



Universitat Autònoma de Barcelona

ADVERTIMENT. L'accés als continguts d'aquesta tesi queda condicionat a l'acceptació de les condicions d'ús establertes per la següent llicència Creative Commons:  http://cat.creativecommons.org/?page_id=184

ADVERTENCIA. El acceso a los contenidos de esta tesis queda condicionado a la aceptación de las condiciones de uso establecidas por la siguiente licencia Creative Commons:  <http://es.creativecommons.org/blog/licencias/>

WARNING. The access to the contents of this doctoral thesis it is limited to the acceptance of the use conditions set by the following Creative Commons license:  <https://creativecommons.org/licenses/?lang=en>



Universitat Autònoma
de Barcelona

**Utilización de la espectroscopia NIR en el análisis de
identificación y control de calidad de productos
farmacéuticos liofilizados**

Laia Garcia Uceda

Tesis Doctoral

Programa de Doctorat de Química

Director: Dr. Jordi Coello Bonilla

Departament de Química

Facultat de Ciències

2020

Memòria presentada per aspirar al Grau de Doctor per Laia Garcia Uceda

Laia Garcia Uceda

Vist i plau

Dr. Jordi Coello Bonilla

Bellaterra, 04 de juny del 2020

Aquest projecte de doctorat s'ha elaborat en col·laboració amb l'empresa Laboratorios Reig Jofre en el marc del Pla de Doctorats Industrials, convocatòria DI 2016.



Generalitat de Catalunya
Departament d'Empresa i Coneixement
Secretaria d'Universitats i Recerca

Responsable i tutor del projecte a l'empresa:

Dr. Saša Nikolić

Sant Joan Despí, 04 de juny del 2020

“La imaginación es más importante que el conocimiento.
El conocimiento es limitado. La imaginación envuelve el
mundo.”

Albert Einstein

AGRAÏMENTS

Aquesta tesi doctoral s'ha elaborat gràcies al finançament de les següents institucions:

- Ajut de Doctorat Industrial amb codi 2016 DI 0034 concedit per l'Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR) del Departament d'Empresa i Coneixement de la Generalitat de Catalunya.
- Departament d'Innovació, Universitat i Empresa de la Generalitat de Catalunya.
- Ajut del Ministerio de Economía, Industria y Competitividad del Gobierno de España, MINECO amb el projecte CTQ2016-79696-P (AEI/FEDER, EU).
- I a l'empresa Laboratorios Reig Jofre.

Juntament amb la col·laboració de molta gent que han fet possible que aquesta tesi arribés fins on ha arribat.

El primer cop que vaig escoltar parlar de la quimiometria va ser al 2014 en el treball de final de carrera, on amb l'ajuda del Dr. Manel Alcalà i la Dra. Vanessa Cárdenas vaig aprendre la base d'aquesta extensa disciplina de la química, desconeguda per mi fins aleshores. El treball final de màster, amb l'ajuda del Prof. Santiago Maspoch, Dr. Saša Nikolić i la Dra. Glòria Clua em va permetre veure com aquesta disciplina es podia aplicar al dia a dia d'una empresa, i com de gran podia ser el impacte del desenvolupament d'un petit model. A tots ells els voldria agrair de tot cor la passió amb la que m'han transmès aquets coneixements i que em van permetre agafar les forces necessàries per començar aquesta tesi.

Tot i això, amb tot aquest coneixement només havia descobert la punta del iceberg i durant aquesta tesi i amb la ajuda incansable del Prof. Jordi Coello he tingut l'oportunitat d'endinsar-me i conèixer molt més aquesta disciplina. Jordi, moltes gràcies per tota la ajuda, comprensió i paciència (molta paciència) que m'has ofert durant aquets casi 4 anys, sense tu no hagués pogut portar a bon port aquest projecte.

Una de les coses que més em va fascinar de la quimiometria va ser el gran ventall d'aplicacions que té i com amb una mica de temps i esforç es poden aplicar a la indústria fent grans canvis. En aquest punt voldria donar les gràcies a l'empresa Reig Jofre, per haver-me permès portar la teoria a la pràctica. En especial voldria donar les gràcies al Dr. Saša Nikolić, gràcies per ensenyar-me tot el que sé sobre liofilització i per confiar en mi.

També voldria agrair a la Dra. Sònia Pérez responsable tècnic del Servei d'Anàlisi Tèrmic i Calorimetria del IQAC, i al director tècnic del Servei de Difracció de Raig X de la UAB Dr. Ángel Álvarez, per l'ajuda que m'han ofert a l'hora d'interpretar resultats d'equips més especialitzats.

Tot aquest projecte no hagués sigut possible sense el suport i ajuda extraordinària de amics i família, que tot i la "mala pata" que he tingut m'han ajudat en tot i més.

En aquets últims anys, Reig s'ha convertit en la meva segona casa i he tingut l'oportunitat de conèixer gent meravellosa que m'ha ajudat el que no està escrit, moltes gràcies a tots QA, IT, Producció i I+D. En especial voldria donar les gràcies al departament de QC per treure temps de la rutina per ajudar-me i aconsellar-me. Gràcies Montse A., Monica, Inma, Núria i en especial al Jesús, sense vosaltres no hagués arribat aquí. També, agrair a tots els analistes l'ajuda (Olga, Rebeca, Alejandro, Merche, ...), fent una menció especial a les dues persones que m'han ensenyat tot el que se sobre el KF i que sempre han estat disponibles per donar-me un cop de mà; gràcies Alba i Montse. També voldria agrair a la Blanca Ayza tot el que ha fet per mi (que no es poc), gràcies per escoltar-me cada dia i per ajudar-me a veure les coses des d'un altre punt de vista.

Tampoc voldria oblidar-me d'agrair a tota la gent del despatx que ha estat allà tant per oferir un mocador com per cantar alguna que altre cançó, moltes gràcies Laura, Pere, Emma, Carmen, Clara, Alicia, Clara E., Carlos, Eszter, Anna Flo, Saša, Anna Anglada i Glòria. També a l'Anna Lopez, a qui li desitjo tota la sort del món en el seu projecte i que sap que sempre podrà contar amb mi.

També voldria agrair a totes les persones del departament de química analítica i altres departaments que m'han ofert el seu suport en el àmbit més acadèmic, i sempre han estat allà per escoltar-me i treure'm un somriure. Gràcies Vanesa, Dong, Diego, Glòria, Aira, Judit Brassier, Judit Puig, Desirèe, Claudia, Elia, Anna, Marta, Kevin i Silvia. Sense oblidar-me de la persona que porta amb mi des del primer dia que vaig començar la carrera i amb la qui he compartit més llàgrimes i riures que ningú. Dioni, moltes gràcies per sempre estar allà, gràcies per escoltar-me i ajudar-me a seguir endavant, juntes ho vam començar i juntes ho acabarem.

Aquesta etapa m'ha permès conèixer molta gent però en especial dues persones. Anna Anglada ("pollet"), moltes gràcies per escoltar-me, per ser tan tossuda i per ajudar-me en tot el que podies i més. Gloria Clua, si no fos per tu no estaria ni escrivint aquestes línies, hem passat per molt les dues juntes i has passat de ser la meva mentora a la meva amiga, moltíssimes gràcies per tot el que has fet per mi.

Per últim voldria agrair a la meva família tot el suport que m'han donat durant aquest projecte. Gràcies, a l'Anna i la Silvia per estar sempre allà, per escoltar-me, per dir-me veritats quan feia falta i oferir-me un somriure quan era necessari. Gràcies, als meus cosins, al Jofre i al Nil,

i a les meves cunyades (Alba i Cristina) per sempre entendre que el ordinador formava part de mi, a les meves avies per estimar-me de forma incondicional tot i no veurem tant, i als meus tiets per sempre estar allà. En especial a la meva tieta Toni, qui sempre ha sigut per a mi un model a seguir. També donar les gràcies a la meva familie de les terres de Lleida que sempre han entès les hores amb l'ordinador. I per últim als meus pares, per què gràcies al seu esforç he arribat fins on he arribat i tinc la sort de ser qui sóc. Pama i mapa, gràcies per donar-me l'educació que tinc i per ensenyar-me que les coses s'han de treballar sempre amb un somriure a la cara. En especial voldria agrair l'ajuda del meu germà, en tot aquest projecte. Marc gràcies per ajudar-me en tot i treure temps dels teus projectes per mi. Cada cop que veig la teva feina i amb la persona en que t'has convertit em sento més orgullosa de tu, i estic segura que arribaràs molt lluny, *GG Max!*

Finalment, al Cesc, qui SEMPRE ha estat allà, nit i dia, qui ha entès i ha passat per totes i cada una de les etapes d'aquesta tesi. T'ho he dit molts cops però aquesta tesi no hagués sigut possible sense tu, i espero poder tornar-te tot el que tu has fet per mi aquests anys.

Moltes gràcies a tots,

Laia

RESUMEN

Esta tesis aborda la mejora de la calidad de productos farmacéuticos liofilizados. A grandes rasgos, el proceso de liofilización consiste en eliminar, por sublimación y desorción, el agua de la solución de medicamento congelada. El contenido de humedad residual (RMC) en un producto liofilizado se considera un atributo crítico de la calidad (CQA) y la cuantificación de este atributo aporta información sobre cómo se ha desarrollado el ciclo de liofilización y sobre la evolución de la calidad del producto durante su vida útil. Por otra parte, las formulaciones de productos farmacéuticos liofilizados que contienen biomoléculas o moléculas orgánicas de gran tamaño acostumbran a presentar una alta sensibilidad al oxígeno; un alto contenido de oxígeno puede tener efectos negativos en la calidad del producto final: disminución de la eficacia, formación de subproductos que pueden tener alta toxicidad o el cambio de coloración del producto.

En esta tesis se han desarrollado métodos analíticos para determinar y evaluar el impacto en el producto liofilizado de la humedad y el oxígeno. Todos los productos liofilizados utilizados han sido fabricados por Laboratorios Reig Jofre.

En un primer trabajo se ha estudiado el impacto de la humedad contenida en los tapones sobre un producto liofilizado hidrolizable, el cual ha sido almacenado durante 6 meses en condiciones de estabilidad acelerada. Para ello, previamente ha sido necesario desarrollar y validar, con la rigurosidad exigida en el entorno farmacéutico, un método analítico basado en la valoración coulombimétrica de Karl Fischer, acoplada a un horno, para la determinación de la humedad residual en los tapones de los viales. El estudio de la humedad y de los productos de degradación por hidrólisis a lo largo de los meses, ha permitido deducir que la degradación por hidrólisis del producto está relacionada con la cantidad de humedad inicial de los tapones.

En el siguiente trabajo se ha desarrollado un método de cuantificación de RMC en otro producto liofilizado. A partir de medidas espectroscópicas en el infrarrojo cercano (NIR), se ha realizado un modelo de regresión por mínimos cuadrados parciales (PLSR). Previamente, ha sido necesario la puesta a punto y validación del método analítico de referencia. El modelo multivariante se ha validado siguiendo las directrices de las diferentes agencias regulatorias de la industria farmacéutica. Finalmente, se ha aplicado a escala industrial para realizar un mapeo de humedad, evaluando la homogeneidad del proceso de secado del ciclo de liofilización.

En el tercer trabajo se ha puesto a punto un método cualitativo para relacionar el contenido de oxígeno en el *headspace* del vial con el espectro de reflectancia en el visible del producto. Se ha realizado la clasificación con PLS-DA de muestras con alto y bajo contenido de oxígeno a partir de un cierto valor umbral. El método se ha validado y se han calculado los parámetros de rendimiento de éste; la zona de falta de fiabilidad del método se ha calculado a partir la curva característica de rendimiento.

En el último trabajo se ha desarrollado y validado un método para la identificación, en el propio envase, de productos terminados antes de su liberación. Se ha utilizado la espectroscopia NIR para crear una biblioteca representativa de los espectros NIR de diez productos y se han utilizado técnicas quimiométricas de reconocimiento de patrones para su verificación y validación.

LISTADO DE ABREVIACIONES

Acrónimo	Significado en inglés	Significado en español
ANN	Artificial Neural Networks	Redes neuronales artificiales
ANOVA	Analysis of Variance	Análisis de la varianza
AOTF	Acousto-Optic Tunable Filters	Filtros acústico-óptico
API	Active Pharmaceutical Ingredient	Ingrediente activo farmacéutico
CC	Correlation Coefficient	Coefficiente de correlación
CCI	Container Closure Integrity	Integridad del cierre del vial
CIE	Commission Internationale de l'Eclairage	Comisión internacional de iluminación
CIP	Clean In Place	Limpieza in situ
CQA	Critical Quality Attribute	Atributos Críticos de calidad
Edw 1	Lyofast 22	--
Edw 2	Lyofast 25	--
Edw 3	Lyofast 15	--
EMA	European Medicines Agency	Agencia europea del medicamento
FDA	Food and Drug Administration	---
FIR	Far Infrared	Infrarrojo lejano
FMEA	Failure Mode and Effects Analysis	Análisis modal de fallos y efectos
FMS	Frequency Modulation Spectroscopy	Espectroscopia de modulación de la frecuencia
FN	False Negative	Falsos negativos
FNR	False Negative Rate	Tasa de falsos negativos
FP	False Positive	Falsos positivos
FPR	False Positive Rate	Tasa de falsos positivos
GMP	Good Manufacturing Practice	Normas de correcta fabricación
ICH	The International Council for Harmonisation	Consejo Internacional para la armonización
KNN	K-Nearest Neighbour	K vecino más cercano
LED	Light Emitting Diode	Diodo emisor de luz
LOD	Limit of Detection	Límite de detección
LOQ	Limit of Quantification	Límite de cuantificación
MIR	Mid-Infrared	Infrarrojo medio
MLR	Multiple Linear Regression	Regresión lineal múltiple

NIPALS	Nonlinear Iterative Partial Least Squares	Mínimos cuadrados parciales iterativos no lineales
NIR	Near Infrared	Infrarrojo cercano
OSC	Orthogonal Signal Correction	Corrección de la señal ortogonal
PC	Principal Component	Componente principal
PCA	Principal Component Analysis	Análisis de componentes principales
PCR	Principal Component Regression	Regresión en las componentes principales
Ph. Eur.	European Pharmacopoeia	Farmacopea europea
PLS-DA	Partial Least Squares Discriminal Analysis	Análisis discriminante por mínimos cuadrados
PLSR	Partial Least Squares Regression	Regresión parcial por mínimos cuadrados
PRM	Pattern Recognition Methods	Métodos de reconocimiento de pautas
PRT	Pressure Rise Test	Test de incremento de presión
QbD	Quality by Design	Calidad por diseño
R ²	Coefficient of Determination	Coefficiente de determinación
RMC	Residual Moisture Content	Contenido de humedad residual
RMSEC	Root Mean Square Error of Calibration	Error cuadrático medio de calibración
RMSECV	Root Mean Square Error of Cross Validation	Error cuadrático medio de validación cruzada
RMSEP	Root Mean Square Error of Prediction	Error cuadrático medio de predicción
RPN	Risk Priority Number	Número de prioridad de riesgo
RSD	Relative Standard Deviation	Desviación estándar relativa
S	Selectivity	Selectividad
SEL	Standard Error of Laboratory	Error estándar de laboratorio
SECV	Standard Error of Cross Validation	Error estándar de la validación cruzada
SIMCA	Soft Independent Modeling of Class Analogy	Modelado suave independiente por analogía de clases
SIP	Sterilization in Place	Esterilización in situ
SNV	Standard Normal Variate	Variable normal estándar
T _{co}	Collapse Temperature	Temperatura de colapso
TDLAS	Tunable Diode Laser Absorption Spectroscopy	Espectroscopia de absorción láser de diodo sintonizable
TDLAS-FMS	Tunable Diode Laser Absorption Spectroscopy with Frequency Modulation Spectroscopy	Espectroscopia de absorción láser de diodo sintonizable con espectroscopia de modulación de frecuencia
Teu	Eutectic Temperature	Temperatura eutéctica
Tg'	Glass Transition Temperature	Temperatura de transición vítrea

TN	True Negative	Negativos verdaderos
TP	True Positive	Positivos verdaderos
U	Threshold	Umbral
USP	United States Pharmacopeia	Farmacopea de Estados Unidos
UV	Ultraviolet	Ultravioleta
WNS	Wavelength Modulation Spectroscopy	Espectroscopia de modulación de longitud de onda

ÍNDICE

I. INTRODUCCIÓN

1. PRODUCTOS FARMACÉUTICOS LIOFILIZADOS.....	21
1.1. Laboratorios Reig Jofre	22
1.2. Liofilización.....	23
1.2.1. Fundamento teórico	24
1.2.2. Instrumentación.....	30
1.3. Atributos críticos de calidad de un medicamento parenteral liofilizado	31
2. KARL FISCHER	32
2.1. Introducción	32
2.2. Fundamento teórico	33
2.2.1. Valoración volumétrica	35
2.2.2. Valoración coulombimétrica	36
2.3. Instrumentación	37
2.3.1. Karl Fischer acoplado a un horno	37
3. ESPECTROSCOPIA DE ABSORCIÓN EN EL INFRARROJO CERCANO.....	40
3.1. Introducción	40
3.2. Fundamento teórico	42
3.3. Instrumentación	45
3.3.1. Fuente de radiación.....	46
3.3.2. Selector de longitudes de onda	46
3.3.3. Presentación de la muestra.....	47
3.3.4. Detector	49
3.4. Aplicaciones en la industria farmacéutica.....	50
3.4.1. Cuantificación de la humedad residual en un producto farmacéutico liofilizado	51
3.4.2. Identificación de productos farmacéuticos	52
4. ESPECTROSCOPIA DE ABSORCIÓN LÁSER DE DIODO SINTONIZABLE.....	54
4.1. Introducción	55
4.2. Fundamento teórico	55
4.2.1. Espectroscopia de modulación de frecuencia.....	57
4.3. Instrumentación	58

4.4.	Aplicaciones.....	59
4.4.1.	Base teórica de la determinación del oxígeno.....	60
5.	QUIMIOMETRÍA	63
5.1.	Introducción	63
5.2.	Análisis cualitativo.....	64
5.2.1.	Similitud	65
5.2.2.	Análisis de componentes principales.....	67
5.2.3.	Análisis discriminante por mínimos cuadrados parciales.....	69
5.3.	Análisis cuantitativo	70
5.3.1.	Calibración univariante	70
5.3.2.	Calibración multivariante	71
5.4.	Etapas del proceso de modelado.....	73
5.4.1.	Definición del objetivo	74
5.4.2.	Preparación de las muestras	74
5.4.3.	Obtención de la información analítica	75
5.4.4.	Pretratamiento de los datos	76
5.4.5.	Selección del intervalo de longitudes de onda	78
5.4.6.	Exploración de los datos	78
5.4.7.	Selección de muestras.....	79
5.4.8.	Construcción y evaluación del modelo quimiométrico.....	81
5.4.9.	Validación del modelo quimiométrico.....	84
5.4.10.	Aplicación y mantenimiento del modelo.....	84
6.	VALIDACIÓN DE MÉTODOS ANALÍTICOS	86
6.1.	Introducción	86
6.2.	Propiedades analíticas estudiadas en la validación de métodos analíticos.....	87
6.2.1.	Especificidad	88
6.2.2.	Linealidad	89
6.2.3.	Intervalo de linealidad	89
6.2.4.	Exactitud	90
6.2.5.	Precisión.....	90
6.2.6.	Límite de detección	91
6.2.7.	Límite de cuantificación	91
6.2.8.	Robustez	92

6.3. Validación de métodos analíticos multivariantes	92
6.3.1. Adaptación de los parámetros de la validación de métodos analíticos multivariantes	93
7. BIBLIOGRAFÍA	95

II. OBJETIVOS

1. OBJETIVOS	109
--------------------	-----

III. REACTIVOS Y MATERIALES

1. REACTIVOS.....	113
2. PATRONES.....	113
3. INSTRUMENTOS.....	113
3.1. Espectrofotómetro NIR.....	114
3.2. Liofilizadores	114
3.3. Valorador Karl Fischer	115
3.4. Espectrofotómetro de absorción láser de diodo sintonizable para la determinación de oxígeno.....	117
3.5. Espectrofotómetro UV-Visible	118
3.6. Otros instrumentos	119
4. PROGRAMAS PARA LA ADQUISICIÓN Y EL TRATAMIENTO DE LOS DATOS	120
4.1. Programas para la adquisición de los datos	120
4.2. Programas para el tratamiento de los datos.....	120
5. BIBLIOGRAFÍA	121

IV. IMPACTO DE LA HUMEDAD DEL TAPÓN DEL VIAL EN EL PRODUCTO TERMINADO

1. INTRODUCCIÓN	125
2. MATERIALES Y MÉTODOS EXPERIMENTALES.....	126
2.1. Muestras	126
2.1.1. Tapones	126
2.1.2. Muestras del producto H.....	126
2.2. Métodos experimentales.....	127
2.2.1. Determinación de la humedad en tapones de clorobutilo de 13 mm	127
2.2.2. Determinación de la humedad residual del producto H	128

2.2.3.	Determinación de la concentración y las impurezas del producto H	128
3.	RESULTADOS Y DISCUSIÓN.....	128
3.1.	Desarrollo del método	128
3.1.1.	Selección de la técnica.....	128
3.1.2.	Preparativa de la muestra	129
3.1.3.	Selección de los parámetros del método	129
3.2.	Validación del método	131
3.2.1.	Especificidad	131
3.2.2.	Exactitud	132
3.2.3.	Linealidad	132
3.2.4.	Intervalo de linealidad	133
3.2.5.	Precisión.....	134
3.2.6.	Límite de cuantificación	135
3.3.	Impacto de la humedad inicial de los tapones en un producto	135
3.3.1.	Planteamiento del experimento	136
3.3.2.	Humedad de los tapones	137
3.3.3.	Degradación del producto.....	138
4.	CONCLUSIONES.....	141
5.	BIBLIOGRAFÍA	141
V.	DETERMINACIÓN DE LA HUMEDAD RESIDUAL EN UN PRODUCTO LIOFILIZADO	
1.	INTRODUCCIÓN.....	145
2.	MATERIALES Y MÉTODOS EXPERIMENTALES.....	146
2.1.	Muestras	146
2.2.	Método de referencia	147
2.3.	Adquisición de los espectros.....	149
3.	RESULTADOS Y DISCUSIÓN	149
3.1.	Análisis de viabilidad del método	149
3.2.	Objetivo del modelo.....	169
3.3.	Desarrollo del modelo cuantitativo NIR	169
3.4.	Validación del modelo cuantitativo	174
3.4.1.	Especificidad.....	175

3.4.2.	Linealidad.....	176
3.4.3.	Intervalo de linealidad.....	176
3.4.4.	Exactitud.....	176
3.4.5.	Precisión.....	177
3.4.6.	Límite de detección y cuantificación.....	178
3.4.7.	Robustez.....	179
3.4.8.	Conclusiones de la validación del modelo.....	179
3.5.	Aplicación del modelo.....	179
4.	CONCLUSIONES.....	183
5.	BIBLIOGRAFÍA.....	184

VI. DETERMINACIÓN DE OXÍGENO EN EL *HEADSPACE* DE UN PRODUCTO LIOFILIZADO

1.	INTRODUCCIÓN.....	189
2.	MATERIALES Y MÉTODOS EXPERIMENTALES.....	190
2.1.	Muestras.....	190
2.2.	Determinación de la absorbancia del producto disuelto.....	190
2.3.	Método de referencia.....	191
2.4.	Adquisición de los espectros.....	191
3.	RESULTADOS Y DISCUSIÓN.....	191
3.1.	Estudios previos.....	191
3.2.	Desarrollo de un modelo cuantitativo.....	193
3.3.	Desarrollo de un modelo cualitativo.....	196
3.4.	Validación del modelo.....	204
4.	CONCLUSIONES.....	206
5.	BIBLIOGRAFÍA.....	207

VII. CONSTRUCCIÓN Y VALIDACIÓN DE UNA BIBLIOTECA DE ESPECTROS NIR

1.	INTRODUCCIÓN.....	211
2.	MATERIALES Y MÉTODOS EXPERIMENTALES.....	212
2.1.	Muestras.....	212
2.2.	Adquisición de los espectros.....	213

3.	RESULTADOS Y DISCUSIÓN	215
3.1.	Descripción de las muestras	215
3.2.	Construcción de la biblioteca de espectros	217
3.3.	Validación de la biblioteca	224
4.	CONCLUSIONES.....	226
5.	BIBLIOGRAFÍA	227

VIII. CONCLUSIONES

1.	CONCLUSIONES.....	231
----	-------------------	-----

Esta tesis se ha realizado dentro del programa de doctorado del Departamento de Química de la Universidad Autónoma de Barcelona. Su ejecución ha sido posible gracias al programa de ayudas de la Generalitat de Catalunya de Doctorats Industrials, que se caracteriza por fomentar la transferencia de conocimiento entre el ámbito académico y el industrial, uniendo ambos entornos.

