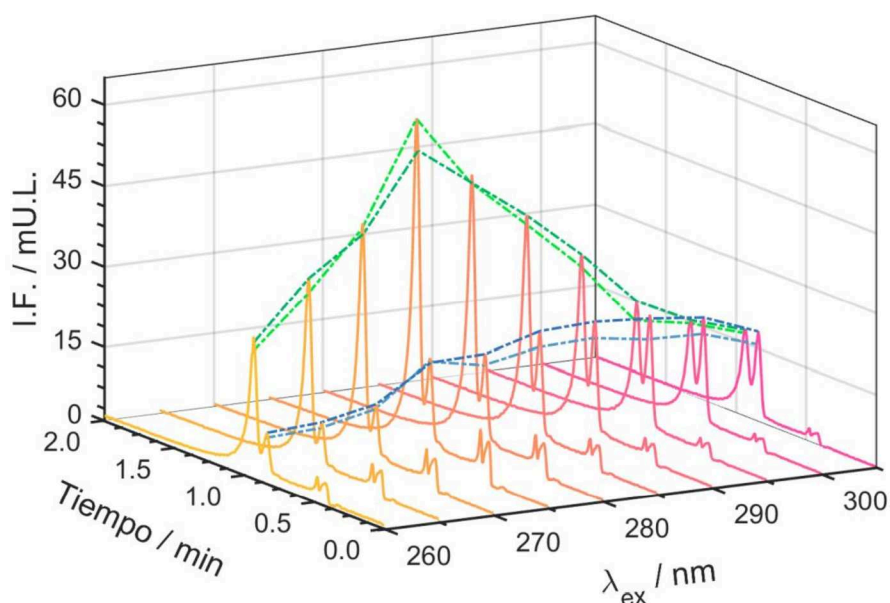


Figura 3.48. Cromatogramas correspondientes a $\lambda_{em} = 350$ nm extraídos de las matrices de tiempo de retención-espectro de emisión para cada longitud de onda de excitación, mostrándose los espectros de excitación obtenidos para OFL a un tiempo de retención $t_R = 0.840$ min (azul claro) y $t_R = 0.852$ min (azul oscuro) y para CPF a un tiempo de retención $t_R = 0.930$ min (verde oscuro) y $t_R = 0.942$ min (verde oscuro).

Sin embargo, ante la posible pérdida de trilinealidad, se pueden proponer distintas estrategias de pre-procesamiento de datos que permiten recuperar la trilinealidad, como por ejemplo, utilizar algoritmos que corrigen corrimientos de picos cromatográficos

cos.

Adicionalmente, en referencia al número de datos obtenidos, se obtuvieron hiper-matrices de $150 \times 10 \times 45$ datos para las dimensiones temporales, espectrales de excitación y de emisión de fluorescencia, respectivamente. Puede observarse que los modos correspondientes a los tiempos de retención y a los espectros de emisión presentan una resolución considerablemente más alta que la correspondiente a los espectros de excitación, obteniéndose una hiper-matriz de dimensiones desproporcionadas. Desafortunadamente, el número de datos correspondiente al modo de excitación solo podrá ser incrementado si se aumenta el número de inyecciones o corridas cromatográficas por muestra, incrementándose, de esta manera, el tiempo de análisis y el consumo de reactivos y de muestra.

Metodología III. La particularidad más destacada que presenta esta metodología es que las EEM de fluorescencia se registran simultáneamente con la elución del analito del sistema cromatográfico. Si bien la rápida velocidad de barrido de fluorescencia del instrumento utilizado permite registrar EEM en un tiempo muy reducido, existe, de manera inherente, una dependencia directa entre el espectro de excitación y la elución del analito, no así entre los espectros de emisión y el tiempo de retención. Este fenómeno se observa en virtud de que la longitud de onda de excitación cambia en el sentido que aumenta el tiempo de retención, es decir, a cada tiempo se corresponde una longitud de onda de excitación diferente. Sin embargo, debido a que el espectro de emisión de una sustancia es invariable y su intensidad depende de la longitud de onda de excitación y de la abundancia del analito y, adicionalmente, es registrado en un tiempo muy corto, en el análisis de una sustancia pura se observarán espectros de emisión idénticos de diferente intensidad, en todo el procedimiento analítico. Estas observaciones se traducen en una dependencia del modo instrumental de excitación de fluorescencia con el modo correspondiente al tiempo de retención, llevando consigo la ruptura de la trilinealidad del dato de tercer orden. Esta ruptura de la trilinealidad se presenta como un gran obstáculo que debe ser superado para lograr el correcto análisis de los datos, ya sea de manera instrumental utilizando espectrofotómetros con barrido de fluorescencia aún más rápido, o matemática utilizando algoritmos que permitan recuperar la trilinealidad de los datos o que permitan analizar datos de tercer orden no trilineales.

En la Figura 3.49.A se muestran las matrices de excitación-emisión registradas a diferentes tiempos de retención para una solución conteniendo un analito, reconstru-

yéndose el perfil temporal correspondiente. Las Figura 3.49.B.1 y Figura 3.49.B.2 muestran los perfiles de excitación y de emisión obtenidos para el analito a diferentes tiempos de retención, respectivamente.

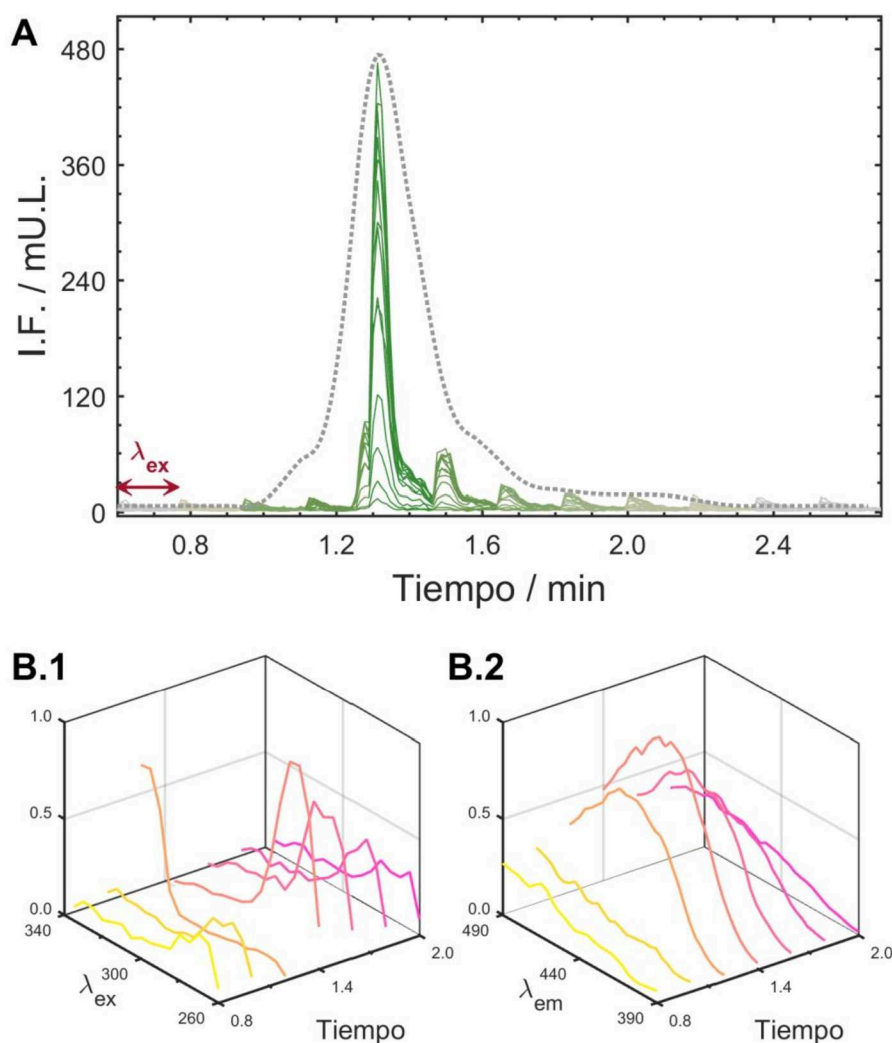


Figura 3.49. Matrices de excitación-emisión consecutivas, vistas desde el modo de excitación, registradas para una muestra conteniendo CPF (verde) (A). En línea de puntos gris se representa el perfil cromatográfico obtenido a partir de los máximos de intensidad registrados para cada matriz. Se muestran las matrices de tiempo de retención-espectro de excitación (B.1) y tiempo de retención-espectro de emisión (B.2) normalizadas, reconstruidas a partir de los espectros de excitación y emisión correspondientes a los máximos de emisión y excitación, respectivamente.

Finalmente, analizando el número de datos obtenidos para cada modo instrumental, se obtuvieron hiper-matrices TEEM de $15 \times 15 \times 28$ datos para las dimensiones temporales, espectrales de excitación y de emisión de fluorescencia, respectivamente. Al igual que lo sucedido para la Metodología I, si bien se obtiene un número balanceado de datos para cada modo, para el modo temporal se obtuvieron 15 matrices consecutivas para una muestra proveniente de un procedimiento cromatográfico,

significando una baja resolución en el modo temporal. Este número solo será incrementado si se reduce el rango espectral de excitación, o su resolución, o se incrementa la velocidad de barrido espectral, requiriéndose, entonces, de un instrumental más sofisticado.

Pre-procesamiento y procesamiento de datos. Con el objetivo de demostrar las propiedades de los datos mencionadas anteriormente, se analizaron diferentes hiper-matrices de datos obtenidas mediante las tres metodologías propuestas de manera individual utilizando un algoritmo de descomposición trilineal (TLD). Se utilizó PARAFAC como algoritmo quimiométrico para el modelado de datos aprovechando las particularidades que presenta: **1-** solo permite el análisis de datos que no presenten desvíos de la trilinealidad; **2-** se obtienen resultados con interpretación física; y **3-** el resultado de la resolución es único. De esta manera, y conociendo *a priori* las propiedades del sistema (tiempos de retención, espectros de excitación y emisión de fluorescencia de los componentes puros), fue posible analizar si los resultados obtenidos de la descomposición reflejan las propiedades reales del sistema. Es importante mencionar que no se aplicaron herramientas de pre-procesamiento a ningún set de datos.