Teniendo en cuenta esta particularidad y para facilitar la lectura y comprensión del trabajo, se ha escrito un capítulo de introducción. Éste se divide en distintos apartados, ofreciendo una descripción detallada tanto del proceso de liofilización, utilizado por Laboratorios Reig Jofre en la fabricación de sus productos, como de las técnicas analíticas empleadas en la tesis (valoración Karl Fischer, espectroscopia NIR y espectroscopia TDLAS-FMS). Además, se ha realizado una introducción a las técnicas quimiométricas utilizadas y un resumen exhaustivo de la forma de trabajar en un entorno industrial farmacéutico, con énfasis en la validación de los métodos analíticos. Al final de la introducción se incluye una extensa bibliografía sobre los temas tratados.

En el siguiente capítulo se describen los objetivos principales de la tesis y, a modo de subobjetivos, los pasos necesarios para alcanzarlos.

A continuación, y con la finalidad de agilizar la lectura de los capítulos en los que se describe en detalle el trabajo realizado, se ha compilado la descripción de los equipos y programas empleados. Gran parte de esta instrumentación se ha utilizado en todos los estudios realizados.

Los últimos cuatro capítulos corresponden a los cuatro trabajos desarrollados en esta tesis. Cada uno de ellos consta de una pequeña introducción, para poner al lector en contexto, exponiendo la necesidad y el interés del desarrollo del estudio. Posteriormente, se detallan los métodos analíticos empleados, característicos de cada estudio, junto con la metodología aplicada, los resultados obtenidos y su discusión. Cada capítulo se cierra con unas breves conclusiones y la bibliografía de interés.

El último capítulo consiste en las conclusiones conjuntas de todo el trabajo realizado, tanto desde un punto de vista académico como industrial.

I. INTRODUCCIÓN

1. PRODUCTOS FARMACÉUTICOS LIOFILIZADOS

Los productos farmacéuticos parenterales son aquellos medicamentos que se administran al paciente por vía intramuscular, subcutánea o intravenosa, entre otras. Esta forma de administrar el medicamento presenta numerosas ventajas respecto a la forma oral (mayor biodisponibilidad, nula degradación gastrointestinal y hepática, ...).¹ No obstante se requiere que el producto sea estéril, lo que dificulta su fabricación y aumenta la complejidad del control de calidad. La presente tesis doctoral se ha centrado en los medicamentos liofilizados para la administración parenteral.

Se puede definir un medicamento como una mezcla de sustancias que, administradas a un organismo, permiten el tratamiento o la prevención de enfermedades, o que ayudan en la restauración, corrección o modificación de algunas funciones fisiológicas de éste.^{2,3} Los medicamentos pueden estar compuestos por uno o varios principios activos (API, siglas del inglés *Active Pharmaceutical Ingredient*) los cuales desarrollan la actividad farmacológica y aportan eficacia al medicamento, y por excipientes, cuya función consiste en aportar distintas propiedades al mismo (forma, biodisponibilidad y estabilidad del API, sabor del producto, ...) con el fin de facilitar su administración.

La industria farmacéutica se encarga de la producción de los medicamentos e invierte parte de su capital para impulsar la investigación de nuevos productos y el desarrollo de procesos de fabricación más eficientes.⁴⁻⁶ Aunque los primeros indicios del uso de medicamentos datan del 3000-4000 a.C., la historia de la industria farmacéutica, tal y como la conocemos hoy en día, se remonta al siglo XIX. Desde entonces hasta la actualidad, este sector ha ido evolucionando a pasos agigantados, y se parece poco a los primeros laboratorios de boticarios, los cuales marcaron el inicio de esta industria.⁷

Durante este tiempo, la regulación aplicada a la industria farmacéutica ha ido evolucionando con el fin de comercializar medicamentos fiables y seguros, hasta llegar a los estándares de alta calidad con que se trabaja hoy en día.⁸

1.1. Laboratorios Reig Jofre

Laboratorios Reig Jofre es una compañía farmacéutica familiar fundada en 1929 por Ramón Reig Jofre en Barcelona. Lo que antaño empezó siendo un pequeño laboratorio en la trastienda de una farmacia; actualmente, cuenta con 4 centros de desarrollo y producción en Europa (uno en Suecia, dos en Toledo y otro en Barcelona) y comercializa productos en más de 70 países del mundo (tanto productos propios como de socios comerciales).⁹



Figura I - 1. A la izquierda se puede ver el logo de la compañía en conmemoración del 90^º aniversario de la fundación de la empresa. A la derecha, se puede ver la farmacia original donde se fundó el primer Reig Jofre.

Reig Jofre, se dedica a la investigación, el desarrollo, la fabricación y comercialización de productos farmacéuticos y complementos nutricionales. Dadas estas características, no es de extrañar que siempre haya mostrado un interés para colaborar con el ámbito académico y así aplicar la investigación para desarrollar y mejorar los procesos de fabricación. Por esto, desde el año 2014 Laboratorios Reig Jofre ha participado activamente en el programa de Doctorados Industriales de la Generalitat de Catalunya. El principal objetivo de esta iniciativa es contribuir a la competitividad y a la internacionalización de la industria catalana, facilitando la transferencia de conocimientos de las universidades a las empresas.

Esta tesis ha sido realizada con una ayuda de este programa (expediente DI 2016-034) y se ha desarrollado en la planta de Barcelona de Laboratorios Reig Jofre, especializada en productos farmacéuticos inyectables líquidos y liofilizados. Este proyecto se ha realizado en colaboración con la Unidad de Química Analítica, del Departamento de Química de la Universidad Autónoma de Barcelona, con el fin de aplicar los nuevos avances de la química analítica a este sector.

1.2. **Liofilización**

Para administrar un medicamento parenteral y que éste pueda realizar su función, debe encontrarse disuelto.¹⁰ Este punto es crítico debido a que no todos los API son estables en solución, lo que produce una rápida degradación del producto. Con el fin de poder prolongar la vida útil de estos API, en el sector farmacéutico se ha empleado, tradicionalmente, el proceso de liofilización.¹¹⁻¹³

A grandes rasgos, el proceso de liofilización consiste en eliminar, por sublimación y desorción, el agua de la solución del medicamento congelada. Al ser un proceso aprobado por las autoridades reguladoras farmacéuticas y, que puede ser realizado en condiciones estériles, cuenta con una amplia gama de aplicaciones.¹² Estas aplicaciones no se limitan sólo a medicamentos parenterales tradicionales, sino que también incluyen formulaciones más novedosas basadas en nanopartículas o estructuras tridimensionales, entre otros.¹⁴

Si hacemos un repaso de los orígenes del proceso de liofilización, tenemos que remontarnos a la época en la que pueblos como los inuit o los samis congelaban los alimentos y empleaban la sublimación del agua para mejorar su preservación. Este proceso, aunque tiene la misma finalidad que el proceso actual de liofilización (eliminar el agua), se diferencia en que se realiza a presión atmosférica.¹⁵ La primera civilización que empleó no sólo una temperatura reducida, sino también una baja presión para sublimar el agua de los alimentos, fue la incaica situada en las altas montañas de los Andes. Los incas gozaban, de forma natural, de unas bajas temperaturas y de una presión más reducida (dada la altitud a la que se encontraban respecto al nivel del mar), características que permiten considerarlos como los precursores del proceso de liofilización.^{16,17}

El interés científico por este proceso empezó a finales del siglo XIX principios del siglo XX, cuando se publicaron diferentes estudios sobre la preparación de tejidos biológicos y se patentaron los primeros liofilizadores.^{18,19} Sin embargo, aunque estos estudios aportaron avances significativos, no llegaron a traspasar el ámbito académico.²⁰ No fue hasta la Segunda Guerra Mundial (1939-1945) cuando la creciente necesidad de plasma sanguíneo y penicilina puso a la liofilización en el punto de mira. Desde entonces la industrialización de la liofilización ha ido creciendo para adaptarse a la nueva tecnología y regulación.¹⁶

Actualmente, el proceso de fabricación industrial de medicamentos parenterales liofilizados suele ser por lotes, en los que la formulación líquida del medicamento se dosifica en viales y éstos se introducen en los liofilizadores de modo que, el proceso tiene lugar de forma independiente en cada vial.²¹

1.2.1. Fundamento teórico

El proceso de liofilización se basa en la transferencia simultánea de calor y de masa de la disolución, para conseguir eliminar el solvente, habitualmente agua.²² El principio fundamental de la liofilización es la sublimación, que consiste en el cambio de estado del solvente de una fase sólida a una gaseosa, sin tener que pasar por el estado líquido.²⁰

Este proceso se acostumbra a dividir en las siguientes operaciones unitarias:²³

- Congelación: la principal función de esta etapa es solidificar la formulación líquida del medicamento para, posteriormente, empezar con el secado.
- Secado primario: en esta etapa se disminuye la presión y se aumenta la temperatura del liofilizador, favoreciendo la eliminación por sublimación del agua que se encuentra en forma de hielo en el producto.
- Secado secundario: en esta etapa, se aplica al producto una menor presión y una mayor temperatura que en la etapa anterior, con el fin de eliminar el agua que se encuentra adsorbida.

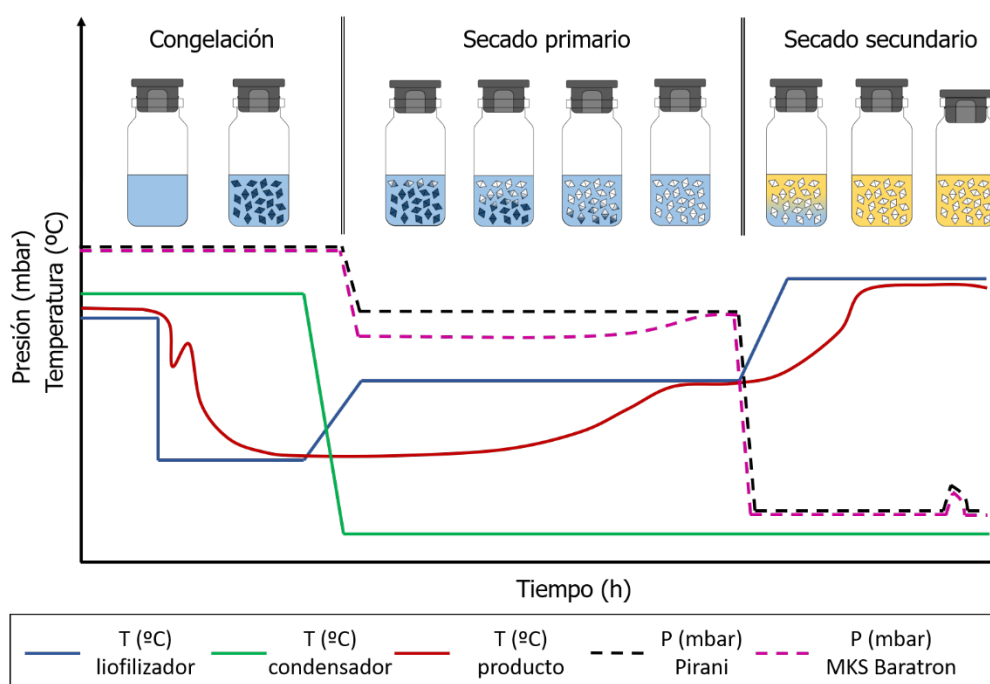


Figura I - 2. Esquema de las distintas etapas de un ciclo de liofilización, indicando la evolución de la temperatura y presión, así como su efecto en un vial dosificado.

La selección de los valores de temperatura o de la presión en estas etapas depende de las características térmicas de la formulación líquida del producto farmacéutico. Estos parámetros tienen un impacto crítico, tanto económico (duración del ciclo de liofilización, gasto energético, entre otros) como en términos de calidad del producto final (estabilidad del producto, posible degradación del API, ...).^{24,25}

Por esta razón, es importante optimizar el ciclo de liofilización para cada formulación de un medicamento. Normalmente esta optimización se ha estudiado aplicando el método de ensayo y error; no obstante, en la última década la utilización de la calidad por diseño (QbD, de las siglas en inglés de *Quality by Design*) ha ido ganando terreno.^{14,23,26}

En las siguientes secciones se describen con más detalle cada una de las etapas principales del proceso de liofilización.

1.2.1.1. Congelación

La congelación se considera la primera etapa del proceso de liofilización, y está relacionada con la eficacia de las siguientes etapas del mismo, así como con las propiedades fisicoquímicas del producto final.^{18,23,27}

La implementación de esta etapa se debe al profesor de farmacología L.F.Shackell, cuando en 1909 se percató de que antes de empezar a secar por vacío era necesario que la muestra estuviera congelada.¹⁹ Desde entonces, el estudio de esta fase, inicialmente considerada trivial, ha ido ganando importancia.

El primer paso de la etapa de congelación es la formación de cristales de hielo. Estos cristales se forman a partir de la nucleación; un fenómeno espontáneo, exotérmico y estocástico, que se ve favorecido a bajas temperaturas.^{21,28} Una vez formados los cristales de hielo, éstos empiezan a crecer provocando un aumento de la concentración en la solución del medicamento que recibe el nombre de crioconcentración.²⁸ La formación de cristales únicamente de hielo puro es debido a que las moléculas del medicamento (API y/o excipiente) tienen un mayor tamaño que las del disolvente y no pueden entrar en la estructura del hielo. La formación de los cristales de hielo de gran tamaño en la crioconcentración facilita la sublimación en el secado, lo que conlleva una reducción del tiempo de esta etapa, aunque puede tener un impacto notable en la degradación del producto.¹⁴

La temperatura para conseguir que el producto se encuentre en estado sólido depende de su estructura:

- Productos cristalinos: en el caso de los medicamentos que presentan una estructura cristalina la temperatura de esta etapa debe ser inferior a la temperatura eutéctica (T_{eu}). Por debajo de esta temperatura toda la solución que todavía se podía encontrar en equilibrio con el hielo, cristaliza.
- Productos amorfos: en este caso los productos no llegan a cristalizar, sino que se quedan en un estado de sólido amorfo. La temperatura en la etapa de congelación para este tipo de productos debe ser inferior a la temperatura de transición vítrea (T_g')^{1,23}

En algunos productos amorfos, para promover el proceso de cristalización se aplica el tratamiento térmico llamado *Annealing*. Este tratamiento consiste en someter a la solución congelada un aumento de la temperatura por encima de la T_g' , con el fin de aumentar la movilidad de las moléculas del producto y promover su redistribución, permitiendo la cristalización del producto inicialmente amorfo.²⁹ Una vez terminado este calentamiento, se vuelve a disminuir la temperatura (por debajo de T_g') para reducir la movilidad y tener las condiciones óptimas para la siguiente etapa.^{12,30}

Dado que la formulación líquida del medicamento se encuentra dosificada en distintos viales, la nucleación del hielo no se produce de forma homogénea en todo el lote, por lo que se debe dejar el tiempo suficiente para que se produzca la solidificación en todos los viales del mismo.^{23,28}

1.2.1.2. Secado primario

La siguiente etapa del proceso de liofilización es el secado primario. Se centra en sublimar la mayor cantidad posible de solvente que queda retenido en los cristales de hielo formados durante la etapa de congelación. La finalidad de esta etapa es obtener un producto seco con una estructura intacta.^{18,31} Es la etapa del proceso de liofilización que requiere más tiempo (puede llegar a durar varios días), por lo que todos los procesos de optimización de ciclo de liofilización se centran en ella.^{20,27}

Para que se produzca la sublimación, la presión de la cámara del liofilizador debe ser menor que la presión de vapor del hielo en el frente de sublimación del producto. Esta diferencia de presión hace que las moléculas de agua migren del producto a la cámara del liofilizador, y de ésta al condensador, donde quedan retenidas.^{11,12}

¹ La denotación (') se emplea para determinar la transición vítrea de soluciones congeladas concentradas.²⁹

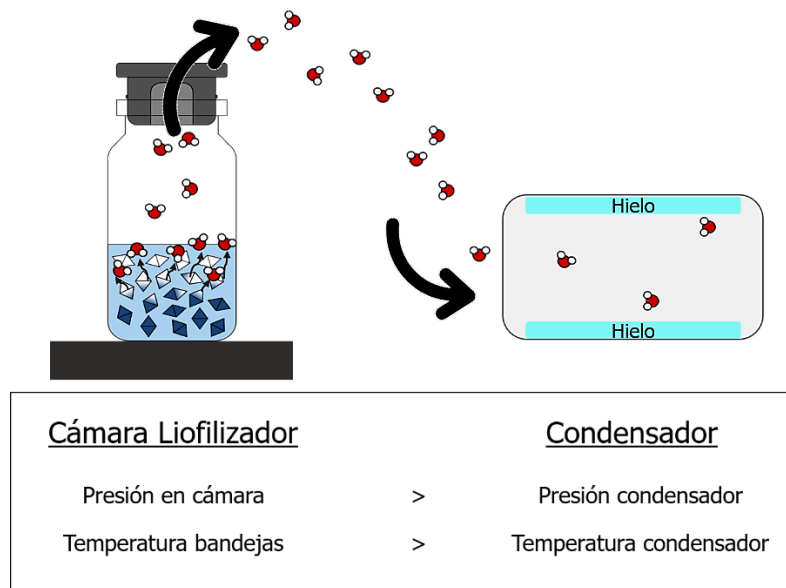


Figura I - 3. Esquema de cómo se desprenden las moléculas de agua retenidas en el producto congelado y se dirigen al condensador, donde quedan atrapadas al congelarse de nuevo.