En la Figura 3.50 se muestran los resultados obtenidos para los distintos arreglos de datos de tercer orden obtenidos mediante las tres metodologías propuestas. En todos los casos se analizaron diferentes parámetros de inicialización, si bien siempre se mantuvo la restricción de no-negatividad en los tres modos. El número de componentes presentes en cada muestra se evaluó mediante el análisis CORCONDIA (*CORCONDIA: CORe CONSistency DIAGnostic*).

Metodología I. En todos los casos, el número de componentes para cada set de datos analizado fue igual a 3, reflejándose el número real de compuestos presentes en las muestras. Los perfiles correspondientes a los espectros de excitación y emisión de fluorescencia mostraron idéntica similitud a los espectros reales obtenidos en las mismas condiciones para cada analito puro (Figura 3.50, *Metodología I*).

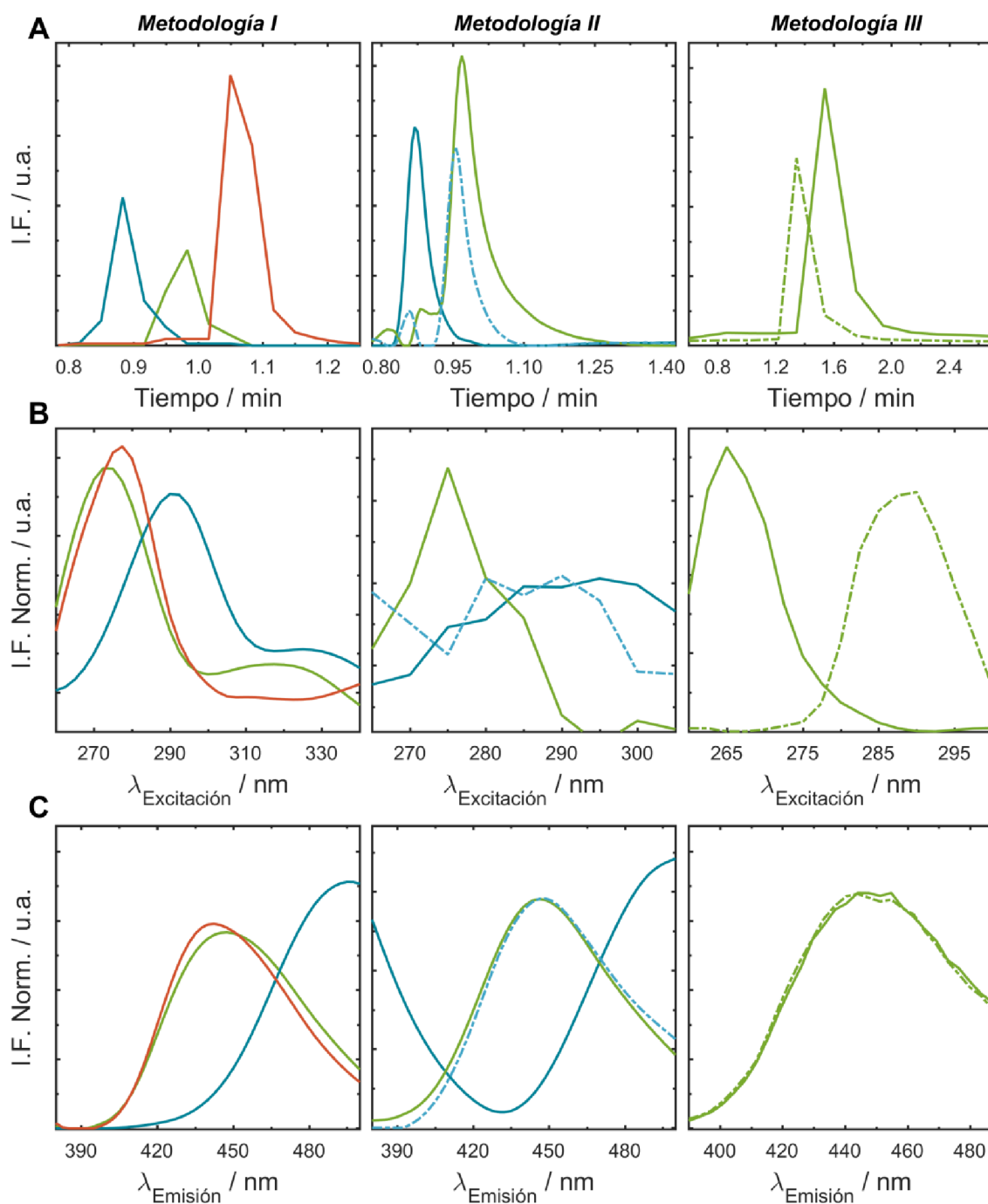


Figura 3.50. Resultados obtenidos por PARAFAC para las 3 metodologías propuestas, mostrándose los perfiles temporales (A), espectrales de excitación (B) y espectrales de emisión (C). Para la Metodología I se muestran los perfiles de CPF (línea continua verde), OFL (línea continua azul) y DNF (línea continua rojo); para la Metodología II se indican los perfiles de CPF (línea continua verde) y OFL (línea continua azul) y un perfil correspondiente a un factor desconocido (línea de puntos y rayas azul verdosa); para la Metodología III se muestran los perfiles obtenidos para CPF (verde), observándose 2 factores para el mismo componente

Es importante mencionar que se registraron, adicionalmente, los cromatogramas obtenidos para cada análisis cromatográfico mediante detección DAD-UV, que luego fueron utilizados como referencia a fin de analizar y comparar los tiempos de retención obtenidos de la descomposición de los datos multidimensionales. Los perfiles tempora-

les obtenidos por PARAFAC para todos los sets de datos generados mostraron perfecta similitud entre ellos, si bien se observaron diferencias entre los tiempos de retención correspondientes a las señales cromatográficas de distintas muestras, indicándose una falta de cuadrilinealidad en una disposición o arreglo de cuatro vías para un análisis de múltiples muestras. Además, la comparación entre el perfil temporal obtenido por PARAFAC y el cromatograma de referencia DAD-UV mostró un alto grado de similitud, si bien se observaron pequeños retardos en los tiempos correspondientes a los datos de tercer orden, considerándose que la detección DAD-UV se realiza antes de la recolección de fracciones.

Sobre la base de los resultados presentados se puede asegurar que los datos generados con la *Metodología I* presentan perfecta trilinealidad, pudiéndose obtener perfiles espectrales de excitación y emisión de alta calidad. Sin embargo, en un arreglo de 4-vías para el análisis de múltiples muestras se observan desvíos de cuadrilinealidad a causa de corrimientos en los tiempos de retención de los analitos entre distintas muestras. Por lo tanto, y debido a que PARAFAC solo es aplicable para arreglos de 4-vías cuadrilineales, el modelado quimiométrico debe ser realizado utilizando algoritmos quimiométricos más flexibles, como U-PLS/RTL, APARAFAC, PARAFAC2 o MCR-ALS con matrices de datos desdobladas.

Metodología II. En este caso, no pudo establecerse un número único de componentes para todos los sets de datos analizados, obteniéndose valores de entre 2 y 4 factores. Esta diferencia entre el número de componentes estimado y el número real de compuestos presentes en la muestra se debe, principalmente, a los corrimientos observados entre los tiempos de retención de los analitos para una misma muestra. En la Figura 3.50, *Metodología II* puede observarse que para una muestra con un número de componentes estimado igual a 3 se obtienen 3 perfiles cromatográficos y espectrales de excitación bien definidos, mientras que para el perfil espectral de emisión se muestran 2 perfiles de características similares y 1 perfil perfectamente resuelto. Estos perfiles son el resultado de una falta de trilinealidad en el dato de tercer orden debido a la dependencia entre los modos correspondientes a los tiempos de elución y espectros de excitación y, en consecuencia, a una deficiencia de rango resultante en el modo correspondiente a los espectros de emisión. Adicionalmente, los perfiles espectrales de excitación no reflejan los perfiles de excitación reales de los analitos, si bien los perfiles correspondientes a los espectros de emisión de los 2 analitos presentes en la muestra presentan un alto grado de similitud con sus espectros reales.

Debido a que la falta de trilinealidad se debe específicamente a los corrimientos observados entre los tiempos de retención de una misma muestra, se pueden utilizar herramientas de pre-procesamiento de datos que permitan recuperar la trilinealidad del dato, como COW (Nielsen *et al.*, 1998) o *i*-coshift (Tomasi *et al.*, 2011), entre otros. Esta metodología, si bien resulta en una estrategia prometedora, demanda un tiempo de procesamiento considerablemente alto y, además, se debe tener la precaución de no introducir modificaciones indeseadas adicionales. Por ello, es importante utilizar algoritmos de corrección de corrimientos de señales cromatográficas que consideren a los datos como un matriz completa bidimensional y no como un arreglo de vectores individuales.

Si bien para fines cuantitativos estos datos pueden ser analizados mediante algoritmos de modelado quimiométrico más flexibles que PARAFAC, como U-PLS/RTL o APARAFAC, es importante considerar el efecto de la falta de trilinealidad para los datos de tercer orden originada por los corrimientos ocurridos entre tiempos de retención de una misma muestra y la falta de cuadrilinealidad en virtud de los corrimientos entre tiempos de retención de distintas muestras, a fin de obtener resultados analíticos confiables. En la bibliografía se pueden encontrar trabajos que reportan falta de cuadrilinealidad para datos de tercer orden dispuestos en un arreglo de 4-vías debido a corrimientos observados entre los tiempos de retención para distintas muestras, pero no se considera el mismo fenómeno para múltiples corridas cromatográficas de una misma muestra (Lozano *et al.*, 2013; Montemurro *et al.*, 2016), subestimándose, así, un efecto por pérdida de trilinealidad.