Para que las moléculas de agua de los cristales de hielo puedan pasar a estado gaseoso necesitan aumentar su movilidad, se debe incrementar la temperatura de la cámara del liofilizador.³² Así pues, es necesario controlar tanto la temperatura como la presión de la cámara de liofilización; la combinación óptima de ambas será la que permita tener la máxima sublimación con el mínimo tiempo y sin que el producto funda (productos cristalinos) o colapse (productos amorfos).^{11,18,23}

El colapso de un medicamento liofilizado es un fenómeno que consiste en una pérdida de la estructura de la pastilla liofilizada (Figura I - 4) y suele estar relacionado con someter el producto a una temperatura superior a su temperatura de colapso en productos amorfos, T_{co} .^{33,34} Este proceso puede ocurrir durante las etapas del secado primario y secundario, o durante el almacenamiento del medicamento liofilizado.³⁵ El colapso del producto depende de la formulación y, en algunos casos, puede tener un impacto significativo en términos de calidad (como por ejemplo, en el aspecto, en un aumento en el tiempo de reconstitución, de la humedad o de la degradación).³⁶ Por lo tanto, debe evitarse el colapso, para garantizar la calidad del producto liofilizado.³⁷



Figura I - 4. Imagen de un producto liofilizado colapsado (los dos viales de la izquierda), frente a un vial del mismo producto sin colapso (vial de la derecha). Imagen extraída de la referencia.³⁸

La selección de presión y temperatura es única para cada producto y depende de las características térmicas de la solución del medicamento. Para diseñar el secado primario de una forma eficiente y reducir la duración de esta etapa, la velocidad de sublimación en el frente de sublimación debe ser lo más alta posible. Para ello, la temperatura del producto debe ser lo más próxima posible a la temperatura de colapso, sin llegar a sobrepasarla.^{23,39,40}

La etapa del secado primario termina cuando toda el agua libre ha sublimado. Finalizarla prematuramente tiene un gran impacto en la calidad del producto (por ejemplo, una alta humedad residual o colapso), mientras que una prolongación excesiva de ésta implica un coste económico innecesario. Por esta razón, se necesita una estrategia adecuada para determinar el punto final del secado primario.^{23,41}

Existen diferentes métodos para determinar el punto final de esta etapa:

- Temperatura del producto: cuando la sonda de temperatura (por ejemplo, termopares) en contacto con el producto, indica una temperatura similar a la del fluido térmico que circula por los pisos del liofilizador, denominada como temperatura de bandejas, significa que no hay transferencia calorífica y que el vial no necesita más energía para realizar la sublimación y, por lo tanto, que el secado primario ha terminado. El principal problema de este método se debe a que sólo algunos viales del lote contienen sonda, por lo que sus lecturas son únicamente una estimación para el conjunto del lote. Además, la propia presencia de la sonda genera un núcleo de congelación en el vial que la contiene, que no está presente en el resto, de modo que lo diferencia de los demás.^{23,41}
- Comparación de presiones: permite solventar el problema del método anterior debido a que no es invasivo. En este caso, la determinación del punto final del secado primario se realiza comparando el valor obtenido por la sonda de presión Pirani

(basada en la conductividad térmica de los gases) con el valor obtenido por la sonda de presión MKS Baratron (presión capacitiva). Cuando el proceso de sublimación ha terminado, la concentración de vapor de agua en la cámara del liofilizador decrece y consecuentemente el valor de la sonda de presión MKS Baratron aumenta. Entonces, cuando la sonda Pirani y la MKS Baratron convergen se considera que la etapa ha finalizado.^{18,23,41}

- Otros: aunque los dos métodos comentados son los más empleados, existen otros basados en diversas propiedades físicas del proceso. Los más importantes están basados en el uso de técnicas espectroscópicas para la determinación del caudal de vapor de agua desde la cámara de liofilización hasta el condensador.^{18,41}

1.2.1.3. Secado secundario

Una vez terminada la etapa del secado primario, el producto muestra un aspecto seco, pero aún presenta un contenido de humedad residual alto que debe reducirse para prolongar la estabilidad del producto.¹¹ Con el fin de disminuir el contenido de agua se realiza la última etapa del proceso de liofilización, el secado secundario. En esta etapa se elimina el agua que se encuentra adsorbida en el producto, la que no se sublimó en la etapa anterior, debido a que no se solidificó en forma de hielo puro sino mezclada con el medicamento. La separación de ambas fases de secado se empezó a emplear en 1952, en un artículo de R. I. N. Greaves.^{18,23,32}

El contenido de humedad residual al finalizar el secado secundario acostumbra a encontrarse alrededor de un 1 %.^{40,42} Sin embargo, el valor óptimo para cada producto depende de la estabilidad del producto final y los valores de los parámetros de temperatura, presión y tiempo de esta etapa dependen del contenido de humedad residual deseado para el producto liofilizado.^{32,39,43}

Aunque el secado secundario, operacionalmente, empieza una vez ha terminado el secado primario, la desorción del agua ya se inicia durante la etapa anterior, cuando en alguna zona del producto ya no queda hielo, por lo que, aunque el proceso de desorción es más lento que el de sublimación, el secado secundario acostumbra a necesitar menos horas que el secado primario para realizar su función.^{11,27,42,44}

Esta reducción del tiempo de secado también se debe a que al tener un contenido de agua inicial menor que en el secado primario, las condiciones de temperatura y presión que se aplican pueden ser más agresivas (considerando el perfil térmico del producto). Estas nuevas condiciones se basan, habitualmente, en aumentar la temperatura y

disminuir la presión hasta el mínimo posible que pueda alcanzar el equipo.⁴⁴ De esta manera, se suministra al sistema más energía para poder liberar las moléculas de agua que seguían adsorbidas en el producto. Este incremento de energía se realiza aumentando la temperatura, lo que a su vez puede llevar a la degradación del producto. Por esta razón la definición del tiempo de duración de esta etapa es crucial para mantener la calidad del producto.^{12,23}

Aunque existen varios métodos para saber cuándo ha terminado el secado secundario, el más empleado, es la prueba de incremento de presión (PRT, del inglés *Pressure Rise Test*). La extensión de este método se debe a la facilidad con la que se puede realizar, a lo económico que resulta y a su fiabilidad.⁴³ Consiste en cerrar la válvula que conecta la bomba de vacío con la cámara del liofilizador durante unos 15 min. Al cerrarse, la presión de la cámara aumenta debido al incremento de la concentración de vapor de agua. Si el producto todavía no está suficientemente seco, este aumento de presión será considerable, mientras que si el producto ya está casi seco el aumento de presión será mínimo.⁴⁵

1.2.2. Instrumentación

Al ser un proceso tan complejo, la liofilización requiere de equipos igual de complejos, denominados liofilizadores.¹⁶ Desde su primera aplicación a escala industrial, las partes principales del liofilizador siguen siendo las mismas, pero como es de esperar, todo el sistema de control de parámetros como la temperatura o la presión han cambiado radicalmente.¹²

Fundamentalmente, estos equipos están dotados de una gran cámara hermética con control de la temperatura donde se introducen los medicamentos ya dosificados (en viales) en los distintos pisos, un condensador para retener el agua que ha sublimado del producto y un sistema de vacío para controlar la presión de la cámara.^{12,18,22} En la siguiente figura (Figura I - 5) se muestra un esquema de las partes principales de un liofilizador:

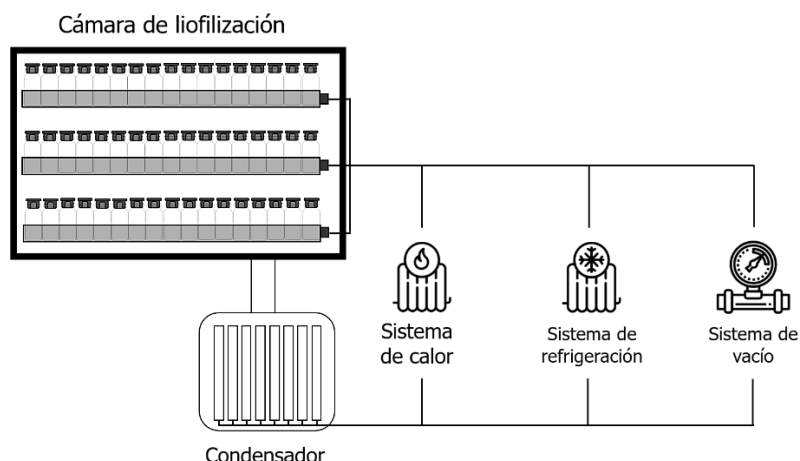


Figura I - 5. Esquema de la estructura de un equipo de liofilización, mostrando los principales bloques que lo componen.

En el caso particular del sector industrial farmacéutico, estos equipos también deben disponer de sistemas de esterilización y limpieza para evitar la contaminación cruzada o la pérdida de esterilidad que afecte a los siguientes lotes de producción (CIP, siglas del inglés *Clean in place* y SIP, siglas del inglés *Sterilization in place*).¹³ También deben disponer de un sistema hidráulico con el fin de taponar los viales dentro de la cámara al finalizar el ciclo de liofilización y, de esta manera asegurar la esterilidad del medicamento.¹⁸

1.3. Atributos críticos de calidad de un medicamento parenteral liofilizado

Al terminar el ciclo de liofilización, se debe comprobar que los atributos críticos de calidad (CQA, siglas del inglés *Critical Quality Attribute*) fisicoquímicos del producto liofilizado (aparición de la pastilla, contenido de agua residual, pH, tiempo de reconstitución, riqueza del producto e impurezas, entre otros) cumplen con las especificaciones.^{29,46}

Sin embargo, estos atributos críticos de calidad pueden variar durante el ciclo de vida o el almacenamiento del producto. Las variaciones se deben principalmente a tres factores externos: el contenido de agua residual del producto, las condiciones de almacenamiento (temperatura y humedad relativa) y la mezcla de gases que se encuentran en el espacio de cabeza del vial.⁴⁷

La presente tesis se centra en desarrollar nuevos métodos analíticos para determinar el impacto de estos factores en los productos parenterales liofilizados.

2. KARL FISCHER

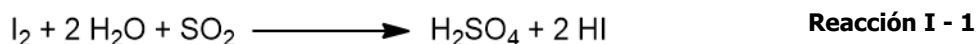
En la industria farmacéutica, y más especialmente en la fabricación de productos farmacéuticos liofilizados, es esencial la determinación del contenido de humedad en el producto final. En estos productos el contenido de agua es considerado un CQA, debido a que aporta información sobre el proceso de liofilización y es un indicador de la calidad final del producto. Por ello, la determinación de esta molécula es uno de los análisis más habituales en la industria farmacéutica de liofilizados.^{48,49}

Dado que el agua es un compuesto presente en una gran variedad de muestras (gases, sólidos insolubles, líquidos, polvos ...), no se dispone de un método universal para su determinación. Por lo tanto, la elección de un método de análisis adecuado para determinarlo depende de las propiedades de la muestra a analizar.^{50,51}

En este apartado se describe uno de los métodos más importantes para la determinación del agua, la valoración Karl Fischer, que ha sido empleado durante el desarrollo de esta tesis.⁵²

2.1. Introducción

La valoración Karl Fischer es un método de análisis directo, basado en una reacción química selectiva, que permite la cuantificación del agua en la muestra, independientemente de su procedencia (superficial o unida). Presenta un seguido de ventajas frente a otros métodos directos usados para la determinación de agua (métodos de secado): es un método específico, rápido, automatizable, permite validar los procedimientos analíticos y salvaguarda la trazabilidad de los análisis. Estas características hacen que sea un método muy interesante para la industria farmacéutica, pudiéndose adaptar a la estricta normativa de calidad a la que está sujeta.^{49,50,52}

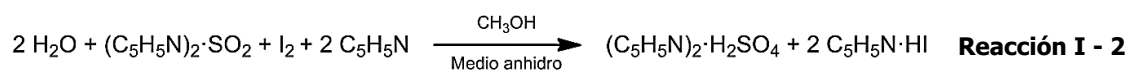


Este método lleva el nombre de su inventor, el químico alemán Karl Fischer (1901-1958), quien, en 1935, publicó el trabajo sobre la determinación del agua a partir de una adaptación de la reacción de Bunsen (Reacción I - 1) usada para determinar el dióxido de azufre (SO₂) en soluciones acuosas.^{53,54} Desde entonces, se ha convertido en un método de análisis de referencia para la determinación del agua en laboratorios de todo el mundo, y se especifica en las farmacopeas líderes del sector, como la de Estados

Unidos (USP, siglas del inglés *United States Pharmacopeia*) o la Europea (Ph. Eur. abreviación del inglés *European Pharmacopoeia*).⁵⁵⁻⁵⁷

2.2. Fundamento teórico

Karl Fischer se dio cuenta que realizando la reacción de Bunsen, en un medio anhidro metanol y con un exceso de dióxido de azufre, podía utilizarse para cuantificar el agua en una muestra. Para ello necesitaba neutralizar el ácido sulfúrico, producto de la reacción, y asegurar que la reacción se completara; con este propósito empleó como base la piridina.^{49,53} En la Reacción I - 2 se puede observar la reacción que Karl Fischer presentó para determinar el contenido de agua:



En esta reacción, el metanol se utiliza como solvente y se observa una relación molar de 2:1 de agua y yodo. Sin embargo, como se demostró en estudios posteriores, esta relación estequiométrica era incorrecta, debido a que el metanol no actuaba meramente como solvente, sino que también desempeñaba un papel como reactivo.⁵⁸⁻⁶⁰ Actualmente, todavía existe polémica para definir el mecanismo y las reacciones que tienen lugar en el medio de una valoración Karl Fischer.⁶¹ A pesar de ello, la reacción global se puede definir como:



Donde Z es una base y ROH es un alcohol. Comúnmente se utiliza imidazol o piridina y metanol, respectivamente; aunque se pueden utilizar otros compuestos con el fin de evitar reacciones secundarias con la muestra, o para mejorar la solubilidad de ésta. Al considerar el metanol (u otro alcohol, ROH) como un reactivo, la relación molar pasa de 2:1 (Reacción I - 2) a 1:1 (Reacción I - 3). Siguiendo esta reacción, cuando toda el agua del medio ha reaccionado, se produce un exceso de yodo que causa un cambio de coloración, de incoloro a amarillento hasta marrón intenso. Esto permite determinar de forma visual el punto final de la reacción. Sin embargo, debido a que esta coloración puede variar dependiendo de la muestra analizada o de los reactivos empleados, la determinación del punto final de forma visual se considera una determinación muy subjetiva, lo que dificulta la validación de los métodos e incluso la automatización de la técnica.^{48,49}

Considerando las desventajas que presenta el método visual, actualmente la mayoría de las valoraciones Karl Fischer realizan la determinación del punto final electroquímicamente. Para ello, controlan la reacción redox del par iónico I_2/I^- utilizando dos electrodos indicadores metálicos inertes a este tipo de reacciones, normalmente de platino.^{51,62} Este control se puede llevar a cabo bipotenciométricamente o biamperométricamente.

Una determinación bipotenciométrica consiste en aplicar una corriente constante y medir el voltaje conforme los electrodos se van polarizando. Esto implica que cuando la concentración de yodo es baja (hay mucha agua en el medio de valoración) se requiere un gran voltaje para mantener los electrodos polarizados, y cuando la concentración de yodo aumenta (no queda agua en el medio de valoración) el voltaje desciende, indicando el punto final de la reacción.

En una determinación biamperométrica en cambio se mantiene un voltaje constante y se mide la corriente durante la polarización. Por lo tanto, al introducir la muestra (concentración elevada de agua) la concentración de yodo es baja, igual que la intensidad de corriente, y cuando toda el agua ha reaccionado, existe un exceso de yodo en el medio de valoración y la intensidad de corriente aumenta, indicando que la reacción ha terminado. Estos métodos electroquímicos permiten una mejor exactitud y precisión en los análisis, facilitando la validación de los métodos.^{48,50,63}

Con la finalidad de poder determinar el contenido de agua en el mayor tipo de muestras, se han ido desarrollando diferentes modificaciones de la valoración Karl Fischer inicial. En las siguientes secciones se describirán las dos vertientes más importantes de este método, la valoración volumétrica y la culombimétrica. Actualmente se utilizan ambas, dado que presentan propiedades complementarias, en la Tabla I - 1 se muestra una breve comparativa entre ellas.

Tabla I - 1. Comparación entre la valoración de Karl Fischer volumétrica y la culombimétrica. Tabla adaptada de la referencia.⁴⁸

Características del método	Valoración volumétrica	Valoración culombimétrica
Fuente de yodo	Reactivo valorante	Oxidación del yoduro <i>in situ</i>
Intervalo de concentración de agua a determinar	10 ppm hasta 100 %, con una cantidad de muestra de 15 g y 0,02 g, respectivamente.	1 ppm hasta 5 %, con una cantidad de muestra de 10 g y 0,05 g, respectivamente.

Características del método	Valoración volumétrica	Valoración culombimétrica
Ventajas	Se pueden usar reactivos valoradores con distintas concentraciones.	Alta exactitud.
	La muestra se introduce directamente en la celda de valoración independiente de su estado físico.	Al generar el yodo <i>in situ</i> no requiere estandarización de la solución valorante.
Desventajas	La concentración de yodo de la solución valorante se debe estandarizar frecuentemente.	Las muestras sólidas se deben disolver previamente, para introducirse en la celda de valoración.
	La matriz de la muestra no puede reaccionar con los reactivos de la valoración.	
	Los solventes se deben renovar frecuentemente.	

2.2.1. Valoración volumétrica

Una valoración volumétrica consiste en determinar el volumen necesario de una disolución de concentración conocida (estandarizada) para que reaccione completamente con el analito estudiado.⁶² La valoración volumétrica de Karl Fischer es la modificación directa del método original (1935).

En los inicios de la técnica, la muestra se solubilizaba en un solvente, previamente titulado, y se determinaba la cantidad de agua empleando una solución valorante de metanol anhidro, la cual contenía todos los reactivos necesarios (Reacción I - 2).⁴⁸ Sin embargo, la solución valorante presentaba una alta inestabilidad que requería su preparación diaria.⁵³ Este inconveniente fue solventado en 1947 por A. Johansson, quien dividió la solución valorante en dos, una en la que se disuelve la muestra, que contiene la base (Z) y el dióxido de azufre disueltos en metanol, y otra solución de metanol con concentración conocida de yodo, que sirve como disolución estándar o solución valorante (Reacción I - 3).^{48,62,64}

El uso de dos soluciones aumentó la estabilidad de los reactivos, aunque la concentración de yodo de la solución valorante debe ser determinada periódicamente, siendo una de las principales desventajas de la valoración volumétrica Karl Fischer.⁶²

Esta modalidad de valoración se utiliza para muestras con un contenido elevado de agua, desde 10 ppm hasta el 100 %, con una cantidad de muestra de 15 g y 0,02 g, respectivamente. No obstante, es posible determinar concentraciones de agua más bajas

modificando la concentración de la solución valorante, la precisión de la bureta con la que esta se añade o la cantidad de muestra que se utiliza. Para cuantificar el contenido de agua con una valoración volumétrica Karl Fischer, es necesario que la muestra se disuelva completamente, independientemente de si se trata de muestras líquidas, sólidas y semisólidas. Una mala disolución de la muestra dificultaría la liberación del agua al medio de valoración, proporcionando resultados erróneos. Sin embargo, este tipo de valoración permite el uso de un amplio abanico de disolventes orgánicos (polares o no polares) para aumentar el grado de disolución de las muestras.⁴⁸⁻⁵⁰

2.2.2. Valoración culombimétrica

Como alternativa a la valoración volumétrica A. S. Meyer y C. M. Boyd presentaron en 1954 una modificación de la técnica basada en la generación culombimétrica del yodo en la celda de valoración según la semirreacción Reacción I - 4.^{48,65}



Una valoración culombimétrica emplea un consumo conocido de corriente eléctrica para determinar la concentración de analito. Por lo que siguiendo la ley de Faraday se puede relacionar la corriente generada con la cantidad de agente valorante necesario (yodo en este caso) para llevar a cabo la reacción.^{48,49} Para este tipo de valoraciones, en lugar de emplear una bureta, como en la valoración volumétrica, se mide el tiempo requerido para completar la reacción.⁶²

En este caso, el yodo se genera *in situ* en la celda de valoración por la oxidación anódica del yoduro (I^-) (Reacción I - 4), y la reducción de protones (H^+) en el cátodo.⁶⁶ La generación del agente valorante directamente en el medio de trabajo, proporciona una mejora en el método aumentando su precisión.⁴⁸ Una vez el yodo ha sido generado en un electrodo generador, el proceso químico de cuantificación del agua es el mismo que en la valoración volumétrica (Reacción I - 3).^{67,68}

La valoración de Karl Fischer culombimétrica se emplea para determinar el agua en muestras en las que esté a baja concentración. Esta técnica tiene un intervalo de trabajo que va, aproximadamente, desde 1 ppm hasta el 5 %, con una cantidad de muestra de 10 g y 0,05 g, respectivamente, por lo que se podría considerar como un método de trazas.⁵⁷ Al contrario que en la valoración volumétrica, la valoración culombimétrica sólo puede determinar la humedad de muestras líquidas o gaseosas, por lo que la

determinación de agua de muestras sólidas requiere una disolución previa con el reactivo de la misma celda o con un disolvente previamente titulado.^{48,49}

2.3. Instrumentación

Habitualmente, los instrumentos de valoraciones de Karl Fischer constan de un recipiente cerrado (celda de valoración) con una abertura para la introducir la muestra, un agitador magnético para solubilizarla y homogenizar la solución, un electrodo indicador para determinar el punto final de la reacción y un sistema para introducir/generar el yodo (dependiendo de la valoración empleada).⁶³

La reacción de valoración (Reacción I - 3) se controla monitorizando la señal del electrodo indicador, y el sistema introduce/genera yodo dependiendo de la concentración del agua sin reaccionar de la celda valoración.

En la Figura I - 6 se puede observar la configuración de los instrumentos de valoración de Karl Fischer volumétricos y coulombimétricos.

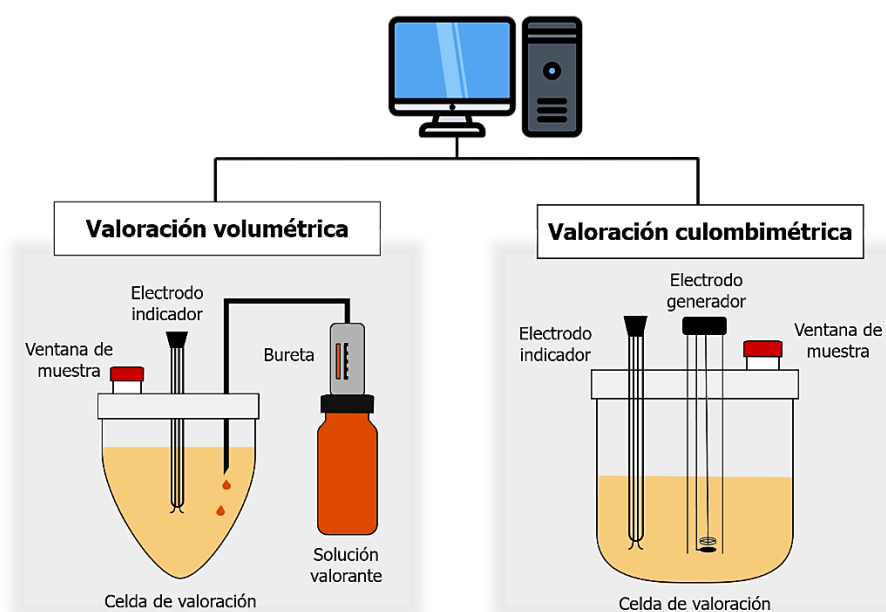


Figura I - 6. Esquema de las partes principales de un instrumento de valoración Karl Fischer tanto volumétrica como coulombimétrica.

2.3.1. Karl Fischer acoplado a un horno

Como se ha comentado anteriormente, con ambas valoraciones (volumétrica y coulombimétrica) se pueden analizar una amplia variedad de muestras (sólidas, líquidas, semisólidas, gaseosas...). Aun así, estas muestras tienen que presentar unas características específicas: poder disolverse en el medio de valoración y no presentar

reacciones secundarias con éste. Dichas restricciones excluyen una gran cantidad de muestras. Este problema puede solventarse empleando distintos reactivos para evitar estas reacciones interferentes o añadiendo disolventes que permitan mejorar la solubilidad. A pesar de esto, hay muestras que aun así siguen quedando fuera del alcance de las valoraciones.

Para este grupo de muestras, se emplea una combinación de los métodos de secado y de la valoración Karl Fischer. Esta combinación permite resolver los problemas de selectividad de los métodos de secado, evita el contacto directo de la muestra con el medio de valoración Karl Fischer (minimiza las reacciones secundarias) y no requiere la solubilización de la muestra. El método resultante consiste en calentar la muestra (métodos de secado) y transportar los vapores resultantes a la celda de valoración con un gas portador seco e inerte para la muestra (aire seco, nitrógeno (N₂), argón, ...), consiguiendo así que no haya interferencias.^{48,63} La combinación de ambas técnicas se conoce como valoración Karl Fischer acoplada a un horno, y puede emplearse tanto en valoraciones volumétricas como coulombimétricas. En la Figura I - 7 se puede observar un esquema del instrumento, indicando sus distintos elementos.

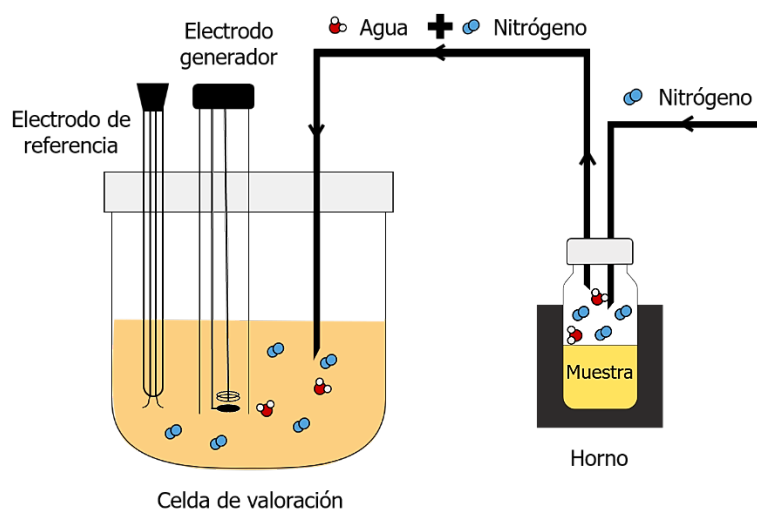


Figura I - 7. Esquema de una valoración Karl Fischer coulombimétrica acoplada a un horno.

Aunque esta técnica permite ampliar el abanico de muestras que pueden analizarse con la valoración Karl Fischer, sigue habiendo muestras que pueden generar problemas. La causa principal de esta exclusión se debe a la selección de una temperatura óptima para calentar la muestra.⁶³ Esta temperatura debe ser lo suficientemente elevada para que se evapore el agua de la muestra (superficial y unida) sin llegar a degradar por combustión el producto. En caso contrario, como subproducto se obtendría agua, lo que se traduciría con un resultado mayor que el esperado.

Para hallar la temperatura óptima del horno para la determinación del agua, se mide la respuesta electroquímica de la reacción de Karl Fischer mientras se aumenta gradualmente la temperatura de la muestra, realizando una rampa de temperatura, hasta llegar a su descomposición. En la Figura I - 8 se muestra un ejemplo de rampa de temperatura; en ella se puede observar dos picos al inicio de la rampa debidos a la liberación del agua (superficial y unida) y un tercer pico debido a la descomposición del producto. Dado que los picos de liberación del agua y de degradación están separados, se puede seleccionar una temperatura entre ambos como temperatura óptima para este producto. El principal inconveniente de este método es que no todas las muestras presentan una separación de picos tan clara como en la figura, y es posible que la temperatura de descomposición esté muy próxima a la temperatura de liberación del agua.

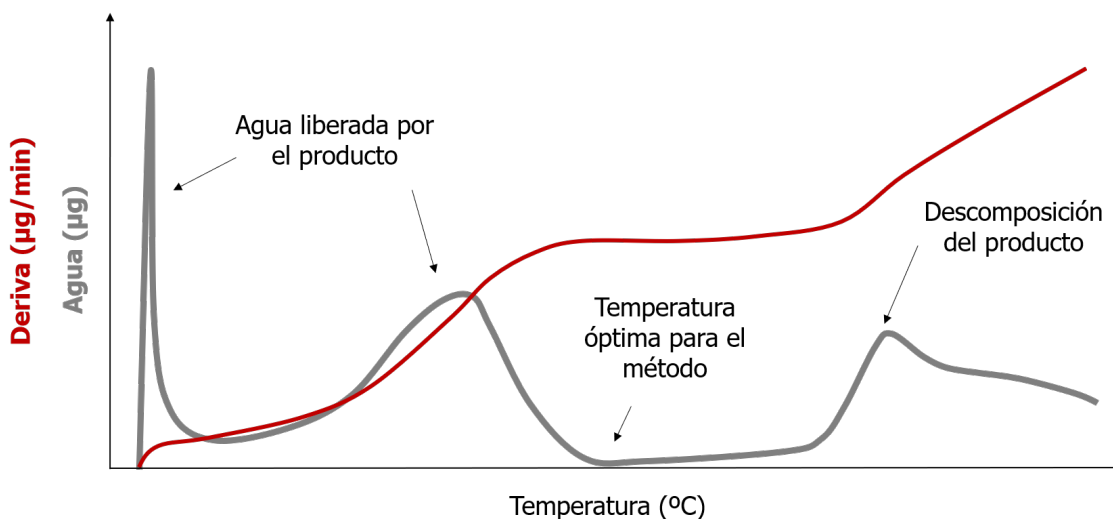


Figura I - 8. Ejemplo de rampa de temperatura de un producto, en el cual se puede ver como al aumentar la temperatura se libera el agua del producto y posteriormente se descompone.

3. ESPECTROSCOPIA DE ABSORCIÓN EN EL INFRARROJO CERCANO

El incremento de los requisitos que exige la regulación relacionados con la calidad de los productos farmacéuticos liofilizados ha potenciado implementación de técnicas rápidas y no destructivas.⁶⁹ Una de ellas es la espectroscopia de absorción en el infrarrojo cercano (NIR, siglas del inglés *Near-Infrared*). El uso de esta técnica se ha incrementado en las últimas décadas y ha ido sustituyendo otras técnicas empleadas tradicionalmente (HPLC, valoración Karl Fischer, etc.) para realizar el análisis del producto.^{70,71}

En este apartado se describe esta técnica espectroscópica utilizada en el desarrollo de esta tesis.

3.1. Introducción

En 1666 Isaac Newton introdujo el término espectro y el inicio de la espectroscopia, con su experimento de la descomposición de la luz solar en una serie continua de colores al pasar a través de un prisma.^{72,73} Desde entonces, se puede definir la espectroscopia como el estudio de la interacción (absorción, emisión o dispersión) de las ondas electromagnéticas con la materia.⁷⁴

Años después de este hito histórico, se descubrieron las zonas del espectro electromagnético contiguas a la franja visible. En el año 1800 Fredrick William Herschel descubrió la zona del infrarrojo y en 1801 Johann Wilhelm Ritter la zona del ultravioleta (conocida por las siglas UV).⁷⁵ Herschel demostró que la dispersión de la luz se extendía de forma no visible para el ojo humano por la zona inferior al color rojo, a la que nombró infrarrojo.^{76,77} No obstante, lo que realmente detectó Herschel fue el infrarrojo cercano (NIR), debido a que el prisma de cristal que empleó para descomponer la luz absorbió la región de la radiación del infrarrojo medio (MIR, siglas del inglés *Mid-Infrared*).⁷⁸

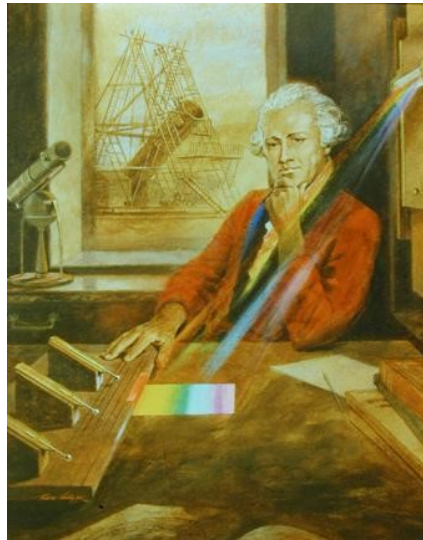


Figura I - 9. En esta imagen aparece Fredrick William Herschel, junto a una representación del experimento que le permitió descubrir el infrarrojo. Imagen extraída de la referencia http://www.observadores-cometas.com/Herschel/About_William_Herschel/About_William_Herschel/Biography/Who_was_Herschel.htm.

Sin embargo, a pesar de su temprano descubrimiento, la espectroscopia basada en la región NIR quedó en un segundo plano frente a la espectroscopia ultravioleta, la visible e incluso la MIR.⁷⁹ Esta desatención se debió principalmente a que la información ofrecida por las estrechas bandas fundamentales obtenidas por la espectroscopia MIR, era más fácil de interpretar que la obtenida a partir de bandas anchas y solapadas de la espectroscopia NIR.^{79,80} A esto, se debe añadir la baja intensidad de absorción de los compuestos orgánicos e inorgánicos en esta región del espectro y el poco desarrollo de los instrumentos de medición de la época. Todos estos factores se ven reflejados en los escasos estudios científicos realizados referentes a la medición de espectros en la zona NIR desde el descubrimiento de la región hasta el año 1930.^{70,80}

Gracias al desarrollo tecnológico producido durante la Segunda Guerra Mundial, y a algunos trabajos pioneros en la determinación cuantitativa de sustancias en longitudes de onda NIR concretas, el estudio de la región NIR empezó a tomar interés.^{70,78,80} A pesar de ello, dado el impulso de otras técnicas, estos métodos basados en la espectroscopia NIR fueron prácticamente abandonados.⁸¹ No fue hasta su aplicación en el sector de la alimentación y agricultura por Ben-Gera y Norris, a finales de los años 60, que se consolidó la espectroscopia NIR como una técnica analítica con un gran potencial.^{80,81}

Desde entonces, se ha considerado la espectroscopia NIR, según la expresión de McClure, como "el gigante dormido de la química analítica".⁸² Este reconocimiento se debe a las ventajas que presenta frente a otras técnicas, como poder analizar todo tipo

de muestras (líquidas, sólidas, semisólidas y gaseosas) con un nulo o mínimo pretratamiento de la muestra, la rapidez con la que se puede realizar un análisis, o que permite el análisis de más de una propiedad a partir de un mismo espectro, entre otras.^{83,84} En vista de todas estas ventajas, actualmente cuenta con un número considerable de aplicaciones, una de las cuales es en el ámbito del control de la calidad en la industria farmacéutica.

Aunque en la industria farmacéutica se empezó a estudiar la aplicación de NIR ya en la última parte de los años 60, no fue hasta 1986 cuando la agencia reguladora de EE. UU. la FDA (FDA, siglas del inglés *Food and Drug Administration*) aprobó el primer método analítico basado en ella para determinar la cantidad de agua y de Lincomicina en un producto liofilizado de ámbito veterinario.^{70,85,86} Desde aquel momento, la implementación de la técnica NIR en este sector se ha ido consolidando, tanto en el análisis cuantitativo como cualitativo.⁶⁹ A pesar de ello, la principal dificultad que limita la expansión en la implementación de esta técnica analítica en este sector es el uso imprescindible de la quimiometría (ver apartado 5) y la validación de los métodos desarrollados, debido a que requieren de personal cualificado con una formación específica.⁷¹

3.2. Fundamento teórico

En la Figura I - 10 se muestra un esquema del espectro electromagnético donde se pueden observar todas las regiones que lo definen.⁸⁷

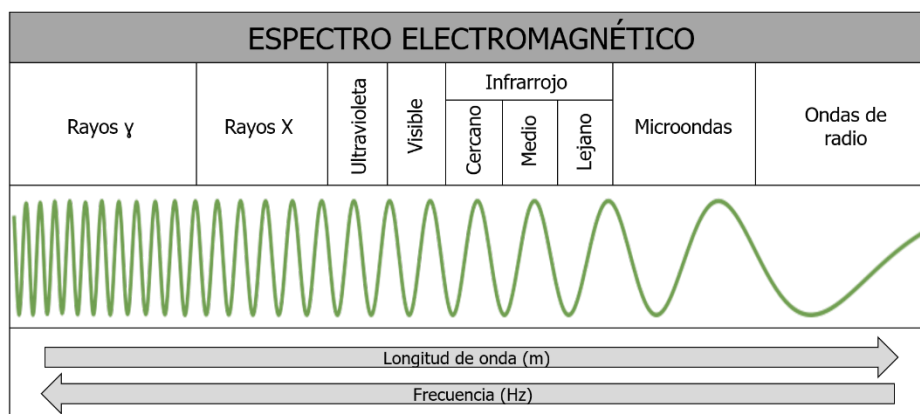


Figura I - 10. Espectro electromagnético con las distintas zonas que lo conforman delimitadas, indicando la variación de la longitud de onda y de la frecuencia.

La región del infrarrojo se encuentra situada entre la zona del visible y la de las microondas, entre 760 nm y 111000 nm aproximadamente, y está relacionada con la energía vibracional de las moléculas. Debido a la gran variabilidad de aplicaciones e

instrumentación necesaria para abarcar la totalidad de la región del infrarrojo, ésta se divide en tres zonas.⁸⁸

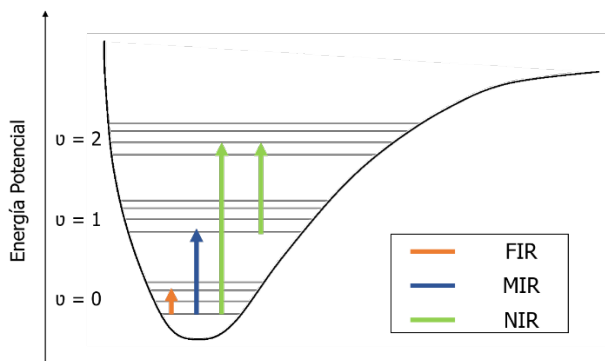


Figura I - 11. Representación de las absorciones de las diferentes zonas del infrarrojo, figura adaptada de la referencia.⁸⁹

- Infrarrojo lejano (FIR, siglas del inglés *Far-Infrared*): esta zona comprende desde 16000 nm hasta 111000 nm, aproximadamente. En ella aparecen vibraciones no armónicas de gran amplitud, vibraciones de redes cristalinas y transiciones de espectros de vibración-rotación.^{89,90} Pocas sustancias orgánicas absorben en este rango de longitudes de onda, que se caracteriza por tener un espectro de absorción con bandas poco intensas y principalmente de moléculas muy grandes, como sustancias organometálicas o inorgánicas.⁸⁷
- Infrarrojo medio (MIR): esta zona se encuentra definida entre 2500 nm y 16000 nm. Aquí aparecen las vibraciones fundamentales de los enlaces moleculares, produciendo bandas de una intensidad mayor que las de FIR. La mayoría de moléculas presentan absorción en esta región del espectro electromagnético.^{87,91}
- Infrarrojo cercano (NIR): esta zona va desde el final del visible hasta los 2500 nm.^{II} En esta zona aparecen las bandas de absorción de las transiciones desde el estado fundamental de vibración al segundo estado excitado o superior, lo que se conoce como sobretonos o armónicos, y las debidas a la excitación simultánea de dos o más estados de vibración, conocidas como bandas de combinación.^{87,91}

Aunque todas estas zonas tienen un amplio interés analítico, en esta tesis se ha utilizado la espectroscopia de absorción NIR.

^{II} De acuerdo con la comisión internacional de la Iluminación (CIE, siglas del francés *Commission Internationale de l'Eclairage*) no es posible definir un límite preciso entre las regiones del visible y el infrarrojo. Sin embargo, toda la bibliografía revisada define este límite entre 700 y 800 nm.

El modelo clásico que describe la vibración de un enlace entre dos átomos es el del oscilador armónico. Este modelo sigue la ley de Hook, donde la relación entre la energía potencial de vibración y la distancia de enlace es una función parabólica centrada en la distancia de equilibrio (Figura I - 12). Si se añade que la energía no es continua, sino que está cuantificada y depende del número cuántico vibracional, ν , se obtiene el modelo cuántico ideal de la vibración. En este modelo sólo están permitidas las transiciones entre estados vibracionales consecutivos ($\Delta\nu = \pm 1$) conocidas como bandas fundamentales.^{78,80,92}

No obstante, en la práctica, el modelo de oscilador armónico ideal presenta algunas limitaciones:

- Las nubes electrónicas de los átomos del enlace crean una barrera de potencial, debida a la repulsión entre ellos, lo que dificulta la compresión de la vibración.⁸⁰
- Contrariamente, si la energía es suficientemente grande, el enlace se rompe.⁹³

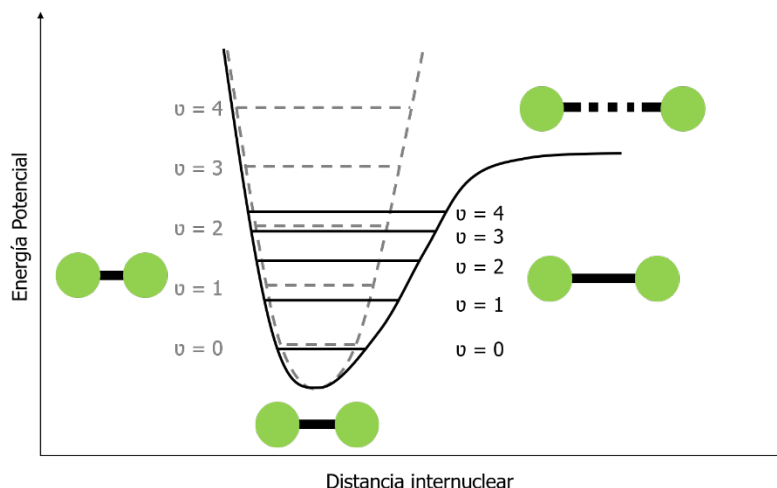


Figura I - 12. Representación de los distintos modelos para describir la vibración molecular entre dos átomos, en gris se puede ver el modelo clásico y en negro el modelo de oscilación no armónico.

Para solventar estas limitaciones, es necesario considerar la no armonicidad de la vibración. Esto significa que la distancia energética entre niveles es distinta, decreciendo a medida que aumenta la energía. También implica la aparición de vibraciones "prohibidas" llamadas armónicos o sobretonos, con transiciones $\Delta\nu = \pm 2, \pm 3, \dots$, que aparecen a longitudes de onda menores que la banda fundamental. Las bandas fundamentales se observan principalmente en el espectro de absorción MIR, mientras que las bandas de los sobretonos aparecen en la región del NIR y son menos intensas que la banda fundamental correspondiente. Esto se debe a que la probabilidad de transición disminuye al aumentar el número cuántico asociado al nivel de vibración, lo

que se traduce en una disminución de la intensidad.^{93,94} La mayor parte de los sobretonos en el NIR están relacionados con el estado de vibración de tensión (en inglés *Stretching*) de los enlaces X—H presentes en la mayoría de moléculas orgánicas (donde X puede ser C, O, N o S principalmente). Estos sobretonos presentan una mayor energía cuanto mayor es la diferencia de masa entre los átomos del enlace.⁸⁹

Este modelo de oscilación no armónico también permite considerar la combinación de dos o más vibraciones para producir una única banda, lo que se conoce como bandas de combinación. Estas bandas, juntamente con los sobretonos, son las bandas de absorción observadas en el espectro NIR.

Independientemente del modelo oscilatorio empleado, para que una vibración molecular sea activa en el espectro de absorción NIR, ésta tiene que aportar un cambio en el momento dipolar de la molécula ($\Delta J \neq 0$).⁸⁹

La temperatura y la humedad ambiental, la interacción de las moléculas entre ellas (por ejemplo, puentes de hidrogeno), el tamaño de la partícula, el grosor de la muestra y el estado físico en que se encuentran las muestras pueden hacer variar la posición y la intensidad de las bandas del espectro de absorción NIR.⁸⁷

3.3. Instrumentación

En términos de instrumentación, un espectrofotómetro NIR no se distingue mucho de un espectrofotómetro visible o infrarrojo.⁹⁴ Por este motivo, el desarrollo de la instrumentación NIR va muy entrelazada con ambas técnicas.

A principios del siglo XX no existía la idea de tener un espectrofotómetro dedicado sólo a la zona NIR. Dado el bajo interés de esta región, la medición en la zona NIR era un accesorio de otros instrumentos de UV-visible o del MIR.⁷⁰

A pesar de esto, gracias al desarrollo de detectores fotoeléctricos con sulfuro de plomo (II) (PbS), a los esfuerzos realizados por el grupo de Karl Norris y al creciente interés en la espectroscopia NIR, nacieron los espectrofotómetros NIR como equipos independientes para su uso en el laboratorio.^{74,82}

La configuración de los espectrofotómetros NIR varía dependiendo de la muestra (líquida, sólida, ...), de la resolución con la que se quiera registrar el espectro y de la aplicación de la técnica. Por lo tanto, dadas las numerosas aplicaciones que existen actualmente, no es de extrañar que existan distintas configuraciones para poder dar respuesta a la demanda.⁹⁵ A pesar de la flexibilidad de estos instrumentos y, aunque la

configuración del equipo varíe, todos los espectrofotómetros NIR cuentan con los siguientes bloques: fuente de radiación, selector de longitud de onda, compartimiento de la muestra y detector.^{74,83} En las siguientes secciones se detalla cada una de estas partes.

3.3.1. Fuente de radiación

La fuente de radiación es la parte del instrumento que genera el haz de radiación que interactúa con la muestra. Esta puede ser continua, emitiendo en todo el espectro, o discreta, y sólo emitir a una o unas pocas longitudes de onda.^{83,88}

Las fuentes de radiación continuas abarcan un amplio intervalo del espectro electromagnético, idealmente, la totalidad de la zona espectral del NIR.⁹⁶ Habitualmente, se emplean las lámparas de halógeno con filamentos de tungsteno que producen una radiación térmica.⁹⁷ Se caracterizan por ser económicas y tener una intensidad de luz estable.⁷⁴

Las fuentes de radiación discreta irradian la muestra con unas pocas longitudes de onda de la zona del espectro electromagnético del NIR y se emplean en aplicaciones donde no es necesario registrar la totalidad del espectro. Utilizan fuentes de radiación no térmicas, como los láseres de diodo NIR o los LED NIR (LED, siglas del inglés *Light-Emitting Diode*) a longitudes de onda fijas.^{74,83} Se caracterizan por ser más eficientes en términos energéticos y por la facilidad con la que se pueden adaptar a instrumentos portátiles sin perder robustez.^{96,97} Una aplicación de este tipo de fuentes se describe en el apartado 4.