Metodología III. En todos los casos, el número de componentes para cada set de datos analizado correspondiente a las muestras conteniendo un analito fue mayor a 1. En la Figura 3.50, *Metodología III* puede observarse que para un mismo compuesto se resolvieron 2 perfiles temporales y espectrales de excitación, si bien los 2 perfiles correspondientes a los espectros de emisión muestran perfecta similitud. Esto se debe a que para una misma muestra se obtienen diferentes espectros de excitación en virtud de su dependencia con la elución del analito. Sin embargo, y como se mencionó anteriormente, los espectros de emisión no presentan dependencia con ninguno de los modos instrumentales y sólo mostraran diferencias en sus intensidades absolutas. Por consiguiente, la falta de trilinealidad mostrada en estos datos lleva consigo una deficiencia de rango en el modo espectral de emisión. Es importante mencionar, además, que ante posibles corrimientos en los tiempos de retención de los analitos se observarían diferencias entre los perfiles temporales de diferentes muestras, pero, adicional-

mente, los perfiles espectrales de excitación mostrarán diferencias para las diferentes muestras.

Es notable la alta complejidad que presentan los datos de tercer orden generados con esta metodología debido a que la fuerte dependencia entre los modos instrumentales conlleva una falta de trilinealidad extrema. Desafortunadamente, los algoritmos disponibles actualmente para el modelado de datos de tercer orden no resultan eficientes para el correcto análisis y descomposición de estos datos. En adición, y a diferencia de lo mencionado para la *Metodología II*, no se encontraron herramientas de pre-procesamiento que permitan corregir los efectos ocurridos. En consecuencia, el análisis de estos datos requiere desarrollar nuevos algoritmos quimiométricos que permitan el modelado de datos, es decir, algoritmos que permitan la descomposición de datos de tercer orden no trilineales, como ocurre análogamente para datos de segundo orden no bilineales generados, por ejemplo, por espectroscopia de fluorescencia sincrónica total (Schenone *et al.*, 2013).

Aplicación analítica. Con base en las observaciones realizadas, la Metodología I se muestra como la estrategia más eficiente para la generación de datos de tercer orden. Con el objetivo de estudiar la aplicabilidad del método y de la metodología de generación de datos seleccionada para llevar a cabo un análisis cuantitativo se propuso el desarrollo de un método analítico modelo basado en la cuantificación de 3 f-QUI en agua de consumo. El modelado de los datos se realizó utilizándose los algoritmos APARAFAC y MCR-ALS, este último por modelado de hiper-matrices TEEM desdobladas, considerándose la falta de cuadrilinealidad que presenta el arreglo de datos de 4 vías.

Modelado quimiométrico por MCR-ALS. Una vez obtenidas la hiper-matriz correspondiente a cada muestra, se llevó a cabo el modelado quimiométrico utilizando ext-MCR-ALS según se describe a continuación:

Cada EEM de dimensiones $K \times L$ (siendo K y L el número de longitudes de onda de emisión y excitación, respectivamente) correspondientes a las diferentes fracciones recolectadas, se desdoblaron de manera de obtener vectores con dimensiones $LK \times 1$. Posteriormente, se construyó una matriz de dimensiones $J \times LK$ para cada muestra, con J fracciones, o J tiempos de retención. Construidas las matrices bidimensionales desdobladas para cada muestra, se construyó una matriz de datos \mathbf{D} aumentada en

columnas de dimensiones $(I+1)J \times LK$, para I muestras de calibración y 1 muestra de predicción, ya sea de validación o de agua de consumo fortificada, obteniéndose, así, un arreglo de 2-vías aumentado.

Durante el ajuste por cuadrados mínimos alternantes del modelado quimiométrico, se utilizaron restricciones de unimodalidad y correspondencia entre especies para el modo temporal y restricciones de no-negatividad para los modos temporal y espectral.

Posterior al modelado quimiométrico se obtuvieron los perfiles correspondientes a los tiempos de retención para cada analito en las diferentes muestras (**C**), los correspondientes a las matrices de excitación-emisión desdobladas (**S^T**) para los N componentes presentes en las muestras y una matriz de residuos (**E**). Las EEM correspondientes a cada componente puro $K \times L$ fueron luego obtenidas por reconstrucción a partir de los perfiles contenidos en **S^T** y utilizadas para la identificación de los componentes reales.

Finalmente, a partir de los perfiles temporales contenidos en **C** se obtuvieron las contribuciones individuales de los analitos correspondientes para cada muestra, que fueron luego utilizadas para el análisis predictivo.

Previo al inicio de la descomposición de los datos, se determinó el número de componentes espectralmente activos presentes en la matriz **D** aplicando SVD, además de obtener las estimaciones espectrales iniciales mediante una metodología basada en SIMPLISMA (Windig *et al.*, 1991). En la Figura 3.51 se muestran los perfiles espectrales (**S**) y temporales (**C**) obtenidos como resultado de la descomposición bilineal por MCR-ALS para una muestra conteniendo los 3 analitos y, también, los perfiles espectrales de excitación y emisión correspondientes a cada analito.

Modelado quimiométrico por APARAFAC. En principio, la aplicación de este modelo quimiométrico solo involucraría un paso de inicialización debido a que las restricciones no serían necesarias dado que la descomposición de un arreglo trilineal de 3-vías es, generalmente, único (Olivieri *et al.*, 2014b). Sin embargo, resultados preliminares basados en la descomposición de datos, sin utilizar restricciones, resultaron inadecuados. Por lo tanto, a fin de mejorar los resultados y obtener perfiles con significancia físico-química, y teniendo en cuenta que APARAFAC implementa el mismo tipo de restricciones que las utilizadas en MCR-ALS, se aplicaron las restricciones de unimodalidad, no-negatividad y correspondencia entre especies al modo temporal.

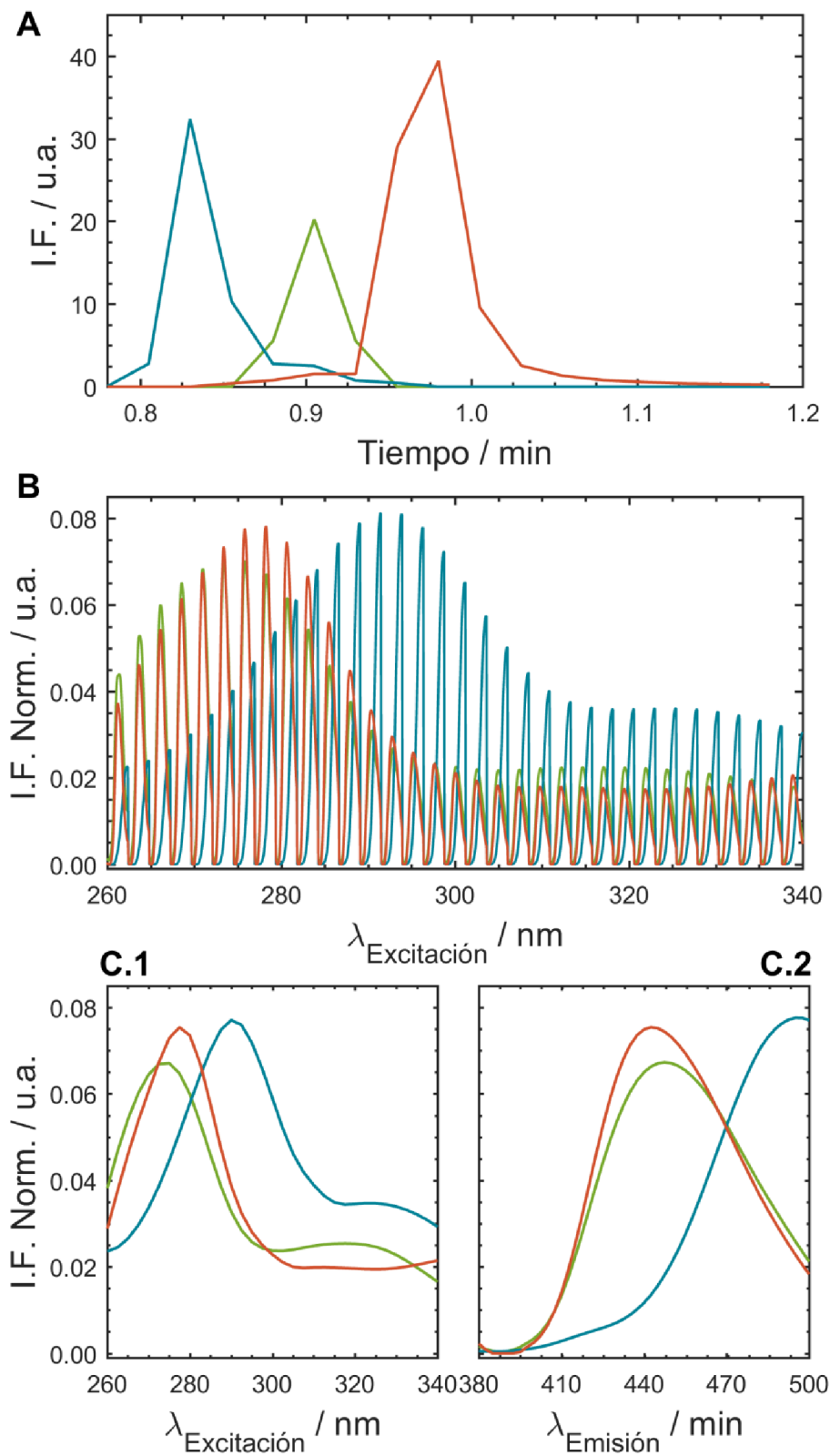


Figura 3.51. Perfiles temporales (A) y espectrales (B) obtenidos por modelado quimiométrico con MCR-ALS de una muestra conteniendo 3 f-QUI. Los perfiles espectrales (B) se corresponden a las matrices de excitación-emisión desdobladas de OFL (azul), CPF (verde) y DNF (rojo). Los perfiles espectrales de excitación (C.1) y emisión (C.2) se obtuvieron a partir de las matrices de excitación-emisión reconstruidas de cada componente

En primer lugar se construyó un arreglo de 3-vías aumentado **X** apilándose una matriz de datos correspondiente a la muestra de validación o de agua de consumo, según corresponda, y las matrices de datos correspondientes a las muestras de calibración. El número de componentes espectroscópicamente activos se obtuvo mediante SVD y la inicialización del proceso iterativo ALS se realizó utilizando los espectros de excitación y emisión reales de los analitos puros.