3.3.2. Selector de longitudes de onda

La longitud de onda con que se tiene que irradiar la muestra se debe seleccionar a partir del espectro continuo que proporciona la fuente de radiación.^{88,98}

Existen distintas clases de selectores de longitudes de onda y cada una de ellas da lugar a un instrumento NIR de características muy distintas.

- **Filtros:** este tipo de selectores consisten en un elemento que interacciona con la radiación procedente de la fuente, bloqueando parte de ésta y permitiendo la iluminación de la muestra con una única longitud de onda. Generalmente se utilizan diversos filtros para conseguir instrumentos que trabajen a longitudes de onda específicas. Estos selectores tienen un bajo precio aunque, según cual sea

la transmitancia del filtro, pueden reducir apreciablemente la intensidad de la señal que llega a la muestra.^{88,97}

- Monocromadores: este tipo de selectores están compuestos por una ranura de entrada, un elemento dispersivo (generalmente una red de difracción) que separa la luz en las longitudes de onda que la constituyen, una ranura de salida, y una rendija que aísla las longitudes de onda deseadas para irradiar la muestra. Estos selectores permiten irradiar la muestra en todo el intervalo de longitudes de onda deseado, lo que ofrece una mayor flexibilidad a los instrumentos que en el caso de utilizar filtros.^{74,88}
- Transformada de Fourier: este tipo de selectores consisten en la codificación de la información de la fuente de radiación empleando un interferómetro (el más empleado es el de Michelson). Esta codificación permite realizar de forma simultánea la medida de varias longitudes de onda, y posibilitan la obtención de una alta precisión y reproducibilidad en los resultados.^{95,96}
- Filtro acústico-óptico (AOTF del inglés *Acousto-Optic Tunable Filters*): este tipo de selectores permite escoger las longitudes de onda de todo el intervalo del espectro electromagnético de la zona NIR usando señales de radiofrecuencia y un cristal de óxido de telurio (IV) (TeO₂). Los AOTF ofrecen a los espectrofotómetros NIR una gran rapidez en la medida y, mejor reproducibilidad y robustez que los sistemas anteriores, gracias a la ausencia de partes móviles.^{83,96,98}

3.3.3. Presentación de la muestra

Como se ha mencionado anteriormente, uno de los puntos fuertes de la espectroscopia NIR es que se pueden analizar una gran variedad de muestras, independientemente de su estado físico y con una manipulación mínima o nula. Por lo tanto, el diseño de esta parte del equipo depende tanto del tipo de muestra como del modo de registro que se emplee.^{80,96,97}

Existen distintos modos para registrar el espectro de las muestras. Los más habituales son:

- Transmitancia: utilizando este modo de registro, la muestra se coloca en un compartimiento entre la fuente de radiación y el detector (Figura I - 13).⁸⁰ De esta manera, la radiación pasa a través de la muestra y la luz que no ha sido

absorbida se detecta al otro lado. En este caso, la absorción de la muestra sigue la ley de Lambert-Beer:

$$A = \log \frac{1}{T} = \log \frac{I_0}{I} \quad \text{Ecuación I - 1}$$

Donde A es la absorbancia, T la transmitancia, I_0 la intensidad de la radiación de la fuente de radiación e I la intensidad de la radiación que no es absorbida por la muestra.

Este modo de registro se emplea en muestras transparentes, como líquidos o gases, las cuales se introducen en una cubeta de un material transparente a la radiación NIR y que mantienen constante el camino óptico. También existen equipos diseñados para la lectura de comprimidos sólidos por transmitancia.^{94,96,97}

- Reflectancia difusa: este modo de registro fue el primero que se empezó a usar en la espectroscopia NIR, gracias a Karl Norris. En las medidas de reflectancia, se estudia la radiación que refleja la muestra. Cuando la luz se refleja con el mismo ángulo que la luz incidente, sin penetrar en la muestra y sin aportar información sobre ella, se conoce como reflectancia especular y se rige por las leyes de Fresnel. La reflectancia difusa aparece en la superficie del sólido y, se propaga en todas las direcciones debido a la absorción y dispersión de la radiación incidente por las partículas del sólido. La reflectancia contiene ambos componentes, el especular y el difuso, y para separarlos se utiliza un detector que forma un ángulo de 45°-60° respecto a la radiación incidente (Figura I - 13). Existen distintas teorías para explicar la reflectancia difusa, pero en la práctica se utiliza la reflectancia relativa (R), que es la relación entre la luz reflejada por la muestra (R_m) y por un estándar (R_{ref}). El estándar es un material con una alta reflectancia absoluta, como placas cerámicas de alta pureza, teflón, sulfato de bario ($BaSO_4$), etc. La relación entre la concentración y la reflectancia relativa es análoga a la ley de Lambert-Beer.

$$A_{aparente} = \log \frac{R_{ref}}{R_m} = \log \frac{1}{R} = a'c \quad \text{Ecuación I - 2}$$

Donde $A_{aparente}$ es la absorbancia aparente, c es la concentración y a' una pseudoconstante de proporcionalidad que contiene la absortividad molar (a) y el

coeficiente de dispersión (s) que depende del tamaño de las partículas del sólido y la longitud de onda de trabajo.

Este modo de registro se emplea con muestras opacas, principalmente sólidos.^{80,94}

- **Transflectancia:** este modo de registro fue desarrollado por Technicon y consiste en un registro que combina los dos modos anteriores. Al igual que en la reflectancia difusa, la fuente de radiación y el detector se colocan en ángulo en el mismo lado respecto a la muestra. Sin embargo, en este modo se coloca una lámina de material reflectante detrás de la muestra y la radiación, después de atravesar la muestra, impacta en el material reflectante y vuelve a atravesarla hasta llegar al detector (Figura I - 13). Este modo de registro se emplea en muestras líquidas, tanto translúcidas como turbias.^{94,96,97}

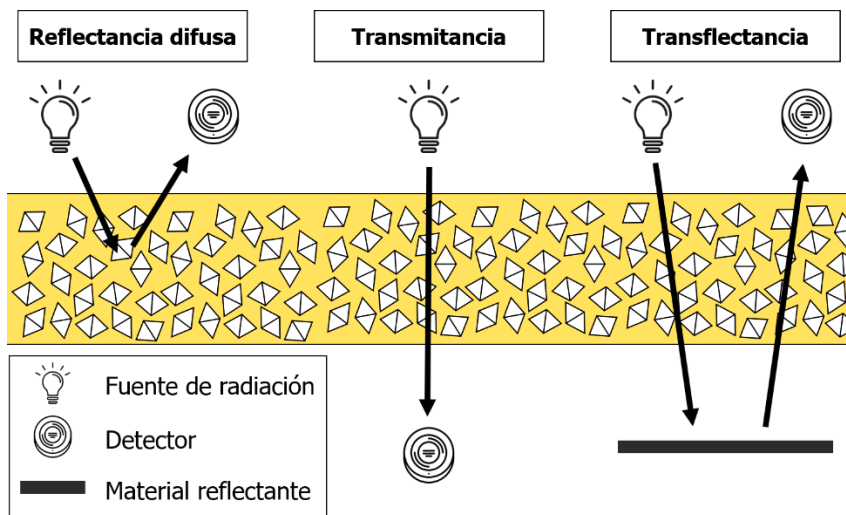


Figura I - 13. Representación de la ubicación de la fuente de radiación y del detector de los principales modos de registro de la espectroscopia NIR.

Todos estos modos de registros pueden adaptarse a medidas *in line* gracias al uso de la fibra óptica, lo que permite medir muestras de forma remota gracias a sus propiedades para transmitir la información.⁷⁰

3.3.4. Detector

El detector es la parte del dispositivo que convierte la energía radiante que proviene de la muestra en una señal eléctrica. El tipo de detector utilizado es crítico en el instrumento, ya que marca el intervalo de longitudes de onda que puede leer el aparato. Se debe conseguir que, con un tiempo de respuesta rápido, las longitudes de onda a las que se

desea trabajar proporcionen una señal con una sensibilidad y relación señal/ruido adecuadas.^{74,88}

Los detectores fotoeléctricos están basados en materiales fotosensibles y semiconductores. Estos materiales absorben la radiación emitida por la muestra, alterando su estado electrónico y produciendo una señal eléctrica.⁸⁸

Los detectores más usados para la región espectral NIR son los basados en silicio (Si), con un intervalo útil desde 400 nm hasta 1100 nm; los de sulfuro de plomo (II), usados entre 1100 nm y 2500 nm; y los de arseniuro de galio e indio (InGaAs) que abarcan el intervalo desde 800 nm hasta 1700 nm.⁸⁰ Algunos equipos combinan la señal en la zona del visible con la del NIR, o bien extienden el intervalo de utilización hasta el MIR. Todos ellos emplean una combinación de detectores.⁹⁷

3.4. Aplicaciones en la industria farmacéutica

Actualmente, la espectroscopia NIR está ampliamente arraigada en la industria farmacéutica, como se ha comentado en la sección 3.1. Esta amplia utilización se basa en el balance positivo de las ventajas sobre las desventajas que presenta esta técnica, y que se resumen en la siguiente tabla.

Tabla I - 2. Recopilación de las ventajas y desventajas de la técnica espectroscópica NIR en el sector farmacéutico.^{69,71,99,100}

Ventajas	Desventajas
- Técnica no destructiva.	- La baja sensibilidad de la señal dificulta la determinación a bajas concentraciones.
- Análisis no invasivo, es decir, se puede analizar el medicamento en el mismo recipiente del producto final.	- El desarrollo de métodos analíticos requiere personal de alta formación.
- Mínima o nula preparativa de la muestra, lo que reduce el error asociado a la manipulación.	- Los métodos analíticos requieren un mantenimiento constante.
- Tiempo de análisis inferior a 1 min.	- Requiere un método de referencia externo para poder desarrollar los métodos analíticos.
- Bajo coste económico, tanto instrumental como procedimental, no requiere solventes.	- Para obtener buenos resultados se necesita una calibración con muchas muestras que cubran la variabilidad natural.
- Análisis de todo tipo de muestras (líquidas, semisólidas, sólidas compactas, sólidas granuladas, gaseosas, ...).	- Dependiendo del modo de registro empleado, la muestra debe tener como mínimo de 5 a 10 mm de espesor.

Ventajas	Desventajas
- Permite el análisis de más de un componente con el mismo espectro.	
- Las medidas se pueden realizar <i>in situ</i> , ya sea <i>on/in/at line</i> .	

Estas características han propiciado el desarrollo de distintos métodos analíticos, que abarcan tanto determinaciones de carácter químico (cuantificación de API, o de solventes, entre otros); como de carácter físico (determinación de polimorfos, control de la homogenización de mezclas, ...).^{71,78,86}

En las siguientes secciones, se describen las aplicaciones de esta técnica espectroscópica empleadas durante el desarrollo de esta tesis.

3.4.1. Cuantificación de la humedad residual en un producto farmacéutico liofilizado

El agua es la molécula más analizada por espectroscopia NIR, debido a que presenta una fuerte intensidad en esta zona espectral, concretamente en las longitudes de onda descritas en la Tabla I - 3.^{78,86} Sin embargo, el máximo de estas bandas puede variar dependiendo del producto o del estado físico de la molécula de agua, así como de otros factores externos como la temperatura de la muestra.¹⁰¹

Tabla I - 3. Principales longitudes de onda de la absorción del agua en la zona NIR.⁸⁵ Donde ν_1 es la frecuencia de vibración de tensión simétrica del enlace O—H, ν_2 es la frecuencia de la vibración de flexión del mismo enlace y ν_3 es la frecuencia de vibración asimétrica de éste.

Longitudes de onda	Características
760 nm ^{III}	$3\nu_1 + \nu_3$
970 nm ^{IV}	$2\nu_1 + \nu_3$
1190 nm	$\nu_1 + \nu_2 + \nu_3$
1450 nm ^V	$\nu_1 + \nu_3$
1940 nm	$\nu_2 + \nu_3$

Esta alta intensidad de las bandas asociadas al agua implica una limitación a las aplicaciones de la técnica NIR para la determinación de compuestos en soluciones

^{III} También se conoce como tercer sobretono de la vibración de tensión del enlace O—H.

^{IV} También se conoce como segundo sobretono de la vibración de tensión del enlace O—H.

^V También se conoce como primer sobretono de la vibración de tensión del enlace O—H.

acuosas, cuya señal queda enmascarada por la del disolvente.^{71,78} Sin embargo, lo que para la determinación de otras especies representa un problema, facilita el desarrollo de métodos analíticos para la cuantificación del agua.¹⁰¹ Así, ya en 1912, F.E. Fowle desarrolló el primer análisis cuantitativo empleando la espectroscopia NIR para determinar el vapor de agua en la atmósfera.¹⁰²

Actualmente, la espectroscopia NIR se presenta como una alternativa analítica a métodos tradicionales para cuantificar el agua, como es la valoración Karl Fischer (descrita en el apartado 2).⁸⁶ Dadas las características resumidas en la Tabla I - 2 y que la utilización de esta técnica en el sector farmacéutico está aceptada por las agencias reguladoras, en las últimas décadas ha aumentado el número de métodos analíticos basados en NIR para la cuantificación del agua.¹⁰¹

Si nos centramos en los productos farmacéuticos liofilizados, el interés por desarrollar métodos analíticos basados en esta técnica es importante, debido al gran impacto de la humedad residual en este tipo de medicamentos (ver sección 1.3). Kamat *et al.* publicaron en 1989 el primer método analítico para determinar la humedad residual de un medicamento parenteral liofilizado de forma no invasiva.¹⁰³ Desde entonces, se han publicados distintos estudios demostrando la idoneidad de la técnica en el sector. En la siguiente referencia se presenta una recopilación de los estudios realizados sobre este tema.¹⁰⁴

3.4.2. Identificación de productos farmacéuticos

Uno de los análisis de rutina realizados en los laboratorios de control de calidad de cualquier empresa farmacéutica es la identificación de todos los materiales que se usen en todas las etapas del proceso; desde la recepción de las materias primas que se van a utilizar para la fabricación, hasta la liberación del producto farmacéutico al mercado.^{86,101,105} En el ámbito farmacéutico, se entiende como identificación a la confirmación, con un rango de exactitud aceptable, de que el producto analizado es lo que realmente se describe en la etiqueta de éste.¹⁰⁶

Para realizar estos análisis de identificación, muchas veces se necesita información proveniente de distintas técnicas analíticas para realizar la confirmación. Estas técnicas analíticas convencionales acostumbran a ser muy laboriosas y destructivas, por lo es necesario encontrar alternativas más rápidas para llevar a cabo estos análisis.⁹² Una de ellas es la espectroscopia NIR combinada con el análisis de datos multivariante.¹⁰⁷

Como ya se ha comentado con anterioridad, los espectros obtenidos en la región NIR se caracterizan por presentar bandas anchas y solapadas.⁸⁶ Esto implica que la extracción de la información para realizar la identificación de los compuestos no se pueda hacer a partir de la mera inspección visual del espectro, identificando las bandas espectrales asociadas a los distintos grupos funcionales de la molécula.¹⁰⁸ Empleando el análisis de datos multivariante cualitativo, es posible clasificar de una forma fácil los espectros de los compuestos desconocidos comparándolos con los espectros que se tienen registrados de compuestos perfectamente identificados y que forman lo que se conoce como "bibliotecas de espectros" (en la sección 5.2 se comentan con más detalle estos métodos de clasificación).^{86,109}

El primer método de identificación en el sector de la industria farmacéutica basado en la espectroscopia NIR fue desarrollado por John Rose en 1982.¹¹⁰ J. Rose utilizó el análisis discriminante para clasificar productos con estructura similar a la Penicilina empleando 19 longitudes de onda. Posteriormente, la implementación de este tipo de métodos fue ganando popularidad.^{78,111}

Actualmente, algunas farmacopeas y agencias reguladoras ya aceptan la técnica NIR para realizar el análisis de identificación tanto en materias primas como en producto terminado.¹¹²⁻¹¹⁴

4. ESPECTROSCOPIA DE ABSORCIÓN LÁSER DE DIODO SINTONIZABLE

La medición de forma precisa, rápida, no destructiva ni invasiva de algunos gases (vapor de agua, CO₂, O₂, CO, NH₃, óxidos de nitrógeno, hidrocarburos...), tiene un gran interés en una amplia gama de sectores industriales, especialmente en la industria farmacéutica de productos farmacéuticos liofilizados. En este sector, la determinación de gases como el oxígeno, el dióxido de carbono o el vapor de agua en el espacio de cabeza (en inglés *headspace*) de los viales, es de gran importancia para asegurar la integridad y estabilidad del producto durante su tiempo de vida útil.

Los sensores basados en la espectroscopia de absorción láser cumplen con todas estas características y son ideales para realizar el análisis del mayor número posible de muestras en un lote. Una de las técnicas más empleadas en este ámbito es la espectroscopia de absorción láser de diodo sintonizable (TDLAS, siglas del inglés *Tunable Diode Laser Absorption Spectroscopy*). A pesar de todas las ventajas comentadas, esta técnica presenta varias fuentes de ruido que, debido a la baja intensidad de las bandas de absorción de los gases, pueden llegar a limitar su sensibilidad.^{115,116} Con el objetivo de conseguir una mayor sensibilidad y mejorar la relación señal/ruido se emplean técnicas de modulación. En éstas, la señal se codifica y detecta a una frecuencia alta, donde el ruido es menor debido a que tiene una relación inversamente proporcional a la frecuencia. Para ello, se puede utilizar la espectroscopia de modulación de la frecuencia (FMS, siglas del inglés *Frequency Modulation Spectroscopy*) relacionada con la espectroscopia de modulación de la longitud de onda (WMS, siglas del inglés *Wavelength Modulation Spectroscopy*) implantada a fuentes TDLAS a principios de la década de 1970. Este tipo de modulación FMS consiste en la codificación de la información de absorción molecular en las bandas laterales del TDLAS y se ha utilizado con este tipo de fuentes desde principios de la década de los 1980. La técnica analítica resultante (TDLAS-FMS) es capaz de medir gases a concentraciones del nivel de trazas (ppm) con una alta relación señal/ruido.^{117,118}

En este apartado se ofrece una descripción de TDLAS-FMS, que ha sido empleada durante el desarrollo de esta tesis.

4.1. Introducción

Los láseres de diodo sintonizables se descubrieron a mediados del 1960 en el MIT Lincoln Laboratory (Massachusetts, EE. UU.), e inicialmente estaban basados en sales de plomo.¹¹⁹ A pesar de esto, no fue hasta el 1970 cuando Hinkley *et al.* demostraron la posibilidad de implementar estos láseres como una fuente de emisión sintonizable para la espectroscopia infrarroja de alta definición. El uso de estos láseres de diodo sintonizables como fuentes de emisión se debe a que son eficientes, fáciles de manipular, robustos, compactos y de bajo coste. Teniendo en cuenta todas estas características y las mejoras significativas que se han realizado en las últimas décadas, tanto en el rendimiento de estas fuentes de emisión como en las técnicas espectroscópicas láser, el establecimiento de la tecnología TDLAS como herramienta para el control de procesos ha aumentado significativamente.¹²⁰⁻¹²³ Actualmente, TDLAS es la técnica de absorción basada en láser más común para la cuantificación de gases en muchos campos (contaminación atmosférica, procesos de combustión, control de procesos industriales...). Además, debido a su fiabilidad de medición, el uso de personal no especializado y la naturaleza no invasiva, la convierten en una técnica muy atractiva para la industria farmacéutica.^{124,125}

4.2. Fundamento teórico

A diferencia de las muestras sólidas y líquidas, las cuales muestran un espectro de absorción con bandas amplias, suaves y muy solapadas; las muestras gaseosas presentan un espectro de absorción caracterizado por bandas muy estrechas, pronunciadas y con un solapamiento mínimo.^{123,126} Si se registra el espectro de una muestra gaseosa empleando un espectrómetro de alta resolución, se puede observar cómo cada una de estas bandas está formada por un elevado número de líneas muy cercanas entre sí. Esta estructura tan definida y diferenciada, se debe a que es posible apreciar tanto los niveles energéticos vibracionales, con un cambio en el número cuántico vibracional (Δv), como los distintos niveles energéticos rotacionales, con un cambio en el número cuántico rotacional (ΔJ). Las transiciones vibracionales con el mismo número cuántico rotacional ($\Delta J = 0$) forman un conjunto de transiciones nombradas como rama Q. Las transiciones con $\Delta J = +1$ se conocen como rama R y aparecen a unas longitudes de onda menores. Por otro lado, las transiciones con $\Delta J = -1$ se conocen como rama P y aparecen con una estructura casi idéntica a la rama R a longitudes de onda mayores (Figura I - 14).

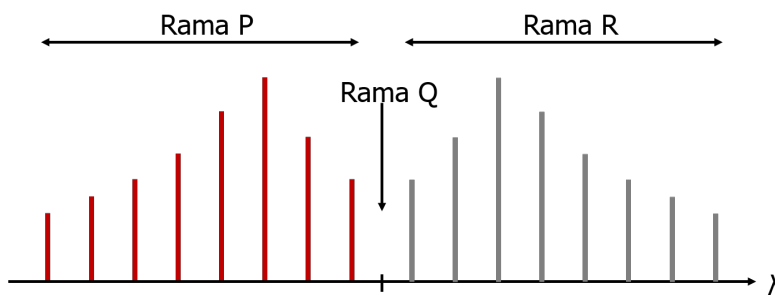


Figura I - 14. Espectro rotacional-vibracional propio de los gases donde se pueden observar las tres ramas comentadas P, Q y R.

La selección de la longitud de onda es un paso crítico para el desarrollo de la técnica, debido a que la banda de absorción del analito debe poder aislarse de otras bandas de la misma especie o de especies interferentes (principalmente gases presentes con una alta concentración en la atmósfera, como el vapor de agua o el CO_2).^{122,127} Por consiguiente, el láser precisa disponer de una longitud de onda que coincida con una banda de absorción rotacional-vibracional intensa y característica del gas que se quiera estudiar. Actualmente, existen distintas bases de datos de líneas espectrales para facilitar esta selección.¹²⁸

La espectroscopia óptica para gases, TDLAS, se basa en la ley de Lambert-Beer, por lo tanto, se mide la atenuación de la luz (de la fuente de emisión) a medida que se propaga a través de un medio absorbente.¹²⁹⁻¹³¹ Cuando la longitud de onda del láser coincide con la banda de absorción característica del gas estudiado, éste absorbe la luz transmitida a través del camino óptico, de acuerdo con la relación conocida de las medidas por transmitancia:^{132,133}

$$I_{\lambda} = I_{\lambda,0} \cdot e^{-S(T)g(\lambda,T,P)Nl} = I_{\lambda,0} \cdot e^{-\alpha l} \quad \text{Ecuación I - 3}$$

Donde:

- I_{λ} es la intensidad transmitida, que depende de la longitud de onda del láser λ .
- $I_{\lambda,0}$ es la intensidad inicial del láser.
- $S(T)$ es la intensidad de la línea de absorción dependiente de la temperatura.
- $g(\lambda, T, P)$ es la forma de la línea de absorción que depende de la longitud de onda del láser, de la temperatura y de la presión.
- N es la densidad de las moléculas absorbentes del gas.
- l es la longitud del camino óptico.
- α es el coeficiente de absorción.

Como se ha mencionado anteriormente, en TDLAS se trabaja con bandas de absorción muy estrechas y nítidas, por lo que el ruido asociado a las fluctuaciones de potencia del láser o al detector, puede afectar severamente a la señal y, por lo tanto, a la sensibilidad de la técnica. Los métodos más comunes empleados para mejorar el rendimiento del sistema TDLAS se pueden clasificar en dos grupos, aquellos que se basan en mejorar la intensidad de la señal y los que se basan en la reducción del ruido. En la siguiente sección se comentará con más detalle uno de los métodos utilizados con este último fin.^{115,116,122}

4.2.1. Espectroscopia de modulación de frecuencia

La aplicación de la FMS en TDLAS, se ve favorecida por la facilidad con la que los láseres de diodo pueden ser modulados y a las bandas estrechas de los gases.¹²⁷

La base de la técnica FMS consiste en modular la frecuencia del láser emisor, superponiéndola a una onda sinusoidal con una frecuencia del orden de varios MHz o GHz.¹²⁴ El concepto es análogo a la forma en que se codifica la información de voz en las bandas laterales de una transmisión de radio.

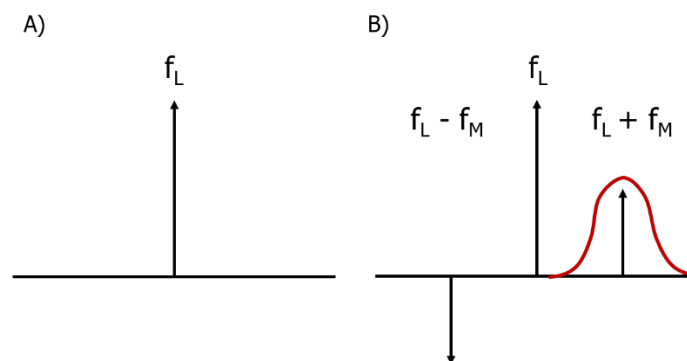


Figura I - 15. En la figura A se muestra la luz emitida por el láser (f_L) en transformada de Fourier antes de ser modulada, en la figura B se muestra la señal modulada y cómo la banda lateral superior ($f_L + f_M$) se encuentra centrada en la banda de absorción del gas.

En el esquema de la Figura I - 15A se muestra el dominio de frecuencia de la transformada de Fourier de la luz emitida por el láser antes de ser modulada, f_L , y en la Figura I - 15B una vez ya ha sido modulada. La frecuencia modulada contiene la frecuencia portadora de la información del láser, f_L , junto con unas bandas laterales que consisten en $f_L \pm f_M$, donde f_M es la frecuencia de modulación. Durante esta modulación la banda lateral superior queda centrada en el máximo de absorción del gas estudiado, mientras que la inferior queda fuera de la banda de absorción (Figura I - 15B).

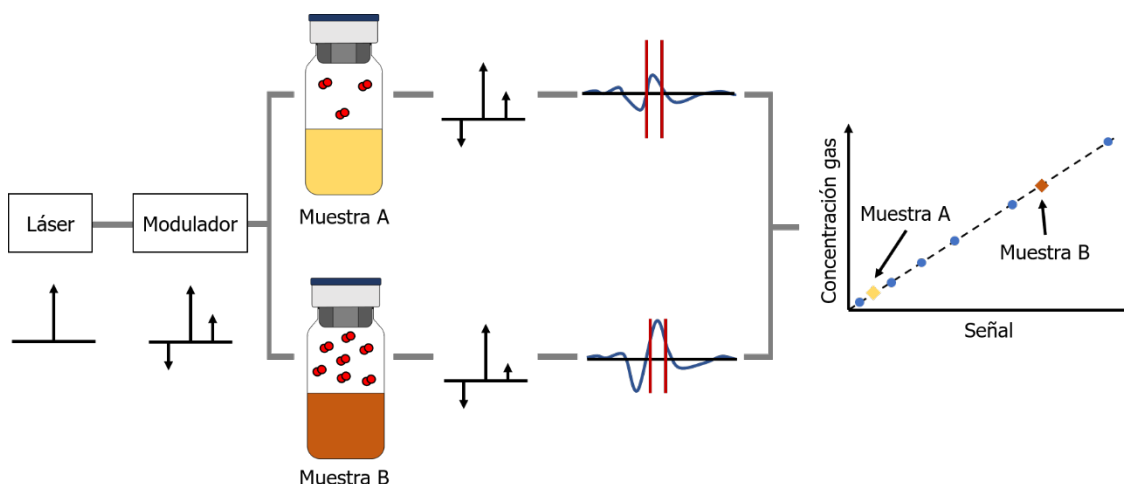


Figura I - 16. Esquema del proceso de modulación aplicado a un análisis. En éste se puede observar cómo varía la señal modulada dependiendo de la concentración del gas y cómo se interpola en una recta de calibración para saber la concentración.

Cuando esta luz modulada pasa a través de la muestra, el gas absorbe una parte de la luz, modificando así la intensidad de la banda lateral superior (Ecuación I - 3).^{130,134} Este proceso se muestra de una forma esquemática en la Figura I - 16. La diferencia entre estas bandas laterales, una dentro de la banda de absorción del gas y la otra fuera, aporta la información sobre la concentración del gas. Para poder obtenerla, primero se debe demodular la señal utilizando un demodulador (Figura I - 17) y, seguidamente, interpolar el resultado en una recta de calibración previamente realizada empleando diferentes patrones de concentración conocidas.¹³⁴

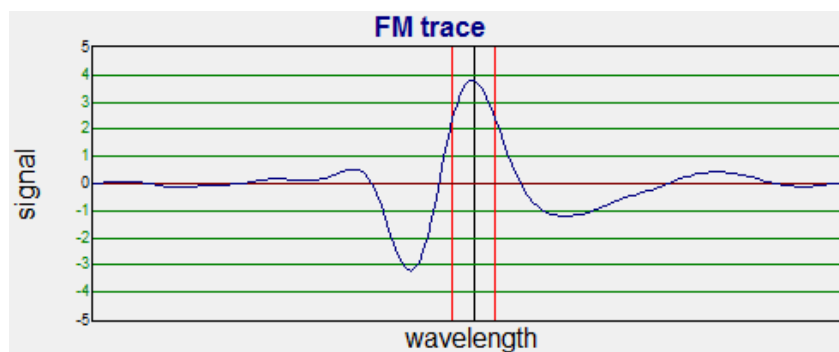


Figura I - 17. Señal demodulada obtenida de la medida del estándar de oxígeno de concentración 25 % (línea azul). La línea gris representa donde debería encontrarse el máximo obtenido y, las líneas rojas el error que se permite, el cual depende del formato del vial.

4.3. Instrumentación

Los espectrómetros láser permiten diseños menos complejos que la instrumentación analítica habitual empleada para la cuantificación de gases (como la cromatografía de gases).¹¹⁸ Las partes más importantes de cualquier configuración TDLAS-FMS son

principalmente, un láser de diodo, un modulador de la frecuencia, un compartimiento de la muestra, un detector y un demodulador.¹²³ En la siguiente figura se muestra un esquema de la instrumentación (Figura I - 18).

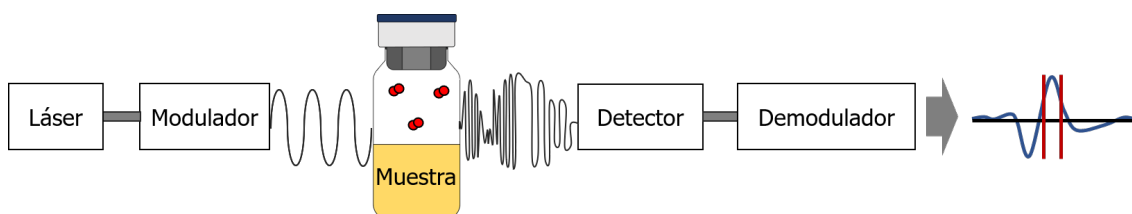


Figura I - 18. Esquema de la instrumentación básica de un TDLAS-FMS donde se pueden observar las distintas partes que lo conforman.

Dado el amplio abanico de aplicaciones de esta técnica (comentadas en la siguiente sección), los equipos pueden presentar distintas configuraciones dependiendo de si se utilizarán en un laboratorio o si estarán implementados en el proceso de fabricación del producto.



Figura I - 19. En esta imagen se muestra a la izquierda un TDLAS para detectar agua en un liofilizador y, a la derecha un equipo para detectar el mismo gas en el espacio de cabeza de viales liofilizados.

Además de la instrumentación comentada, este tipo de técnica requiere de unos patrones de concentración certificada del gas en estudio, con el fin de calibrar el equipo para su uso. Estos estándares deben estar en el mismo recipiente que las muestras a estudiar, para así tener en cuenta la interacción del láser con el formato del envase.^{127,134,135}

4.4. Aplicaciones

Dependiendo del láser empleado como fuente de radiación, la técnica TDLAS puede emplearse para cuantificar distintos gases (vapor de agua, CO₂, CO, NH₃, óxidos de nitrógeno, hidrocarburos, ...), por lo que las aplicaciones que podemos encontrar comprenden un sinfín de posibilidades.^{124,125} Además de cuantificar la concentración de determinados gases, la técnica también puede ser empleada para obtener otros parámetros como temperaturas, flujos o velocidades.¹³⁶

En el sector de los productos liofilizados, TDLAS se utiliza, principalmente, para determinar la concentración de los siguientes gases:

- Vapor de agua: este gas, al igual que ocurre en su estado líquido o sólido, presenta unas bandas de absorción de alta intensidad en el infrarrojo (935 nm y 1393 nm en el caso del vapor).¹³⁷ Esta característica, juntamente con el hecho de ser una de las sustancias más abundantes en la atmósfera, lo convierten en una especie interferente que puede dificultar la cuantificación de otros gases.¹¹⁵ En el caso de los liofilizados, al tratarse de uno de los solventes más empleados en el sector, convierte la cuantificación del vapor de agua en un requisito indispensable, tanto durante el proceso de liofilización (monitorización del ciclo de liofilización) como después (determinar la humedad en el producto liofilizado).^{24,125,138}
- CO₂: el dióxido de carbono, al igual que el vapor de agua, presenta una banda de absorción muy intensa en el infrarrojo (1604 nm para el estado gaseoso) y, como es muy abundante, se considera una especie interferente.¹¹⁵ Por lo que se refiere a su interés en la industria farmacéutica de productos liofilizados, la cuantificación del CO₂ se utiliza principalmente para modelar el crecimiento de microorganismos durante los llenados asépticos.^{117,125,139}
- O₂: en la industria de productos farmacéuticos liofilizados, la cuantificación del oxígeno se puede emplear para modelar (juntamente con el CO₂) el crecimiento de microorganismos durante los llenados asépticos, así como para determinar la estanqueidad de las muestras (CCI, siglas del inglés *Container Closure Integrity*) o confirmación del vacío.^{117,140} La determinación de oxígeno se describe con mayor profundidad en la siguiente sección.

4.4.1. Base teórica de la determinación del oxígeno

Según las reglas de selección espectroscópicas, para poder observar una absorción en el infrarrojo en el caso de una molécula diatómica heteronuclear, se tiene que observar un cambio simultáneo en ΔJ y Δv . Además, dicha molécula debe mostrar un cambio en el momento dipolar eléctrico. Esto quiere decir, que la transición del estado fundamental con $\Delta v = \pm 1$ y $\Delta J = 0$ no está permitida, por lo tanto, no se dispone de la rama-Q.

El caso del oxígeno molecular, el cual consiste en una molécula diatómica homonuclear, ya no muestra un cambio en el momento dipolar eléctrico, por lo que no esperaríamos ninguna actividad en la zona del infrarrojo. Sin embargo, las bandas de absorción del

oxígeno de las transiciones visibles/infrarrojas se han estado estudiando desde el 1927.^{141,142}

La explicación de que el oxígeno sea activo en la región infrarroja se debe a su carácter paramagnético y su configuración electrónica (Figura I – 20). En su estado fundamental la molécula de oxígeno consiste en un triplete ($X^3\Sigma_g^-$) con dos electrones desapareados en los orbitales π_x^* y π_y^* (Figura I – 20A). Sin embargo, también existen dos estados singletes los cuales son $a^1\Delta$ y $b^1\Sigma_g^+$ (se puede observar su configuración electrónica en la Figura I – 20B y C, respectivamente), con una energía sobre el estado fundamental de 94,7 kJ (1269 nm) y 158 kJ (762 nm).¹⁴³

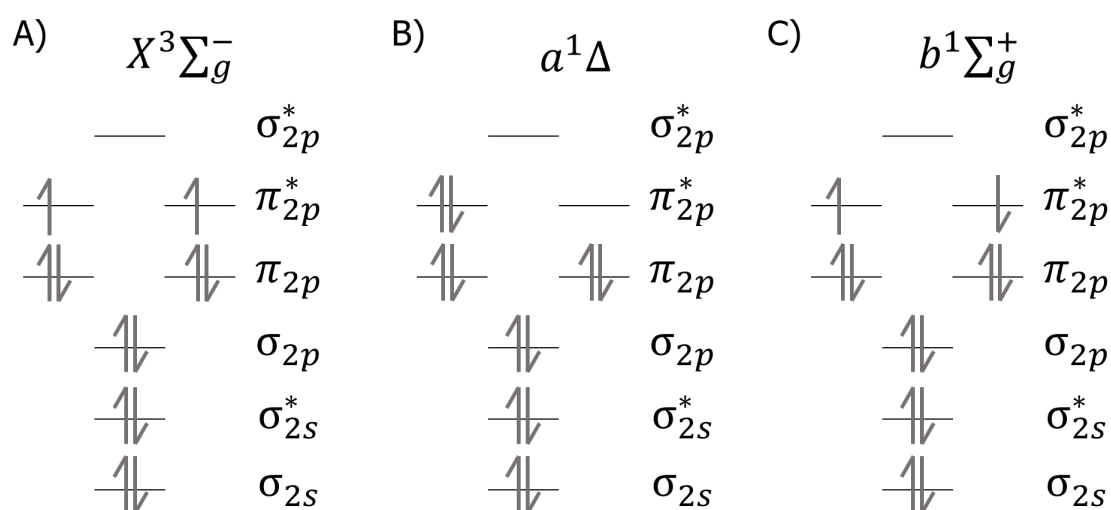


Figura I – 20. Configuración electrónica del oxígeno. En la figura A se puede observar el estado fundamental formado por un triplete con dos electrones desapareados. En la figura B se puede observar el singlete con menor energía $a^1\Delta$ donde ambos electrones están el mismo orbital molecular. En la figura C se puede observar el singlete $b^1\Sigma_g^+$ en el cual ambos electrones se encuentran en distintos orbitales moleculares, pero con el espín inverso.

Entonces el cambio en el espín de uno de los electrones entre los estados triplete y singlete ($b^1\Sigma_g^+ \leftarrow X^3\Sigma_g^-$) es una transición permitida por el cambio del dipolo magnético.

Cuando se acopla el momento rotacional angular (N) con el espín electrónico (S) se proporciona a la molécula un momento angular total (J , donde $J = N + S$). Esta interacción entre el momento rotacional angular con el espín electrónico se desdobra en tres subniveles, de forma que cada nivel rotacional se divide en tres estados con $J = N + 1$, $J = N$ y $J = N - 1$. Por consiguiente, las transiciones entre los términos consecutivos de los tres estados ($\Delta J = \pm 1$) son permitidas por las reglas de selección espectroscópicas mencionadas al inicio de esta sección.¹⁴¹

Consecuentemente, para cada número cuántico N del momento rotacional angular aparecen dos transiciones. Por lo tanto, la molécula de oxígeno presenta un espectro de

rotación-vibración. Las mediciones de oxígeno con la técnica TDLAS se llevan a cabo utilizando un láser centrado a 762 nm, aproximadamente. Esta longitud de onda aproximada se debe a que dependiendo del equipo empleado se puede medir la banda de mayor intensidad de la rama P centrada a 763,42 nm o la banda de la rama R centrada a 760,88 nm (ambas bandas están representadas en la Figura I - 21).^{136,137,144,145}

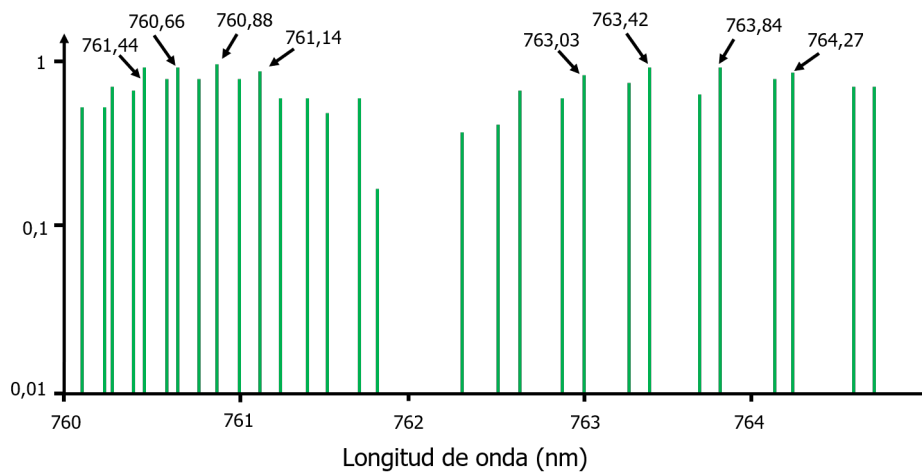


Figura I - 21. Ampliación en la parte de 760 nm a 765 nm del espectro del oxígeno donde se pueden observar las ramas P y R.

5. QUIMIOMETRÍA

La Ph. Eur. define la quimiometría como “la disciplina química que usa métodos matemáticos y estadísticos, (a) para diseñar o seleccionar procedimientos de medida y experimentos óptimos, (b) para proporcionar la máxima información química mediante el análisis de datos químicos”.¹¹³

La quimiometría utiliza datos de todo tipo de técnicas analíticas, siendo especialmente útil en aquéllas que presentan señales multivariantes como, por ejemplo, HPLC, NMR o técnicas espectroscópicas, entre ellas la espectroscopia NIR.¹⁴⁶ En esta última, tal y como se ha mencionado en el apartado 3, la quimiometría es crucial para la interpretación de los datos y ha permitido consolidar su aplicación en la industria.¹⁴⁷

En este capítulo, se ofrece una descripción de las herramientas quimiométricas empleadas en esta tesis para el desarrollo de modelos, tanto en análisis cualitativo como cuantitativo, de manera que cumplan los requisitos regulatorios aplicados a la industria farmacéutica.

5.1. Introducción

Como se desprende de la definición anterior, la quimiometría se entiende como la aplicación de la estadística y de la matemática a datos químicos. La aplicación de pruebas estadísticas clásicas a la química, como la t de Student (1908) o la F de Fischer (1925), son consideradas por algunos autores como la prehistoria de la quimiometría. A pesar de ello, no fue hasta los años 70 cuando debido a la complejidad de los datos multivariantes se empezó a requerir de métodos estadísticos y matemáticos más complejos para poder extraer la información química. Se considera el inicio formal de esta disciplina el año 1972 cuando el sueco Svante Wold introdujo el término quimiometría. Desde entonces, su rápida extensión ha llevado a la creación de organizaciones como la *International Chemometrics Society* (1974) y a la publicación de numerosos libros y revistas especializados en el tema, por ejemplo, *Chemometrics and Intelligent Laboratory Systems* (1986) y *Journal of Chemometrics* (1987).¹⁴⁸⁻¹⁵⁰

En las últimas décadas, el interés en la quimiometría ha aumentado significativamente, debido a la necesidad de enfrentarse a problemas analíticos cada vez más complejos. Estos requieren de técnicas analíticas que proporcionen cada vez más información, lo que implica datos multivariantes y herramientas para extraerla. El rápido desarrollo de

hardware más potente, que permite una mayor rapidez y volumen de cálculos, ha permitido a su vez, que las técnicas quimiométricas estén instauradas como software prácticamente de rutina para la resolución de problemas en numerosos sectores industriales, entre ellos el farmacéutico.^{94,146}

5.2. Análisis cualitativo

El análisis cualitativo tiene como objetivo identificar o clasificar las muestras dependiendo de su composición o de una propiedad específica.¹⁵¹

Estos tipos de análisis, en el ámbito multivariante, se conocen como métodos de reconocimiento de pautas (PRM, siglas del inglés *Pattern Recognition Methods*). Los PRM fueron los primeros métodos quimiométricos que se publicaron al aparecer esta disciplina.¹⁴⁶

Dependiendo de la estrategia utilizada, los análisis cualitativos multivariantes se pueden dividir en dos grupos: no supervisados y supervisados (Figura I - 22).