Finalizada la descomposición quimiométrica se obtuvieron los perfiles correspondientes a los tiempos de retención de cada analito en las diferentes muestras (**A**), los correspondientes a los espectros de excitación (**B**) y los correspondientes para los espectros de emisión (**C**). A diferencia de MCR-ALS, los perfiles espectrales de excitación y emisión se obtuvieron separadamente en matrices individuales. Posteriormente, y análogamente a un modelado por MCR-ALS, se utilizó la información contenida en la matriz **A** para construir un modelo de calibración y obtener las contribuciones de los analitos en las muestras de validación o de agua de consumo.

En la Figura 3.52 se muestran los perfiles espectrales de excitación, de emisión y temporales obtenidos para cada componente como resultado de la descomposición por APARAFAC de una muestra conteniendo 3 analitos y 2 interferentes.

Análisis cuantitativo. A fin de comparar el desempeño de los dos modelos quimiométricos utilizados para el modelado de datos de tercer orden, en términos de capacidad predictiva y estimación de cifras de mérito, se realizó un estudio de predicción en diferentes muestras de validación y muestras de agua de consumo fortificadas. En la Tabla 3.11 y Tabla 3.12 se resumen los valores de predicción y recuperación obtenidos para las 10 muestras de validación y las 6 muestras de agua de consumo en presencia de interferentes.

Los resultados de predicción expuestos en la Tabla 3.11 y Tabla 3.12 obtenidos por modelado con APARAFAC muestran una satisfactoria coincidencia a los obtenidos por MCR-ALS. Adicionalmente, se obtuvieron valores de REP % aceptables para ambos modelos.

Cifras de mérito analíticas. Se estimaron las cifras de mérito analíticas del método desarrollado, para ambos modelos, y, además, se realizó un análisis comparativo entre las cifras de mérito estimadas para el modelo de tercer orden desarrollado y un modelo de segundo orden, este último con modelado por MCR-ALS y PARAFAC. Es importante mencionar que si bien las estimaciones de las cifras de mérito para un

método analítico basado en modelado de datos por MCR-ALS se obtienen a partir de expresiones matemáticas bien definidas y estudiadas (Bauza *et al.*, 2012), las expresiones apropiadas para la estimación de cifras de mérito para un modelo APARAFAC aún no han sido desarrolladas (Olivieri *et al.*, 2012). Por lo tanto, en este trabajo se utilizó una extensión de las expresiones derivadas para PARAFAC (Olivieri *et al.*, 2012), teniendo en cuenta que pueden presentarse sobreestimaciones de las cifras obtenidas.

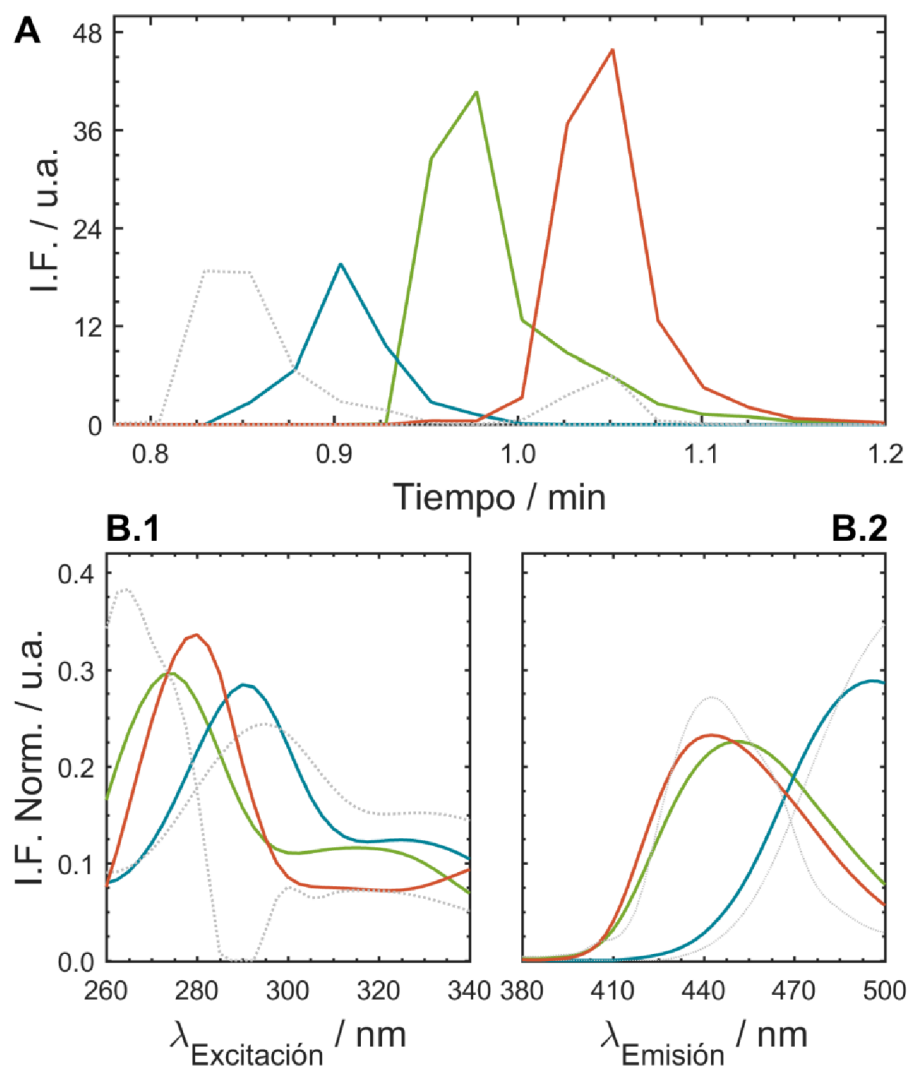


Figura 3.52. Perfiles temporales (A) y espectrales de excitación (B.1) y emisión (B.2) obtenidos por modelado quimiométrico con APARAFAC de una muestra conteniendo OFL (azul), CPF (verde) y DNF (rojo), en presencia de 2 interferentes (línea de puntos gris).

Tabla 3.11. Valores de concentración predicha para para las 3 f-QUI en muestras de validación determinados por modelado con MCR-ALS y APARAFAC ^a

Muestra	OFL			CPF			DNF		
	NOMINAL	PREDICHA		NOMINAL	PREDICHA		NOMINAL	PREDICHA	
		MCR-ALS	APARAFAC		MCR-ALS	APARAFAC		MCR-ALS	APARAFAC
M01	20.0	21.1	20.1	90.0	99.3	92.5	25.0	27.1	23.2
M02	20.0	19.3	19.7	150.0	131.0	121.7	15.0	16.4	15.9
M03	60.0	51.1	52.1	30.0	44.9	58.5	5.0	5.3	5.1
M04	100.0	101.0	99.7	90.0	95.8	90.9	5.0	8.6	8.9
M05	60.0	68.1	70.1	150.0	144.8	147.0	25.0	28.4	28.8
M06	100.0	98.9	99.1	150.0	132.7	136.9	15.0	17.8	18.2
M07	100.0	104.1	101.0	30.0	21.0	22.0	5.0	7.4	7.6
M08	20.0	31.0	31.7	30.0	58.0	51.7	2.0	4.0	4.0
M09	60.0	45.3	55.2	30.0	19.8	25.6	8.0	9.6	9.4
M10	60.0	55.1	72.3	60.0	54.2	44.3	2.0	5.3	2.6
REP % ^b		14.5	13.8		19.0	21.5		19.9	19.1
\bar{R}_{exp} ^c		102.7	107.4		105.8	107.4		146.0	131.5

^a Las concentraciones están dadas en ng mL⁻¹;

^b REP %: error relativo de predicción, expresado en %; $REP \% = 100 \times \frac{\sqrt{\frac{1}{l} \sum_i (c_{nom} - c_{pred})^2}}{\bar{c}}$, para l= 10;

^c \bar{R}_{exp} : valor de recuperación media, expresado en %.

Tabla 3.12. Valores de concentración encontrados de OFL, CPF y DNF en muestras de agua de consumo fortificadas aplicando el método desarrollado y utilizando modelado con MCR-ALS y APARAFAC ^a

<i>Muestra</i> ^b	<i>OFL</i>			<i>CPF</i>			<i>DNF</i>		
	ADICIONADA	ENCONTRADA		ADICIONADA	ENCONTRADA		ADICIONADA	ENCONTRADA	
		MCR-ALS	APARAFAC		MCR-ALS	APARAFAC		MCR-ALS	APARAFAC
MN_01	20.0	30.6	17.4	30.0	26.3	25.0	3.5	3.2	2.9
MN_02	60.0	81.2	63.8	90.0	78.6	91.2	5.5	7.5	7.6
MR_01	60.0	61.9	65.3	90.0	86.6	89.5	2.2	2.7	2.6
MR_02	40.0	32.2	26.2	60.0	60.5	62.9	9.0	12.8	12.4
MM_01	20.0	12.5	19.1	30.0	19.7	17.0	2.2	1.9	1.9
MM_02	40.0	41.1	49.9	60.0	91.1	81.3	9.0	9.5	9.2
\bar{R}_{exp} ^c		106	98		98	97		116	113

^a Las concentraciones están dadas en ng mL⁻¹. Cada valor de concentración encontrado se expresa como el valor medio de tres réplicas de preparación;

^b MN: muestra de agua de napa subterránea apta para consumo; MR: muestra de agua de red; MM: muestra de agua mineral comercial;

^c \bar{R}_{exp} , valor de recuperación media, expresado en %.