- No supervisados: tienen como objetivo detectar similitudes entre muestras de las que, previamente, se dispone de muy poca o ninguna información.^{146,152} Este tipo de métodos se acostumbran a emplear en las etapas iniciales de una investigación y sirven, por ejemplo, para saber si las muestras estudiadas se pueden agrupar como pertenecientes a la misma clase o no. Dentro de este grupo están los análisis de agrupamientos (del inglés *Cluster Analysis*).¹⁵³
- Supervisados: con estos métodos PRM se parte de unas muestras de las que ya se conoce a qué clases pertenecen, por lo que el objetivo consiste en establecer reglas matemáticas para definir fronteras que separen las diferentes clases o modelos que las definan. Estas reglas se aplican a nuevas muestras para saber a cuál de las clases estudiadas se parecen más y poder clasificarlas (este tema se desarrolla con más detalle en la sección 5.4).^{105,153}

En las siguientes secciones se describirán los principales métodos quimiométricos utilizados durante el desarrollo de esta tesis.

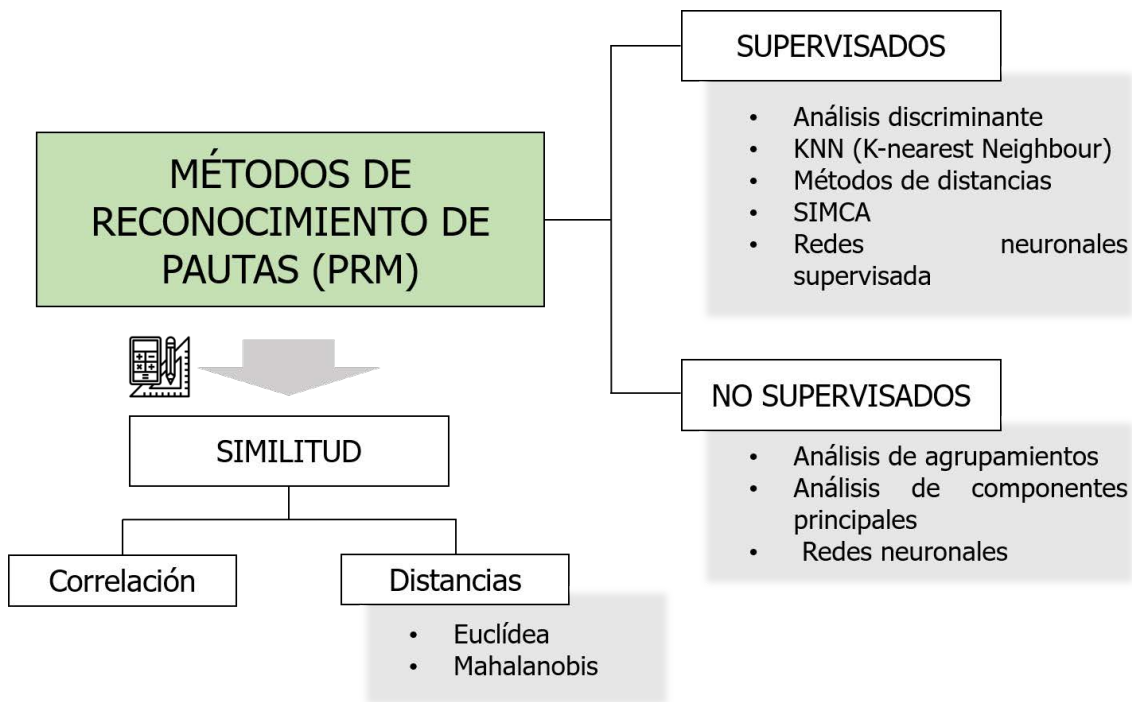


Figura I - 22. Esquema de la clasificación de los diferentes métodos de reconocimiento de pautas del análisis multivariante cualitativo.¹⁰⁵

5.2.1. Similitud

La mayoría de los PRM dependen de la medida de similitud, es decir, del grado de parecido que existe entre dos muestras. La similitud se pueden expresar usando el coeficiente de correlación o la distancia entre las dos muestras estudiadas.¹⁰⁵

5.2.1.1. Coeficiente de Correlación

En general, dos muestras representadas en un sistema de coordenadas son más parecidas entre ellas cuanto menor es el ángulo entre los dos vectores que unen el origen de coordenadas con cada una de las muestras (Figura I - 23). El coeficiente de correlación (CC) del ángulo, es el índice numérico que proporciona la similitud entre las muestras.¹⁵⁴ Se calcula siguiendo la Ecuación I - 4:

$$r_{x,y} = \frac{\sum_i x_i y_i}{\sqrt{\sum_i x_i^2 \sum_i y_i^2}} \quad \text{Ecuación I - 4}$$

Donde $r_{x,y}$ es el coeficiente de correlación entre los vectores x e y . En el caso de datos espectroscópicos, por ejemplo, estos vectores son los espectros de cada una de las muestras, con una dimensión igual al número de longitudes de onda del espectro.

Cuando dos espectros son completamente diferentes, el valor de r es 0, mientras que cuanto más similares sean las muestras más próximo a 1 será el valor de r . Pueden existir CC negativos, cuando las muestras que se parecen entre sí están en cuadrantes opuestos del sistema de coordenadas. A pesar de ello, es habitual utilizar el valor absoluto del CC, el cual puede tomar valores entre 0 y 1.¹²¹

Este índice se considera la herramienta estadística más simple para determinar el grado de semejanza entre muestras, y tiene la ventaja de ser independiente de la concentración de cada muestra.¹¹³

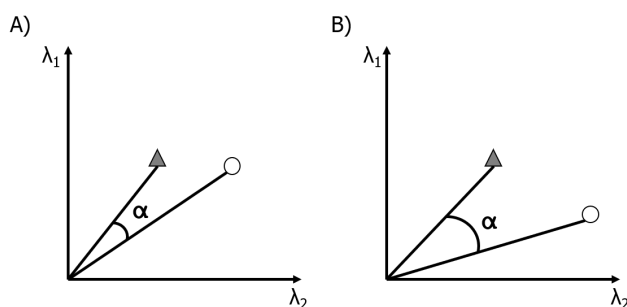


Figura I - 23. En la figura A se puede ver dos muestras similares, con un valor pequeño del ángulo entre ellas y r elevado y en la B, dos muestras más diferentes, con un ángulo mayor entre ellas y r menor.

5.2.1.2. Distancia Euclídea y de Mahalanobis

Además del coeficiente de correlación, la similitud también se puede cuantificar utilizando el concepto de distancia entre las muestras estudiadas.

Cuando las muestras están representadas en un sistema de coordenadas, ya sea el de las variables originales o el de cualquier transformación que se haya realizado, las muestras que se encuentran más próximas entre sí en el espacio se parecerán más entre ellas que las que están más alejadas. La distancia, realmente mide lo que se diferencian las muestras, lo que se podría llamar disimilitud. Así, cuanto menor sea la distancia entre las muestras, mayor será la similitud entre ellas.¹⁵²

Existen muchos tipos de distancias y se debe escoger una u otra dependiendo de las características del problema a estudiar. Las principales distancias utilizadas al aplicar PRM:

- Distancia Euclídea: esta distancia es la más habitual en la vida cotidiana y una de las más empleadas en quimiometría cuando se utilizan variables continuas. Para determinar la distancia entre dos muestras k y l se utiliza la Ecuación I - 5 (expresada en notación escalar), mientras que si se desea determinar la distancia

entre k y el centroide de una clase c , la ecuación se modifica a la Ecuación I - 6 (expresada en notación matricial).

$$d_{k,l} = \sqrt{\sum_{i=1}^m (x_{k,i} - x_{l,i})^2} \quad \text{Ecuación I - 5}$$

Donde $d_{k,l}$ es la distancia Euclídea entre las muestras k y l ; y m es el número de dimensiones del espacio.

$$d_{k,c} = \sqrt{(x_k - \bar{x})(x_k - \bar{x})'} \quad \text{Ecuación I - 6}$$

Donde $d_{k,c}$ es la distancia Euclídea entre la muestra k y el centroide de una clase c , \bar{x} es la media de las coordenadas del centroide y el superíndice ' indica la matriz transpuesta.^{113,146} A pesar de ser la distancia más utilizada, la distancia Euclídea proporciona valores erróneos cuando existe una correlación entre las variables utilizadas, debido a que su valor varía si se mide en la dirección de la correlación o no.

- Distancia de Mahalanobis: este tipo de distancia se utiliza en lugar de la distancia Euclídea cuando se debe tener en cuenta la correlación entre las variables. Por lo que la distancia entre una muestra y el centroide de una clase viene dada por la Ecuación I - 7.

$$md_{k,c} = \sqrt{(x_i - \bar{x})C_x^{-1}(x_i - \bar{x})'} \quad \text{Ecuación I - 7}$$

Donde $md_{k,c}$ es la distancia de Mahalanobis entre la muestra k y el centroide de la clase c , y C_x^{-1} es la matriz la varianza-covarianza, simétrica respecto a la diagonal principal (cuyos elementos son todos igual a 1, ya que contienen la varianza de cada variable consigo misma), mientras que los otros términos de la matriz contienen la covarianza entre las variables respectivas.¹⁴⁶

5.2.2. Análisis de componentes principales

El análisis en las componentes principales (PCA, siglas del inglés *Principal Component Analysis*) es una técnica utilizada para la reducción de variables sin una pérdida de información. Esta técnica se emplea de manera generalizada en el tratamiento de datos de múltiples disciplinas, desde economía a sociología.

En química, PCA empezó a tomar importancia en los años 70 con la aparición de la quimiometría. En esta disciplina, PCA es una herramienta indispensable para poder analizar datos multivariantes continuos en dos dimensiones (por ejemplo, espectros, cromatogramas de HPLC registrados a una longitud de onda del detector, etc).¹⁴⁶ Estos tipos de datos contienen un gran número de variables que están relacionadas entre sí y que contienen la misma información de la muestra expresada desde distintos puntos de vista. Por lo tanto, PCA es una técnica que permite eliminar la información redundante o irrelevante.^{92,155}

PCA transforma las variables originales en unas nuevas variables, no correlacionadas entre sí (ortogonales), y conocidas como componentes principales (PC, siglas del inglés *Principal Components*). La mayor información sobre la variabilidad de los datos originales aparece en las primeras PC, y va disminuyendo paulatinamente hasta que las últimas PC únicamente describen la variabilidad asociada al error experimental o al ruido, y pueden ser descartadas. Así, el número original de variables queda reducido manteniendo toda la información original en unas pocas PC.

Previamente a realizar la reducción de variables, los datos acostumbran a estar centrados y autoescalados (estos pretratamientos de los datos se describen en la sección 5.4.4). Desde el punto de vista matemático, el PCA consiste en una transformación de las variables originales a un nuevo sistema ortogonal de coordenadas, las componentes principales, y que viene descrito por la Ecuación I - 8.

$$X = TP' + E \qquad \text{Ecuación I - 8}$$

Donde:¹⁵²

- X es la matriz de los datos originales, de dimensiones $(n \times m)$. Cada una de las n filas corresponde a un objeto (las muestras), y cada una de las m columnas contienen una de las variables. En esta tesis serán los espectros NIR de las muestras.
- T es la matriz de los *scores*, de dimensiones $(n \times a)$, siendo a el número de PC necesarios para contener la información deseada. Los *scores* contienen la información de las muestras.
- P es la matriz de los *loadings*, de dimensiones $(m \times a)$, conteniendo la información sobre las variables de la matriz original. Matemáticamente son los cosenos directores entre las nuevas variables (PC) y las variables originales.

- E es la matriz de los residuales, la cual contiene la información de los datos originales que no ha sido explicada con las a componentes principales elegidas en el modelo.

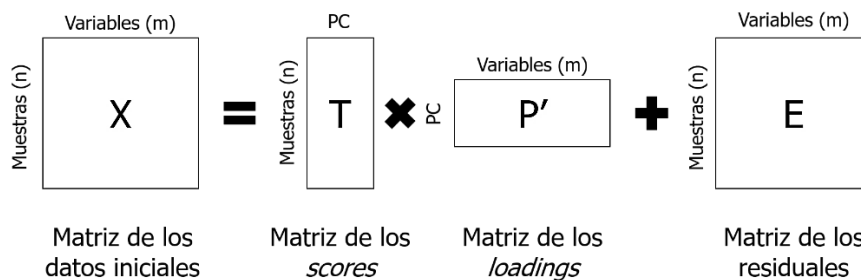


Figura I - 24. Representación de las matrices de la Ecuación I - 8 utilizadas para realizar el PCA.

Existen distintos algoritmos para realizar la descomposición de la matriz original. Uno de los más empleados en el campo de la quimiometría es el algoritmo de mínimos cuadrados parciales iterativos no lineales (NIPALS, siglas del inglés *Nonlinear Iterative Partial Least Squares*). Este algoritmo cada vez que calcula una PC elimina la información explicada por ésta de la matriz original y de esta nueva matriz, con la información sustraída, calcula la siguiente PC, hasta alcanzar a explicar el 100 % de la información contenida en la matriz original.^{92,152}

5.2.3. Análisis discriminante por mínimos cuadrados parciales

El análisis discriminante por mínimos cuadrados (PLS-DA, siglas del inglés *Partial Least Squares-Discriminant Analysis*) es un método supervisado de PRM basado en la reducción de variables. En el PLS-DA se descomponen las matrices X e Y simultáneamente siguiendo el mismo procedimiento que en la regresión por mínimos cuadrados parciales, explicada con más detalle en la sección 5.3.2.1.

En el caso de una clasificación en dos clases, habitualmente la matriz Y es una columna en la que se codifican las muestras con un "1" y un "0" según la clase a la que pertenecen. Si existen más de dos clases, la matriz Y consta de tantas columnas como clases, codificando las muestras con "1" en la clase a la cual pertenecen y "0" en las otras. El resultado de la regresión son valores de Y cercanos "1" y a "0", indicando la pertenencia o no a la clase considerada.

Una muestra del conjunto de predicción se considera que está correctamente clasificada si se cumplen dos criterios: su espectro es similar al de las muestras que forman el conjunto de calibración y el valor calculado de la variable que indica la categoría está

dentro de un intervalo de valores alrededor del "1" también conocido como el límite de decisión (traducción del inglés *cut-off value*). Una forma habitual de marcar este valor límite entre clases es coger el valor medio entre ambas. Con la codificación indicada, todas las muestras cuyo valor calculado por el modelo sea $>0,5$ serían de la clase estudiada.^{156,157}

5.3. Análisis cuantitativo

El análisis cuantitativo tiene como objetivo ofrecer un resultado numérico para expresar una propiedad concreta de una muestra.^{151,158}

Para realizar este propósito, el análisis cuantitativo emplea la calibración, cuyo objetivo es desarrollar un modelo matemático que relacione la señal (o señales) obtenida por un instrumento con la propiedad estudiada. El modelo de calibración se construye tratando la señal instrumental, obtenida a partir de unas muestras de las que se conoce el valor de la propiedad estudiada. En una segunda etapa, para conocer el valor de esta propiedad de nuevas muestras, distintas de las utilizadas en el proceso de calibración, sólo hay que registrar su señal instrumental y aplicar el modelo matemático calculado. En la mayor parte de las ocasiones, el modelo de calibración utiliza ecuaciones lineales en los coeficientes, muchas veces una línea recta, aunque también hay casos en que es necesario utilizar modelos no lineales.¹⁵⁹

5.3.1. Calibración univariante

La calibración univariante relaciona una única variable de respuesta (y) con una única variable predictora (x) utilizando la ecuación de una recta (Ecuación I - 9).¹⁶⁰ La calibración consiste en obtener los coeficientes de la ecuación mediante un ajuste por mínimos cuadrados de las señales de muestras de las que se conoce el valor de la propiedad estudiada.

$$y_i = a + bx_i \qquad \text{Ecuación I - 9}$$

Desde un punto de vista estadístico, y_i son los valores de la propiedad estudiada, por ejemplo, la concentración del analito, y x_i son los valores de la variable seleccionada para cada muestra; por ejemplo, la absorbancia a una longitud de onda concreta. b es la pendiente y a la ordenada en el origen. Tradicionalmente, el uso de la calibración univariante en química ha sido al contrario que en estadística, utilizando como variable dependiente (y) el valor de la señal, mientras que el valor de la propiedad estudiada

(por ejemplo, la concentración) se utiliza como variable independiente (x). Esto ha llevado a cada una de las comunidades (químicos y matemáticos) a llamar "regresión inversa" a la que utiliza la ecuación de forma no habitual en su respectiva comunidad.

Dada la sencillez de este tipo de calibración y que la mayoría de instrumentos, tradicionalmente, han proporcionado una señal analítica única, no es de extrañar que la calibración univariante sea uno de los algoritmos de calibración más empleado.¹⁴⁶ Aunque, se ha utilizado durante el desarrollo de esta tesis, no se describen los detalles del método ya que son ampliamente conocidos.¹⁵¹

5.3.2. Calibración multivariante

No todos los análisis cuantitativos se pueden solventar estudiando la relación entre dos variables, sino que muchas veces es la combinación de distintas variables (X) lo que permite cuantificar una o varias propiedades (y)^{VI} de la muestra.¹⁵³ La calibración multivariante se diferencia de la calibración univariante en que la información procedente de la medida experimental contiene múltiples variables.¹⁶¹

Existen diferentes métodos matemáticos para poder realizar una calibración multivariante cuantitativa, aunque durante el desarrollo de esta tesis se ha empleado la calibración basada en la Regresión por Mínimos Cuadrados Parciales (PLSR, siglas del inglés *Partial Least Squares Regression*), que se explicará con más detalle en la siguiente sección. En la Tabla I - 4 se muestran algunos de los más utilizados.

Tabla I - 4. Listado y breve descripción de los principales métodos de calibración multivariantes cuantitativos.

Método	Descripción
Regresión lineal múltiple (MLR, siglas del inglés <i>Multiple Linear Regression</i>)	Requiere la selección de un número limitado de variables que tengan una correlación con la propiedad a determinar. Se construye una ecuación de regresión, que es lineal en los coeficientes, y se ajusta por mínimos cuadrados, con el fin de establecer una relación lineal entre ellas.
Regresión en las componentes principales (PCR, siglas del inglés <i>Principal Components Regression</i>)	Realiza una reducción de variables utilizando un análisis PCA de los datos X . En la regresión lineal se utilizan los <i>scores</i> calculados en el PCA en lugar de las variables originales.
Regresión por mínimos cuadrados parciales (PLSR, siglas del inglés <i>Partial Least Squares Regression</i>)	También es un método lineal que realiza una reducción de variables, pero la descomposición de los datos busca la máxima correlación entre las variables X y las variables Y a determinar.

^{VI} Si sólo se analiza una propiedad la variable se expresará como y mientras que si se analizan más de una propiedad se expresará como Y .

Método	Descripción
Redes neuronales artificiales (ANN, siglas del inglés <i>Artificial Neural Networks</i>)	Simula la transmisión de la información de las neuronas biológicas, y permite obtener modelos de calibración de sistemas no lineales.

5.3.2.1. Regresión por mínimos cuadrados parciales

La regresión por mínimos cuadrados parcial fue introducida como una alternativa a la regresión clásica por Herman Wold en 1975. A pesar de ello, no fue hasta 1982 cuando Martens y Jensen sugirieron implementar este método para el desarrollo de modelos multivariantes a partir de espectros NIR. Actualmente, PLSR es una de las técnicas más empleadas en la calibración multivariante NIR.¹¹¹

PLSR se puede clasificar como un método de calibración indirecto (los parámetros de calibración se obtienen de la señal obtenida de una mezcla de componentes), flexible (los términos de la ecuación de calibración se deben optimizar) y, desde el punto de vista químico, inverso (la propiedad determinada es la variable dependiente y la señal analítica obtenida es la variable independiente).¹⁶²

Como en PCA, PLSR realiza una reducción de variables, pero a diferencia del primero, éste descompone la matriz X teniendo en cuenta la información de la matriz Y . Por tanto, los factores obtenidos combinan la máxima varianza de X manteniendo una máxima correlación con la propiedad Y a determinar.¹⁵²

Al igual que en PCA, los datos por tratar suelen estar centrados o autoescalados (sección 5.4.4). Seguidamente, se realiza de forma simultánea la descomposición de ambas matrices:

$$X = TP' + E \quad \text{Ecuación I - 10}$$

$$Y = UQ' + F \quad \text{Ecuación I - 11}$$

Donde T y U son las matrices de los *scores*, P' y Q' las matrices de los *loadings*, y E y F las matrices de los residuales de los datos iniciales de X (matriz de los datos espectrales) e Y (matriz de la propiedad a determinar), respectivamente.¹⁶² Uno de los distintos algoritmos que existen para realizar la descomposición de las matrices originales, es el ya comentado NIPALS.

A partir de esta descomposición se obtienen los coeficientes de regresión de la recta de calibración (Ecuación I - 12) que permiten predecir la propiedad determinada al introducir nuevos espectros (Ecuación I - 13).¹⁶²

$$\hat{B} = W(P'W)^{-1}Q' \quad \text{Ecuación I - 12}$$

$$\hat{y}_i = x_i'\hat{B} \quad \text{Ecuación I - 13}$$

Donde \hat{B} son los coeficientes de regresión de la recta de calibración, W la matriz de pesos de los *loadings*, que representa la relación entre X e Y y, x_i e y_i los datos referentes a la nueva muestra.

5.4. Etapas del proceso de modelado

El proceso de modelado de un conjunto de datos ya sea para desarrollar un modelo de análisis cualitativo o uno de análisis cuantitativo, consiste en diversas etapas que finalizan con la aplicación del modelo a su objetivo final. Estas etapas sirven para asegurar que el modelo obtenido es robusto y capaz de predecir correctamente la propiedad de interés de la muestra.^{92,114,163}

En la Figura I - 25 se muestra un esquema del proceso de modelado. En las siguientes secciones se describen con más detalle cada una de estas etapas, incluyendo las peculiaridades de su uso en la industria farmacéutica.