En la Tabla 3.13 se muestran las cifras de mérito estimadas para el método analítico desarrollado utilizando los diferentes algoritmos quimiométricos para el modelado de datos de tercer orden.

Tabla 3.13. Cifras de mérito analíticas obtenidas para los diferentes modelos quimiométricos de datos de tercer orden ^a

Cifra de mérito	OFL		CPF		DNF	
	MCR-ALS ^a	APARAFAC ^b	MCR-ALS	APARAFAC	MCR-ALS	APARAFAC
SEN	10.4	21.0	2.7	20.0	22.9	83.0
SEL	0.68	0.65	0.21	0.29	0.88	0.30
LOD	0.25	0.20	0.99	0.15	0.12	0.02
LOQ	0.75	0.60	2.97	0.47	0.36	0.08

^a Valores estimados según referencia (Bauza *et al.*, 2012), dado en ng mL⁻¹;
^b Valores estimados según referencia (Olivieri *et al.*, 2012), dado en ng mL⁻¹.

Adicionalmente, la Tabla 3.14 exponen las cifras de mérito estimadas para el método analítico desarrollado utilizando los MCR-ALS y PARAFAC para el modelado de datos de segundo orden.

Tabla 3.14. Cifras de mérito obtenidas para OFL utilizando MCR-ALS y PARAFAC/APARAFAC para el modelado de datos de segundo y tercer orden

Cifra de mérito	MCR-ALS ^b		PARAFAC ^{a,c}	
	Segundo orden	Tercer orden	Segundo orden	Tercer orden
SEN	5.2	10.4	7.6	21.0
SEL	0.23	0.68	0.25	0.65
LOD	0.4	0.25	6.9	0.20
LOQ	1.1	0.75	21.0	0.60

^a Para el modelado de datos de segundo orden se utilizó PARAFAC y para el modelado de datos de tercer orden se utilizó APARAFAC;
^b Valores estimados según referencia (Bauza *et al.*, 2012);
^c Valores estimados según referencia (Olivieri *et al.*, 2012).

Es destacable el aumento en los valores de SEN y las mejoras en los LOD y LOQ observadas para el método analítico utilizando modelado por APARAFAC, en comparación con los obtenidos por MCR-ALS, para datos de tercer orden. Además, es

importante destacar el aumento de la SEN y la drástica reducción de los LOD y LOQ al incrementar el orden o dimensión de los datos utilizando PARAFAC y APARAFAC para datos de segundo y tercer orden, respectivamente. Sin embargo, las cifras de mérito estimadas para datos de segundo y tercer orden modelados con MCR-ALS no presentaron diferencias notables. Por otro lado, la marcada diferencia entre los valores de LOD y LOQ obtenidos para MCR-ALS y PARAFAC para datos de segundo orden radica, principalmente, a la ruptura de la trilinealidad a causa de diferencias en los perfiles temporales de diferentes muestras, inconveniente que puede ser fácilmente superado por modelado con MCR-ALS, pero no con PARAFAC.

Estas observaciones se basan, principalmente, en que el modelado quimiométrico de datos de tercer orden utilizando APARAFAC muestra varias ventajas en comparación con MCR-ALS, destacándose la posibilidad de procesar los datos de tercer orden en su estructura tridimensional original, es decir, sin necesidad de desdoblar matrices de datos ni utilizar arreglos de datos con dimensiones disminuidas, y a la posibilidad de superar la falta de cuadrilinealidad, lo que lleva consigo una mejora en las predicciones analíticas y una mejora en las cifras de mérito. Adicionalmente, APARAFAC explota la ventaja de segundo orden aún en presencia de diferencias en los perfiles temporales entre distintas muestras, de manera similar a MCR-ALS. En consecuencia, APARAFAC se muestra como una alternativa adecuada para el procesamiento de datos cromatográficos de tercer orden, lográndose resultados aceptables para muestras multi-componentes, aún en presencia de componentes no calibrados.

Se debe mencionar, también, que se utilizó PARAFAC2 como algoritmo quimiométrico para el modelado de datos, obteniéndose resultados poco satisfactorios en calidad de los valores de predicción obtenidos para las muestras de validación y de agua de consumo. Si bien PARAFAC2 es un algoritmo más flexible que PARAFAC con respecto a las diferencias observadas en los perfiles temporales entre diferentes muestras, no resulta un algoritmo eficiente para el modelado de datos de muestras multi-componentes en presencia de componentes no modelados (Bortolato *et al.*, 2014).

CONCLUSIÓN

Se analizaron 3 metodologías analíticas diferentes para la generación de datos de tercer orden basadas en un procedimiento cromatográfico acoplado a detección de matrices de excitación-emisión de fluorescencia. Se evaluaron los requerimientos ins-

trumentales para cada metodología individualmente en virtud de la complejidad, costos y accesibilidad de equipamiento. Adicionalmente, se realizó un análisis detallado de los datos generados, evaluándose cualitativamente sus propiedades desde el punto de vista de trilinealidad y complejidad de análisis en el pre-procesamiento y procesamiento de datos. Finalmente, se modelaron los datos generados mediante un algoritmo de descomposición trilineal (*TLD*, *PARAFAC*), sin realizar etapas de pre-procesamiento, con el objetivo de evaluar los efectos causados por la pérdida de trilinealidad, y, además, se propusieron alternativas de pre-procesamiento y procesamiento de datos que permitan superar este fenómeno y así obtener datos inequívocos.

La *Metodología I*, basada en la recolección de fracciones a la salida del cromatógrafo, requiere de equipamiento de baja complejidad, necesitándose de un dispositivo que permita la recolección de fracciones en placas de ELISA, el que puede ser construido en el laboratorio de manera simple y económica. El tiempo consumido para el análisis de una muestra estuvo limitado, principalmente, por la etapa de detección, en dependencia a los parámetros instrumentales y características del espectrofotómetro de fluorescencia utilizado. Los datos generados mostraron perfecta trilinealidad debido a la independencia que presentan los modos instrumentales entre sí y en virtud de la propiedad de bilinealidad/trilinealidad característica que muestran las EEM. Los resultados obtenidos de la descomposición trilineal resultaron altamente satisfactorios, obteniéndose perfiles temporales y espectrales con extrema similitud a los perfiles cromatográficos y a los espectros reales, respectivamente. Por otro lado, debido a las diferencias observadas entre tiempos de retención de los analitos entre diferentes muestras, un arreglo de 4-vías mostró falta de cuadrilinealidad. Sin embargo, se demostró que datos cromatográficos no cuadrilineales pueden ser correctamente analizados utilizándose los algoritmos quimiométricos *APARAFAC* o *MCR-ALS*.

La *Metodología II* se muestra como una alternativa que requiere de un cromatógrafo complejo que disponga de inyector automático y un detector de fluorescencia de barrido rápido. Debido a que se requiere de un alto número de inyecciones por muestra, el tiempo de análisis para cada muestra es considerablemente elevado y solo puede ser reducido a expensas de una pérdida de información espectral. Esto se debe a que el espectro de excitación, en este caso, se construyó a partir de las múltiples inyecciones realizadas a diferentes longitudes de onda de excitación. Adicionalmente, esta metodología demanda un alto consumo de solventes y de muestra a causa de las múltiples inyecciones, ocasionado un gran impacto medioambiental y un aumento consecuente de los costos. Se demostró, además, que pequeñas diferencias en los tiem-

pos de retención entre diferentes corridas cromatográficas para una misma muestra llevan consigo modificaciones en los espectros de excitación resultantes, indicándose una dependencia directa entre ambos modos instrumentales. Este efecto, traducido como una falta de trilinealidad en los datos de tercer orden, se muestra como un obstáculo que debe ser superado para la resolución del sistema y la obtención de resultados inequívocos. Sin embargo, en trabajos encontrados en la bibliografía, no se ha evaluado el efecto de la falta de trilinealidad por diferencias entre tiempos de retención en los perfiles cromatográficos de una misma muestra, si bien se reporta falta de cuadrilinealidad en un arreglo de 4-vías ocasionado por el mismo fenómeno, ignorándose, de esta manera, los errores ocasionados por la pérdida de trilinealidad en cada hipermatriz individual (Lozano *et al.*, 2013; Montemurro *et al.*, 2016). Aquí, se propusieron alternativas de pre-procesamiento de datos que permitirían devolver la trilinealidad de los datos utilizando, por ejemplo, algoritmos que corrijan las diferencias entre los tiempos de retención de una misma muestra. Luego, la pérdida de cuadrilinealidad observada en un arreglo de datos de 4-vías puede ser fácilmente superada mediante el uso de algoritmos quimiométricos como U-PLS/RTL, APARAFAC o MCR-ALS.

Finalmente, se presentó la *Metodología III* como una nueva propuesta para la generación de datos cromatográficos de tercer orden, la que resultaría conveniente desde el punto de vista instrumental, tiempo de análisis, costos, consumo de solventes y de muestra e impacto ambiental. Sin embargo, los datos generados mostraron una alta complejidad en virtud de la fuerte dependencia entre los modos instrumentales correspondientes a los espectros de excitación y tiempo de retención. Desafortunadamente, estos datos no pudieron ser analizados con los algoritmos quimiométricos desarrollados hasta el momento y, adicionalmente, no se encontraron estrategias de pre-procesamiento eficientes que permitan recuperar la trilinealidad de los datos. De esta manera, el desarrollo de un nuevo algoritmo quimiométrico que permita el análisis de este tipo de datos resulta, particularmente, necesario.

Sobre la base de todas las observaciones y resultados obtenidos, *la Metodología I* se presenta como una alternativa eficiente y sencilla para la obtención de datos cromatográfico de tercer orden trilineales.

Utilizando la *Metodología I*, se desarrolló un método analítico modelo para la determinación de 3 f-QUI en agua de consumo, obteniéndose resultados altamente satisfactorios. Se utilizaron 2 algoritmos quimiométricos diferentes para el modelado de los datos, MCR-ALS y APARAFAC, realizándose un estudio comparativo en base a las capacidades predictivas y de cifras de mérito estimadas. Se demostró que es posible

obtener mejores cifras de mérito utilizando el algoritmo APARAFAC debido, principalmente, a que se mantiene la estructura tridimensional original de los datos, a diferencia de MCR-ALS que, al utilizar datos desdoblados, se utilizan datos de dimensiones inferiores, es decir, datos de segundo orden desdoblados y aumentados en columna.

Finalmente, es importante notar que, a diferencia de los datos cromatográficos de segundo orden con detección espectral de emisión de fluorescencia, con datos cromatográficos de tercer orden se pueden determinar múltiples analitos que presenten grandes diferencias espectrales, debido a que se registran EEM y no espectros de emisión a una longitud de onda de excitación de compromiso, pudiendo así incluir los máximos de intensidad de todos los componentes de interés y disminuyendo, en consecuencia, los niveles de concentración límites detectables.

G4

Gonclusiones

4. CONCLUSIONES GENERALES

"I have not failed. I have just found 10.000 ways that don't work"

Edison, T. A.

Sobre la base de los resultados expuestos se pueden reportar diferentes observaciones desde el punto de vista analítico, instrumental y/o quimiométrico. Sin embargo, desglosar un método analítico y evaluar las etapas individualmente contradice, paradójicamente, al concepto de análisis multivariado. El desarrollo de un nuevo método analítico para el análisis de sistemas complejos no requiere de una planificación minuciosa de los experimentos o de un procedimiento analítico determinado o de un análisis de datos definido, sino que, para evaluar su desarrollo, un método analítico debe ser evaluado en su conjunto, desde su planificación hasta la obtención de los resultados.

En este trabajo se demostró que la selección del algoritmo para el procesamiento de los datos obtenidos para un método analítico determinado depende del propósito planteado inicialmente, de la técnica analítica utilizada, del método de detección empleado y de las propiedades inherentes del sistema. En otras palabras, no es posible empezar la evaluación de un método desde uno de los extremos. Sin embargo, se

deberá establecer un orden lógico que facilite la interpretación y la comprensión del análisis.

Para análisis cuantitativo, se desarrollaron métodos analíticos separativos basados en metodologías de cromatografía líquida de alto rendimiento y electroforesis capilar zonal acopladas a detección espectral de emisión de fluorescencia y absorción UV, respectivamente, para la determinación de contaminantes emergentes en matrices acuosas. Se observó que un correcto diseño de experimentos, una evaluación previa en detalle del sistema en estudio, la aplicación del algoritmo adecuado para el modelado quimiométrico de los datos y el desarrollo de una minuciosa estrategia de post-análisis resultan una excelente combinación para superar la complejidad de los sistemas en estudio y obtener resultados altamente satisfactorios para el análisis simultáneo de diferentes constituyentes presentes en muestras acuosas. Además, se demostró que para mejorar las propiedades analíticas de un método, por ejemplo, disminuir los límites de detección o aumentar la sensibilidad, es posible utilizar etapas de pre-concentración simples o emplear aditivos que permitan la exaltación de la señal nativa de los componentes. Asimismo, es necesario considerar las propiedades del sistema de calibración y el modelo matemático empleado a fin de obtener estimaciones realistas de las cifras de mérito analíticas.

Por otro lado, el desarrollo de un método basado en el análisis por inyección en flujo con generación de gradiente de pH y modelado de datos de segundo orden permitió concluir que la estrategia instrumental es útil no solo para la determinación de fármacos en sistemas biológicos complejos, sino también, para la obtención de información físico-química acerca del sistema de equilibrio químico en estudio.

Con otros propósitos, se demostró que los modelados quimiométricos de datos de orden superior permiten analizar y estudiar sistemas biológicos desde el punto de vista químico-conformacional. La combinación entre una metodología analítica descriptiva, como espectroscopia infrarroja o dicroísmo circular, y un análisis quimiométrico apropiado permite inferir acerca de los cambios estructurales y el comportamiento de determinadas macromoléculas bajo perturbaciones externas o procesos de desnaturalización inducidos. Sin embargo, es importante considerar la necesidad de un análisis exploratorio previo, a fin de evitar errores en la selección del modelo quimiométrico que pudiesen derivar a una malinterpretación de los resultados.

En cuanto al uso de datos de tercer orden para fines predictivos, se demostró que la generación de los datos de tercer orden no solo depende de la instrumentación utilizada o de la técnica analítica seleccionada sino que depende de la disposición y/o

configuración instrumental usada. Se demostró, además, que los datos generados presentan diferentes estructuras y propiedades matemáticas que deben ser consideradas al momento de seleccionar el modelo quimiométrico a implementar e incluso se comprobó que es necesario el desarrollo de nuevos algoritmos quimiométricos para el modelado de datos multidimensionales.

Por último, pero no menos importante, desde el punto de vista quimiométrico, se puede concluir que no existe un algoritmo único que permita la resolución completa y absoluta del problema, sino que hay algoritmos que permiten obtener diferente información y que existen metodologías que resultan más adecuadas para el análisis de un sistema particular. Se demostró que la estructura de los datos multidimensionales generados es la primera consideración que permite inferir sobre el modelo quimiométrico a implementar. Sobre estas observaciones se puede concluir que para la selección del algoritmo quimiométrico se debe considerar la estructura de los datos multidimensionales generados y no sobreestimar o pasar por alto sus propiedades matemáticas inherentes.

Por todo esto, se puede concluir que si bien la Quimiometría es una poderosa herramienta que permite, entre otras bondades, la obtención de información valiosa acerca de un sistema de estudio, la reducción de tiempos de análisis y la disminución del consumo de reactivos, no puede prescindir del criterio analítico y de la correcta interpretación del analista. Además, es importante no perder de vista que la selección de la técnica analítica a implementar para el desarrollo de un nuevo método analítico repercutirá sobre la información que se obtiene del sistema y condiciona, además, el uso de las herramientas quimiométricas para el análisis. Por consiguiente, y recuperando los conceptos iniciales, *“un método analítico debe ser evaluado en su conjunto, desde su planificación hasta la obtención de los resultados”*, entendiéndose la interconexión entre las etapas involucradas en todo el procedimiento.



05

R

Referencias



5. REFERENCIAS

- Abraham, S. A., *et al.* In *Methods Enzymol*; Academic Press, 2005, pp 71-97.
- Álvarez-Cedrón, L., *et al.* *J Chromatogr B* **1999**, *721*, 271-278.
- Ando, H. Y. and T. Heimbach. *J Pharm Biomed Anal* **1997**, *16*, 31-37.
- Armenta, S., *et al.* *TrAC, Trends Anal Chem* **2008**, *27*, 497-511.
- Arroyo-Manzanares, N., *et al.* *Anal Methods* **2015**, *7*, 253-259.
- Babić, S., *et al.* *TrAC, Trends Anal Chem* **2007**, *26*, 1043-1061.
- Bagińska, K., *et al.* *J Pept Sci* **2008**, *14*, 283-289.
- Bal Ram, S. In *Infrared Analysis of Peptides and Proteins*; American Chemical Society, 1999, pp 2-37.
- Barth, A. *BBA-Bioenergetics* **2007**, *1767*, 1073-1101.
- Bauza, M. C., *et al.* *Anal Chem* **2012**, *84*, 8697-8706.
- Bezemer, E. and S. C. Rutan. *Chemom Intell Lab Syst* **2001**, *59*, 19-31.
- Bogomolov, A., *et al.* In *Progress in Chemometrics Research*, A. L. Pomerantsev, Ed.; Nova Science Publishers, Inc.: New York, USA, 2005, pp 119-135.
- Booksh, K. S. and B. R. Kowalski. *Anal Chem* **1994**, *66*, 782A-791A.
- Borges, A., *et al.* *Anal Chim Acta* **2005**, *544*, 159-166.
- Bortolato, S. A., *et al.* *Chemom Intell Lab Syst* **2015**, *141*, 1-11.
- Bortolato, S. A. and A. C. Olivieri. *Anal Chim Acta* **2014**, *842*, 11-19.
- Brandstetter, M., *et al.* *Analyst* **2010**, *135*, 3260-3265.
- Brandstetter, M., *et al.* (2013a). *Measures for optimizing pulsed EC-QC laser spectroscopy of liquids and application to multi-analyte blood analysis*.
- Brandstetter, M. and B. Lendl. *Sens Actuators, B* **2012**, *170*, 189-195.
- Brandstetter, M., *et al.* *Appl Phys B* **2013b**, *110*, 233-239.

- Bro, R. *Multi-way Analysis in the Food Industry*. Royal Veterinary and Agricultural University, Copenhagen, Denmark, 1997.
- Cañada-Cañada, F., et al. *J Chromatogr A* **2009**, *1216*, 4868-4876.
- Carrier, D., et al. *Biopolymers* **1990**, *29*, 837-844.
- Carvalho, C., et al. *Curr Med Chem* **2009**, *16*, 3267-3285.
- Cattell, R. B. *Psychometrika* **1944**, *9*, 267-283.
- Clayton, C. A., et al. *Anal Chem* **1987**, *59*, 2506-2514.
- Cooley, J. W. and J. Tukey, W. *Math Comp* **1965**, *19*, 297-301.
- Crawford, L. M., D. o. h. a. h. services, Ed.; FDA: Food and Drug Administration, HHS.: Silver Spring, Meryland, 2005, p 44105.
- Curl, R. F., et al. *Chem Phys Lett* **2010**, *487*, 1-18.
- Czarnik-Matusewicz, B. and S. Pilorz. *J Mol Struct* **2006**, *799*, 211-220.
- Chi, E. Y., et al. *Pharm Res* **2003**, *20*, 1325-1336.
- Chirgadze, Y. N. and N. A. Nevskaya. *Biopolymers* **1976**, *15*, 607-625.
- Danzar, K. and C. L.A. *Pure Appl Chem* **1998**, *70*, 993-1001.
- Davidson, B. and G. D. Fasman. *Biochemistry* **1967**, *6*, 1616-1629.
- de Juan, A., et al. *Chemom Intell Lab Syst* **2000**, *54*, 123-141.
- de Juan, A., et al. In *Comprehensive Chemometrics*, S. D. Brown, R. Tauler and B. Walczak, Eds.; Elsevier: Oxford, 2009, pp 325-344.
- de Oliveira Mendes, T., et al. *Food Chem* **2016**, *213*, 647-653.
- DeFlores, L. P., et al. *J Am Chem Soc* **2009**, *131*, 3385-3391.
- Derringer, G. and R. Suich. *J Quality Technol* **1980**, *12*, 214-219.
- Dong, A., et al. *Biochemistry* **1990**, *29*, 3303-3308.
- Dong, A., et al. *Anal Biochem* **2006**, *351*, 282-289.
- Dong, A., et al. *Arch Biochem Biophys* **1998**, *355*, 275-281.
- Dong, A., et al. *Arch Biochem Biophys* **2000**, *383*, 148-155.
- Dong, A., et al. *J Pharm Sci* **1995**, *84*, 415-424.
- Drevenšek, P., et al. *J Inorg Biochem* **2006**, *100*, 1705-1713.
- Dzwolak, W., et al. *Biopolymers* **2004**, *73*, 463-469.
- Dzwolak, W. and V. Smirnovas. *Biophys Chem* **2005**, *115*, 49-54.
- Eilers, P., Boelens, H. Amsterdam, The Netherlands, 2005.
- El-Didamony, A. M. *Luminescence* **2011**, *26*, 112-117.
- EMA (2011). "Guideline on bioanalytical method validation." Retrieved 31-08-2015, 2015, from http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2011/08/WC500109686.pdf.
- Escandar, G. M., et al. *Anal Chim Acta* **2014**, *806*, 8-26.
- Escandar, G. M., et al. *TrAC, Trends Anal Chem* **2007**, *26*, 752-765.
- Fabian, H. and W. Mäntele. In *Handbook of Vibrational Spectroscopy*; John Wiley & Sons, Ltd, 2006.
- FDA and DHHS. In *Guidance for the Industry: Fish and Fishery Products Hazards and Controls Guidance* Florida, USA, 2011, pp 183-208.
- Feng, X., et al. *Anal Methods* **2013**, *5*, 5286-5293.
- Fragoso, W., et al. *Anal Chim Acta* **2016**, *933*, 43-49.
- Gałaszka, A., et al. *TrAC, Trends Anal Chem* **2013**, *50*, 78-84.
- García, M. D. G., et al. *J Chromatogr A* **2009**, *1216*, 5489-5496.
- Giuliani, P., et al. *J Chromatogr B* **2016**, *1009–1010*, 114-121.

- Glasoe, P. K. and F. A. Long. *J Phys Chem* **1960**, *64*, 188-190.
- Greenfield, N. J. *Nat Protoc* **2007a**, *1*, 2733-2741.
- Greenfield, N. J. *Nat Protoc* **2007b**, *1*, 2876-2890.
- Grigsby, J. J., *et al.* *Biophys Chem* **2002**, *99*, 107-116.
- Hajian, R., *et al.* *Journal of Pharmaceutical Analysis* **2016**.
- Han, Y., *et al.* *J Pharm Biomed Anal* **2005**, *38*, 528-531.
- Harrison, R. S., *et al.* In *Rev Physiol, Biochem Pharmacol*, S. G. Amara *et al.*, Eds.; Springer Berlin Heidelberg: Cham, Switzerland, 2007, pp 1-77.
- He, K. and L. Blaney. *J Hazard Mater* **2015**, *282*, 96-105.
- Herrera-Herrera, A. V., *et al.* *J Pharm Biomed Anal* **2013**, *75*, 130-137.
- Hortobágyi, G. N. *Drugs* **1997**, *54*, 1-7.
- Ibañez, G. A., *et al.* *Chem Educ* **2007**, *12*, 22-28.
- ICH (2005). "Validation of analytical procedures: text and methodology Q2(R1). ." *International conference on harmonization of technical requirements for registration of pharmaceuticals for human use*. Retrieved 07 January, 2015, from http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q2_R1/Step4/Q2_R1_Guideline.pdf.
- Immaneni, A. and A. J. McHugh. *Biopolymers* **1998**, *45*, 239-246.
- Ioannou, J. C., *et al.* *Food Hydrocolloids* **2015**, *46*, 216-225.
- Jackson, M., *et al.* *BBA-Protein Struct M* **1989**, *998*, 75-79.
- Jasanoff, A. and A. R. Fersht. *Biochemistry* **1994**, *33*, 2129-2135.
- Jellema, R. H. In *Comprehensive Chemometrics*, S. D. Brown, R. Tauler and B. Walczak, Eds.; Elsevier: Oxford, 2009, pp 85-108.
- JiJi, R. D., *et al.* *Biochemistry* **2006**, *45*, 34-41.
- Keller, H. R. and D. L. Massart. *Chemom Intell Lab Syst* **1992**, *12*, 209-224.
- Kelly, S. M., *et al.* *BBA-Proteins Proteom* **2005**, *1751*, 119-139.
- Khanmohammadi, M. *Multiway Calibration Approaches to Handle Problems Linked to the Determination of Emergent Contaminants in Waters*; Nova Publishers: New York, 2015, p 279.
- Khetan, S. K. and T. J. Collins. *Chem Rev* **2007**, *107*, 2319-2364.
- Khodarahmi, R., *et al.* *Protein J* **2009**, *28*, 349-361.
- Kiers, H. A. L. *J Chemom* **2000**, *14*, 105-122.
- Kucheryavskiy, S., *et al.* In *Data Handling in Science and Technology*, R. Cyril, Ed.; Elsevier, 2016, pp 53-99.
- Lakowicz, J. R. *Principles of fluorescence spectroscopy*; Springer: New York, 2006.
- Le, M. D., *et al.* *Talanta* **2016**, *160*, 512-520.
- Liu, Q., *et al.* *TrAC, Trends Anal Chem* **2014**, *58*, 10-22.
- López-Cobo, A., *et al.* *LWT-Food Sci Technol* **2016**, *73*, 505-513.
- Lozano, V. A., *et al.* *Chemom Intell Lab Syst* **2013**, *125*, 121-131.
- Lu, H., *et al.* *J Mol Biol* **1997**, *265*, 112-117.
- Lukman, S., *et al.* *Geochim Cosmochim Acta* **2012**, *88*, 199-215.
- Maeder, M. and A. de Juan. In *Comprehensive Chemometrics*, S. D. Brown, R. Tauler and B. Walczak, Eds.; Elsevier: Oxford, 2009, pp 261-274.
- Maeder, M. and A. Zilian. *Chemom Intell Lab Syst* **1988**, *3*, 205-213.
- Malik, A. and R. Tauler. *Chemom Intell Lab Syst* **2014**, *135*, 223-234.
- Martí-Aluja, I. and M. S. Larrechi. *Chemom Intell Lab Syst* **2013**, *127*, 49-54.
- Martí, R. B. and J. F. Baldrich. In *Data Handling in Science and Technology*, H. C. G. M. E. Arsenio Muñoz de la Peña and C. O. Alejandro, Eds.; Elsevier, 2015, pp 7-35.
- Massart, D. L., *et al.* *Data Handling in Science and Technology*; Elsevier Science: Amsterdam, The Netherlands, 1998; Vol. 30.

- Massart, D. L., *et al.* *Chemometrics: a textbook*; Elsevier Science Publishers: Amsterdam, The Netherlands, 1988.
- Matsuo, K., *et al.* *Proteins* **2012**, *80*, 281-293.
- Mbogning Feudjio, W., *et al.* *Talanta* **2014**, *130*, 148-154.
- Meloun, M., *et al.* *J Pharm Biomed Anal* **2016**, *120*, 158-167.
- Montemurro, M., *et al.* *Talanta* **2016**, *154*, 208-218.
- Montemurro, M., *et al.* *Sensor Actuat B-Chem* **2017**, *239*, 397-404.
- Moraes, I., *et al.* *BBA-Biomembranes* **2014**, *1838*, 78-87.
- Müller, M., *et al.* *J Phys Chem* **1996**, *100*, 10810-10825.
- Muñoz de la Peña, A., *et al.* *Fundamentals and Analytical Applications of Multiway Calibration*; Elsevier: Amsterdam, 2015; Vol. 29.
- Murayama, K. and M. Tomida. *Biochemistry* **2004**, *43*, 11526-11532.
- Musil, K., *et al.* *J Pharm Biomed Anal* **2016**, *117*, 240-246.
- Myers, R. H., *et al.* *Response Surface Methodology*, 3era ed.; Wiley & Sons, Inc: New York, Estados Unidos, 2009.
- Navea, S., *et al.* *Anal Chem* **2002**, *74*, 6031-6039.
- Navea, S., *et al.* *Anal Chem* **2006**, *78*, 4768-4778.
- Nielsen, N.-P. V., *et al.* *J Chromatogr A* **1998**, *805*, 17-35.
- Oberg, K. A., *et al.* *Eur J Biochem* **2004**, *271*, 2937-2948.
- Oláh, K. *Period. Polytech-Chem.* **1960**, *4*, 125-140.
- Olivieri, A. C. *Anal Methods* **2012**, *4*, 1876-1886.
- Olivieri, A. C. *Chem Rev* **2014**, *114*, 5358-5378.
- Olivieri, A. C. *Anal Chim Acta* **2015**, *868*, 10-22.
- Olivieri, A. C., *et al.* In *Data Handling in Science and Technology*, H. C. G. G. M. E. Arsenio Muñoz de la Peña and C. O. Alejandro, Eds.; Elsevier, 2015, pp 541-575.
- Olivieri, A. C. and G. M. Escandar. In *Practical Three-Way Calibration*, A. C. Olivieri and G. M. Escandar, Eds.; Elsevier: Boston, 2014a, pp 93-107.
- Olivieri, A. C. and G. M. Escandar. *Practical Three-Way Calibration*; Elsevier: Waltham, USA, 2014b.
- Olivieri, A. C. and N. M. Faber. *Anal Chem* **2012**, *84*, 186-193.
- Ortiz, M. C., *et al.* In *Data Handling in Science and Technology*, A. M. de la Peña, H. C. Goicoechea, G. M. Escandar and A. C. Olivieri, Eds.; Elsevier, 2015, pp 37-81.
- Oshokoya, O. O., *et al.* *Anal Methods* **2014**, *6*, 1691-1699.
- Pallarès, I., *et al.* *J Mol Biol* **2004**, *342*, 321-331.
- Parson, W. W. In *Modern Optical Spectroscopy*, W. W. Parson, Ed.; Springer-Verlag Berlin Heidelberg: Heidelberg, 2015a, pp 383-416.
- Parson, W. W. In *Modern Optical Spectroscopy*, Springer, Ed.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2015b, pp 1-29.
- Pasamontes, A. and P. Callao. *Chemom Intell Lab Syst* **2006**, *83*, 127-132.
- Pasquini, B., *et al.* *Talanta* **2016**, *150*, 7-13.
- Patel, M. N., *et al.* *Pharm Chem J* **2008**, *42*, 687-692.
- Pelton, J. T. and L. R. McLean. *Anal Biochem* **2000**, *227*, 167-176.
- Pellegrini, G. E., *et al.* *Anal Chim Acta* **2004**, *520*, 13-18.
- Pérez-Blanco, J. S., *et al.* *J Chromatogr B* **2014**, *955-956*, 93-97.
- Polzi, L. Z., *et al.* *J Phys Chem B* **2012**, *116*, 3353-3360.
- Pribic, R., *et al.* *Anal Biochem* **1993**, *214*, 366-378.
- Qi, X. L., *et al.* *BBA-Protein Struct M* **1995**, *1248*, 43-49.
- Qing, X.-D., *et al.* *Talanta* **2013**, *103*, 86-94.
- Ramer, G., *et al.* *Anal Chem* **2015**, *87*, 4415-4420.

- Rehman, M. S. U., *et al. Chemosphere* **2015**, *138*, 1045-1055.
- Rezaei-Ghaleh, N., *et al. Biophys J* **2008**, *95*, 4139-4147.
- Rezaei-Ghaleh, N., *et al. Arch Biochem Biophys* **2007**, *457*, 160-169.
- Rodríguez-Díaz, R. C., *et al. Anal Chim Acta* **2003**, *494*, 55-62.
- Rothman, L. S., *et al. J Quant Spectrosc Radiat Transfer* **2013**, *130*, 4-50.
- Rusu, A., *et al. Acta Medica Marisiensis* **2014**, *60*, 109-115.
- Rusu, A., *et al. Environ Chem Lett* **2015**, *13*, 21-36.
- Sadeek, S. A., *et al. Spectrochim Acta A Mol Biomol Spectrosc* **2011**, *78*, 854-867.
- Sajjadi, S. M., *et al. Spectrochim Acta A Mol Biomol Spectrosc* **2016**, *156*, 63-69.
- Sánchez-Ponce, R. and S. C. Rutan. *Chemom Intell Lab Syst* **2005**, *77*, 50-58.
- Santiago, P. S., *et al. Biophys Chem* **2010**, *152*, 128-138.
- Sarroukh, R., *et al. BBA-Biomembranes* **2013**, *1828*, 2328-2338.
- Saurina, J. *TrAC, Trends Anal Chem* **2010**, *29*, 1027-1037.
- Saurina, J., *et al. Anal Chim Acta* **2000**, *408*, 135-143.
- Saurina, J., *et al. Anal Chim Acta* **2001**, *432*, 241-251.
- Savitzky, A. and M. J. E. Golay. *Anal Chem* **1964**, *36*, 1627-1639.
- ScienceDirect. *Tuberculosis* **2008**, *88*, 151-154.
- Schaefer, A. *Environ Sci Technol* **2002**, *36*, 339A-340A.
- Schenone, A. V., *et al. Anal Bioanal Chem* **2013**, *405*, 8515-8523.
- Shi, S., *et al. J Pharm Sci* **2013**, *102*, 2471-2483.
- Siano, G. G. *Desarrollo y aplicación de herramientas quimiométricas para resolución de muestras de origen biológico-químico*. Universidad Nacional del Litoral, Santa Fe, 2013.
- Siegmann-Hegerfeld, T., *et al. Chem. Eng. Trans.* **2013**, *35*, 979-984.
- Speltini, A., *et al. J Sep Sci* **2010**, *33*, 1115-1131.
- Sriraman, S. K., *et al. Eur J Pharm Biopharm* **2016**, *105*, 40-49.
- Sturgeon, R. J. and S. G. Schulman. *J Pharm Sci* **1977**, *66*, 958-961.
- Sturini, M., *et al. Water Res* **2012**, *46*, 5575-5582.
- Sukul, P. and M. Spiteller. In *Reviews of Environmental Contamination and Toxicology*; Springer New York: New York, NY, 2007, pp 131-162.
- Szyc, Ł., *et al. J Mol Liq* **2008**, *141*, 155-159.
- Tacar, O., *et al. J Pharm Pharmacol* **2013**, *65*, 157-170.
- Tajabadi, F., *et al. Talanta* **2016**, *160*, 400-409.
- Tauler, R. *Chemom Intell Lab Syst* **1995**, *30*, 133-146.
- Tauler, R. and A. de Juan. In *Data Handling in Science and Technology*, A. Muñoz de la Peña, H. C. Goicoechea, G. M. Escandar and A. C. Olivieri, Eds.; Elsevier, 2015, pp 247-292.
- Tauler, R., *et al. In Comprehensive Chemometrics*, S. D. Brown, R. Tauler and B. Walczak, Eds.; Elsevier: Oxford, 2009, pp 473-505.
- Tistaert, C., *et al. J Chemom* **2012**, *26*, 474-486.
- Tomasi, G., *et al. J Chromatogr A* **2011**, *1218*, 7832-7840.
- Tomoaia, G., *et al. Colloids Surf B Biointerfaces* **2015**, *135*, 726-734.
- Tong, C., *et al. J Lumin* **2010a**, *130*, 2100-2105.
- Tong, C., *et al. Talanta* **2010b**, *82*, 1858-1863.
- Trevisan, M. and R. J. Poppi. *Anal Chim Acta* **2003**, *493*, 69-81.
- Uivarosi, V. *Molecules* **2013**, *18*, 11153-11197.
- USEPA. In *Analytical Methods: Contaminants of Emerging Concern*; EPA: Washington, D. C., 2007, pp 1-72.

- Vera-Candioti, L., *et al. Electrophoresis* **2008**, *29*, 4527-4537.
- Wang, L., *et al. J Lumin* **2009**, *129*, 90-94.
- Wang, T. and C. Jiang. *Anal Chim Acta* **2006**, *561*, 204-209.
- Wang, W., *et al. In Aggregation of therapeutic proteins*, W. Wang and S. Roberts, Eds.; John Wiley & Sons, Inc.: Hoboken, N.J., 2010, pp 119-205.
- Webb, S., *et al. Toxicol Lett* **2003**, *142*, 157-167.
- Windig, W. In *Comprehensive Chemometrics*, S. D. Brown, R. Tauler and B. Walczak, Eds.; Elsevier: Oxford, 2009, pp 275-307.
- Windig, W. and J. Guilment. *Anal Chem* **1991**, *63*, 1425-1432.
- Wu, H.-L., *et al. In Data Handling in Science and Technology*, H. C. G. G. M. E. Arsenio Muñoz de la Peña and C. O. Alejandro, Eds.; Elsevier, 2015, pp 83-165.
- Wu, H.-L., *et al. J Chemom* **1998**, *12*, 1-26.
- Wysocki, G., *et al. Appl Phys B* **2005**, *81*, 769-777.
- Yáñez-Jácome, G. S., *et al. J Chromatogr A* **2015**, *1405*, 126-132.
- Yang, H., *et al. Nat Protoc* **2015**, *10*, 382-396.
- Yang, X., *et al. J Pharm Biomed Anal* **2016**, *117*, 118-124.
- Yao, Y., *et al. Nat Photon* **2012**, *6*, 432-439.
- Ye, Z., *et al. Anal Chem* **2007**, *79*, 1135-1144.
- Yu, T.-J., *et al. Biopolymers* **1973**, *12*, 2161-2176.
- Zhang, M., *et al. Vib Spectrosc* **2011**, *57*, 319-325.
- Zhang, S., *et al. Chem Eng J* **2016**, *299*, 167-176.
- Zhu, X., *et al. Spectrochim Acta A Mol Biomol Spectrosc* **2008**, *69*, 478-482.