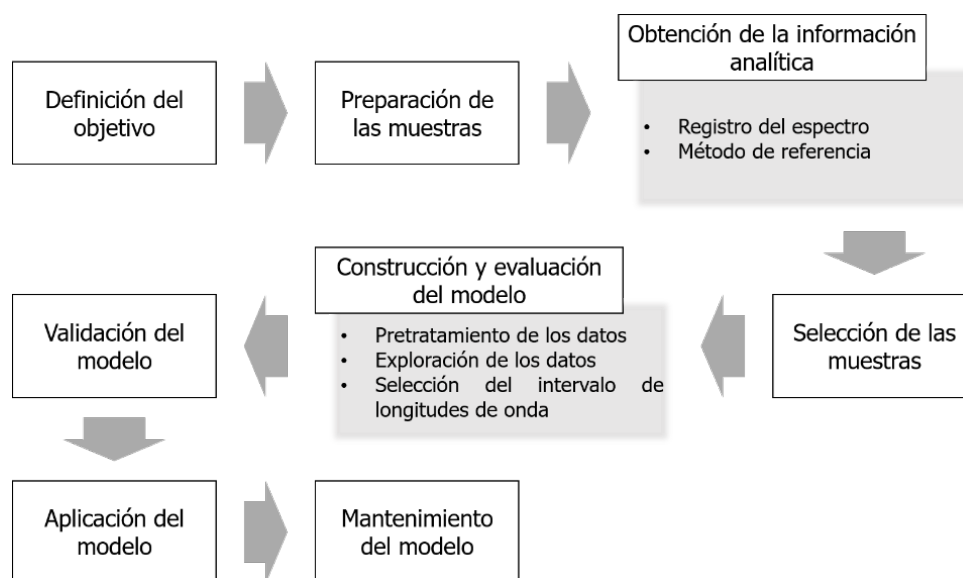


Figura I - 25. Etapas del proceso de construcción, validación y uso de un modelo quimiométrico.

5.4.1. Definición del objetivo

Debido a que se trabaja en un entorno altamente regulado, previamente al desarrollo del modelo multivariante se debe estudiar la viabilidad del mismo para poder definir y documentar cuál será su objetivo y el alcance de éste.¹¹⁴ Es decir, se debe indicar a qué producto(s) va a aplicarse, cómo se van a realizar las medidas (*on/in/at line*), si el modelo va a emplearse en rutina sustituyendo un método analítico ya existente, o si la información obtenida se empleará sólo para ampliar el conocimiento del producto o del proceso. Una vez se ha definido todo esto, se puede proceder a su desarrollo.

5.4.2. Preparación de las muestras

El primer paso para desarrollar un modelo quimiométrico es disponer de un conjunto de muestras que sean representativas del objetivo del estudio, es decir, que incluyan todas las posibles fuentes de variabilidad del proceso (distintos lotes de API y excipiente, diferentes equipos empleados en el proceso de fabricación, etc.) que puedan afectar a la señal. Esto es particularmente importante en el caso de la espectroscopia NIR, debido a que la señal depende tanto de las variaciones químicas de la muestra como de las físicas de ésta (tamaño de partícula, grado de compactación, temperatura, etc.). Esta inclusión permite no sólo aumentar la capacidad predictiva del modelo, sino también su capacidad de detección de muestras anómalas.^{92,153}

El procedimiento habitual es disponer de muestras de distintos lotes de fabricación para así poder analizarlas. Estas muestras contienen la información intrínseca del proceso y cuentan con la variabilidad natural del mismo. Cuanto mayor sea el número de lotes incluidos en el modelo más robusto será.

En el caso de desarrollar un modelo para un análisis cuantitativo, estas muestras, además, deben estar distribuidas de forma homogénea en el rango de concentraciones del analito que se quiere estudiar. El intervalo de concentraciones debe ser lo suficiente amplio como para que contenga el de las futuras muestras que se calculen con el modelo.⁸⁷

Dado que los procesos de fabricación están diseñados para obtener muestras con muy poca variación respecto al valor nominal del analito, cuando es necesario obtener un modelo cuantitativo es difícil que se pueda llegar a cubrir el intervalo de concentraciones deseado empleando solamente muestras procedentes de producción. Por consiguiente, para poder extender el intervalo de concentraciones, además de las muestras de

producción, es necesario preparar un conjunto de muestras de forma distinta a la habitual en la fabricación. Para poder alcanzar este fin manteniendo una similitud en el espectro entre las muestras preparadas y las muestras de fabricación, existen distintas estrategias:⁹²

- Planta piloto: utilizando una planta piloto se mimetiza el proceso de fabricación; de este modo se pueden obtener muestras con similar estructura física, pero a pequeña escala, lo que permite preparar lotes variando la concentración del analito.
- Muestras de producción modificadas: esta estrategia consiste en alterar la concentración del analito en muestras de producción con el fin de cubrir el rango de concentración estudiado. Por lo tanto, para cubrir el rango superior al valor nominal de la concentración, se introduce en la muestra cantidades conocidas del analito (muestras sobredosificadas, traducción del inglés *overdosed*), y para cubrir el rango inferior, se añade una mezcla de los excipientes presentes en la muestra (muestras subdosificadas, traducción del inglés *underdosed*).
- Muestras sintéticas: esta estrategia consiste en preparar muestras por todo el rango de concentraciones estudiado mezclando distintas cantidades conocidas de los componentes presentes en la muestra.

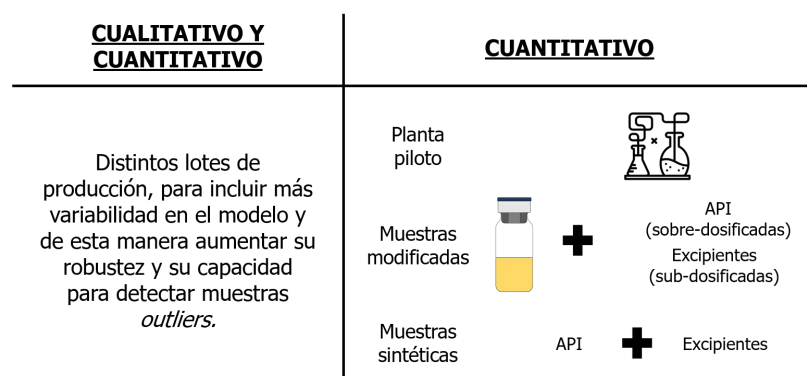


Figura I - 26. Esquema de la preparación de las muestras dependiendo del tipo de análisis cualitativo o cuantitativo.

5.4.3. Obtención de la información analítica

Cuando ya se dispone de las muestras, si se desea realizar un modelo de análisis cuantitativo, la siguiente etapa consiste en adquirir su información analítica.

Como se ha mencionado anteriormente, el proceso de calibrado necesario para la aplicación de un método cuantitativo requiere hallar la concentración de las muestras utilizando un método de referencia. Estos métodos de referencia acostumbran a estar

basados en técnicas destructivas, por lo que, obviamente, se debe registrar el espectro de la muestra previamente a realizar el análisis por el método de referencia.⁸⁷

El registro del espectro de las muestras que se van a usar para desarrollar el modelo multivariante se debe hacer siguiendo el mismo procedimiento normalizado de trabajo que el análisis de una muestra en rutina, es decir, utilizando el mismo equipo, la misma presentación de muestra, el mismo modo de registro, etc.⁹²

Es necesario que los métodos de referencia estén validados y que proporcionen resultados fiables, debido a que la calidad de éstos tiene un impacto directo en la calidad de los resultados de predicción del modelo multivariante.⁹²

5.4.4. Pretratamiento de los datos

Los datos instrumentales, no sólo contienen información sobre el analito o la característica de interés, sino que también contienen una parte de información que es irrelevante para el desarrollo del modelo.⁹⁴ En el caso de los espectros NIR, por ejemplo, sería la dispersión de la luz, interacciones entre compuestos, desplazamientos de la línea base, entre otros. Esta información puede interferir en el modelo multivariante haciéndolo menos robusto.^{74,92}

Con el fin de reducir o eliminar estas interferencias presentes en los datos, se acostumbra a realizar un pretratamiento de éstos previamente al desarrollo del modelo multivariante. El objetivo es simplificar el modelo y maximizar la relación señal/ruido reduciendo el ruido presente en los espectros.¹¹¹ Por lo tanto, el mejor pretratamiento (o combinación de éstos) no puede conocerse a priori y ha de ensayarse durante el desarrollo del modelo.

Existe una amplia variedad de pretratamientos; a continuación, se describen los utilizados durante el desarrollo de esta tesis, aplicados a medidas espectroscópicas.

- Centrado en la media: este pretratamiento se usa para que el origen del modelo esté en el centro de éste y se observe mejor el cambio en el valor de las variables debido a las variaciones de las muestras. Para ello se aplica la siguiente ecuación:

$$X_c = X - m_x \quad \text{Ecuación I - 14}$$

Donde X_c es la matriz con los datos finales centrados en la media, X es la matriz de los datos iniciales y m_x el vector que contiene los valores de la media de cada variable.⁹² Equivale a trabajar con la matriz de varianza-covarianza de los datos.

- Escalado a varianza unidad: este pretratamiento se aplica con el fin de igualar la influencia de las diferentes unidades de las variables. Al aplicarlo, cada una de ellas tendrá un promedio igual a 0 y una varianza de 1. Si todas las variables tienen las mismas unidades, la influencia de las variables con poca intensidad aumenta. Para realizar este pretratamiento, se aplica la siguiente ecuación:

$$X_E = \frac{X_C}{s} \quad \text{Ecuación I - 15}$$

Donde X_E es la matriz con los datos finales escalados, X_C es la matriz de los datos centrados en la media y s el vector que contiene los valores de las desviaciones estándar de cada variable.¹⁵⁵ Equivale a trabajar con la matriz de correlación de los datos.

- Derivadas: este pretratamiento está ampliamente extendido debido a que ayuda a resolver problemas de solapamiento de bandas y efectos de línea de base. Las derivadas acentúan la señal del analito, manteniendo la proporcionalidad de ésta con la concentración, aunque, a la vez, aumenta el ruido del espectro. Al aplicar una primera derivada a un espectro se elimina el desplazamiento vertical constante de la línea base; mientras que, al aplicar una segunda derivada, se elimina la variación lineal de la línea de base con la longitud de onda.^{91,161}

El método de derivación utilizado en esta tesis ha sido el de Savitzky-Golay, el cual realiza un suavizado del espectro juntamente con el cálculo de la derivada.¹⁶⁴

- Variable normal estándar (SNV del inglés *Standard Normal Variate*): este pretratamiento se emplea para minimizar las variaciones en los espectros debidas a cambios físicos en la muestra que influyen en la dispersión de la señal; por ejemplo, diferencias en el tamaño de las partículas.¹⁶⁵ El SNV se aplica a cada espectro individualmente y realiza un autoescalado de los valores de absorbancia. Para ello se aplica la siguiente ecuación:

$$x_{ik}^* = \frac{x_c}{s_i} \quad \text{Ecuación I - 16}$$

Donde, para cada muestra i y variable k , x_{ik}^* es el valor de la absorbancia resultante de aplicar el SNV y x_c el valor de los datos centrados en la media y s_i a la desviación estándar del mismo. Cada espectro resultante de aplicar un SNV tiene una media de 0 y una desviación estándar de 1.^{152,153}

- Corrección de la señal ortogonal (OSC del inglés *Orthogonal Signal Correction*): este pretratamiento desarrollado por S. Wold permite corregir la matriz de datos

original eliminando de esta toda la información de los espectros ortogonal con la propiedad que se desea determinar en el modelo multivariante.¹⁶⁶ El número de factores OSC corresponde a las veces que se someten los datos originales al pretratamiento. La forma que tiene de hacerlo es muy parecida a la del algoritmo NIPLAS aplicado a un PLSR o PCA, pero en este caso los factores OSC son más interpretables. Al igual que ocurre con las técnicas de reducción de variables, utilizar demasiados factores produce un sobreajuste del modelo (el sobreajuste se comenta con más detalle en la sección 5.4.8).

5.4.5. Selección del intervalo de longitudes de onda

Aunque, con el pretratamiento de los espectros se elimina mucha información irrelevante, la selección de un intervalo óptimo de longitudes de onda permite simplificar el modelo multivariante final.⁹²

El intervalo de longitudes de onda debe contener la información relacionada con la propiedad a determinar en el modelo. Si es demasiado amplio, se añade ruido innecesario al modelo, mientras que si es demasiado estrecho se reduce la posibilidad de detectar muestras anómalas durante la predicción.⁹⁴

5.4.6. Exploración de los datos

La exploración de los datos se debe hacer durante todo el proceso de desarrollo de un modelo quimiométrico, para poder detectar así muestras anómalas (en inglés *outliers*).

Una muestra *outlier* se define como una muestra que no sigue la misma tendencia que el resto de las muestras "normales"; es decir, que se diferencia mucho de las demás muestras estudiadas. Este tipo de muestras se pueden clasificar en dos grupos:

- *Y-outlier*: este tipo de muestras son debidas a un valor anómalo en la variable *Y*, es decir, de la medida del método de referencia.
- *X-outlier*: este tipo de muestras se describen como muestras que presentan un espectro, variable *X*, distinto a las demás muestras.

La incorporación de este tipo de muestras durante el modelado puede afectar en la complejidad del modelo y causar problemas en la robustez y en la capacidad predictiva de éste. Por lo tanto, la detección y eliminación de estas muestras debe realizarse durante la construcción del modelo.^{74,153}

Algunos métodos para detectar este tipo de muestras están basados en la visualización de los datos mediante la herramienta multivariante PCA. Esta técnica permite agrupar las muestras dependiendo de su naturaleza, por lo que si una muestra presenta una información distinta a las demás (*outlier*) puede detectarse tanto visualmente, a partir de la distancia. Como estadísticamente mediante el test T^2 de Hotelling, el cual acota un espacio que tiene una cierta probabilidad de contener todas las muestras similares, las muestras que quedan fuera se consideran estadísticamente distintas a las demás para el nivel de probabilidad elegido.^{98,167,168}

5.4.7. Selección de muestras

Con el fin de poder construir y posteriormente evaluar el modelo quimiométrico desarrollado, las muestras se deben dividir en los siguientes grupos:¹⁵²

- Conjunto de muestras de calibración/entrenamiento: en el caso del análisis cuantitativo se acostumbra a denominar a este conjunto como "muestras de calibración", mientras que en el PRM se denomina "muestras de entrenamiento". La finalidad de estas muestras es crear el modelo multivariante de calibración o de identificación, es decir, la ecuación matemática que relaciona la señal registrada con la propiedad estudiada o con la clase a la que pertenece la muestra. Este conjunto contiene la mayor cantidad de muestras (aproximadamente el 50 % del total de muestras disponibles) incluyendo así el máximo de variabilidad posible y cubriendo todo el intervalo de interés del analito en el caso del análisis cuantitativo.
- Conjunto de muestras de evaluación del modelo: el objetivo de este conjunto, formado aproximadamente por un mínimo del 25 % del total de las muestras, es optimizar y evaluar el modelo para obtener unos buenos resultados en la predicción o en la clasificación.
- Conjunto de muestras de predicción o validación: el propósito de este último conjunto de muestras es evaluar y validar el modelo final con unas muestras independientes, que no se han empleado durante la construcción del modelo para obtener una estimación realista de la capacidad predictiva de éste.

Se puede realizar la selección de qué muestras deben ir a cada conjunto utilizando algoritmos, como el de Kennard-Stone, basado en la distancia de Mahalanobis, o de forma aleatoria.¹⁶⁹ La idea de estas estrategias es que cada conjunto de muestras

contenga la máxima variabilidad y presente una distribución de las muestras homogénea en el intervalo de concentraciones estudiado.⁷⁴

Sin embargo, si se dispone de un número de muestras pequeño, sólo se utilizan dos grupos de muestras: calibración/entrenamiento (aproximadamente el 70 % del total de las muestras) y de predicción (aproximadamente el 30 % del total de las muestras). En este caso la evaluación y optimización del modelo se realiza empleando métodos de remuestreo (del inglés *resampling*) del conjunto de calibración/entrenamiento, como la validación cruzada (del inglés *cross validation*).

La validación cruzada permite evaluar y optimizar el modelo a partir de un único grupo de muestras; para ello separa un número determinado de muestras del conjunto de calibración y se construye un modelo sin tenerlas en cuenta.¹⁷⁰ Se evalúa la capacidad predictiva del modelo construido con las muestras separadas. Las muestras separadas se vuelven al conjunto de calibración y se reservan otras muestras, construyendo un nuevo modelo y volviendo a evaluar su capacidad predictiva con las muestras reservadas. Este proceso se realiza de forma iterativa hasta que todas las muestras han formado parte, en algún momento, del conjunto de calibración y del de evaluación, para encontrar el mejor modelo (Figura I - 27).¹⁵³

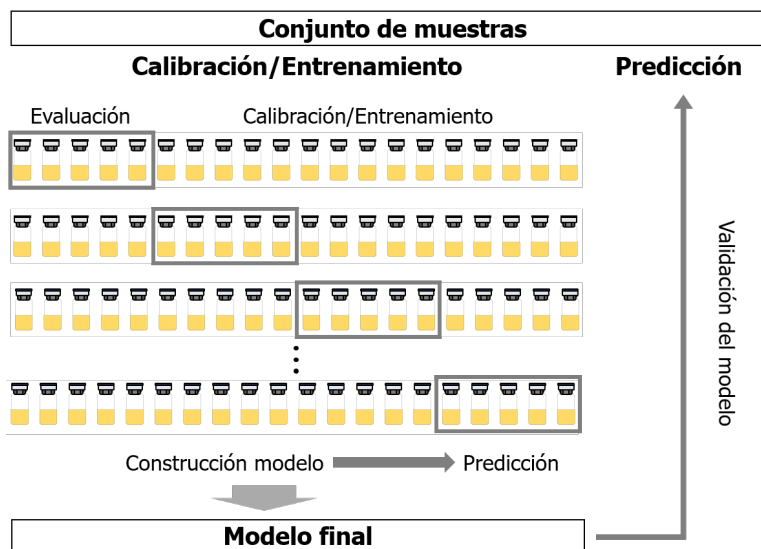


Figura I - 27. Representación gráfica del proceso de validación cruzada.

Se puede realizar la selección de qué muestras deben ir a cada conjunto utilizando algoritmos, como el de Kennard-Stone, basado en la distancia de Mahalanobis, o de forma aleatoria.¹⁶⁹ La idea de estas estrategias es que cada conjunto de muestras contenga la máxima variabilidad y presente una distribución de las muestras homogénea en el intervalo de concentraciones estudiado.⁷⁴

5.4.8. Construcción y evaluación del modelo quimiométrico

En el proceso de construcción del modelo se utilizan las muestras del conjunto de calibración para ir creando y ensayando distintos modelos con un número de variables latentes, α , que va desde 1 hasta A (número máximo de variables latentes ensayado). Si se está efectuando un PCA cada una de estas variables será una PC, mientras que si se realiza un PLSR cada una de ellas representa un factor PLSR. El modelo óptimo en el caso de un PCA será el que presente un número de variables latentes que permita reconstruir la matriz X original con el mínimo error. Mientras que cuando se trabaja con PLSR, como se realiza la descomposición simultánea de las dos matrices de calibración (X e Y), para conseguir el modelo óptimo, no sólo ha de ser posible explicar la matriz X , sino calcular los valores de Y de las nuevas muestras con el mínimo error.

La selección del número óptimo de variables latentes que se van a usar en el modelo final es crucial. No existe un método único que permita identificar cuál es el número óptimo de PC o factores PLSR a utilizar. En general, se aplican criterios estadísticos y, en el caso del PCA, una manera habitual consiste en evaluar la varianza explicada acumulada, seleccionar el número de PC que explica un porcentaje de varianza explicada que sea suficiente para satisfacer las necesidades, o el número de PC a partir del cual el incremento de la varianza explicada es insignificante.

Para un total de A componentes, la varianza explicada acumulada $V_{Exp.ac.}(\alpha)$ para cada componente, expresada en porcentaje se calcula según:

$$V_{Exp.ac.}(\alpha) = \frac{V_{Res}(0) - V_{Res}(\alpha)}{V_{Res}(0)} \quad \text{Ecuación I - 17}$$

Donde $V_{Res}(\alpha)$ es la varianza de la matriz X después de extraerá cada PC y $V_{Res}(0)$ la varianza de los datos originales.

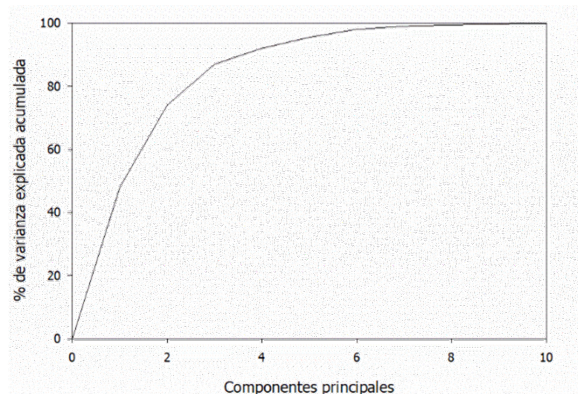


Figura I - 28. Gráfico de la varianza explicada acumulada en Y en función de las componentes principales.

En el caso del PLSR, el objetivo principal es la predicción de nuevas muestras, por lo que el número de factores PLS óptimo está relacionado con obtener bajos errores en las Y , aunque el modelo también debe de ser capaz de explicar la matriz X . La siguiente discusión está centrada en los valores de Y , aunque las ecuaciones son análogas si nos referimos a los valores de las X , sin embargo, en este caso los residuales serían la diferencia entre los valores experimentales (x_i) y los calculados por el modelo (\hat{x}_i).

Se puede evaluar el error de los valores Y de las muestras calibración utilizando el error cuadrático medio de la calibración (RMSEC, siglas del inglés *Root Mean Square Error of Calibration*).

$$RMSEC = \sqrt{\frac{\sum_{i=1}^M (\hat{y}_i - y_i)^2}{M - \alpha}} \quad \text{Ecuación I - 18}$$

Donde M es el número de muestras del conjunto de calibración, α el número de factores de cada modelo ensayado, \hat{y}_i es el valor calculado con el modelo de α factores para la muestra i , e y_i el valor hallado por el método de referencia para la misma muestra.

Cada uno de los modelos construidos se ha de evaluar para decidir cuál es el mejor y, en consecuencia, cuál es el número óptimo de factores PLSR a utilizar. Para ello se necesita calcular el error de predicción que se comete en el cálculo de las muestras. La manera de hacerlo es ligeramente distinta según el método de evaluación utilizado. Cuando se utiliza *cross-validation*, el error cuadrático medio de la predicción de los valores Y de la regresión se denomina RMSECV (siglas del inglés *Root Mean Square Error of Prediction by Cross-Validation*) y se calcula para cada variable latente siguiendo la Ecuación I - 19.¹⁵³

$$RMSECV = \sqrt{\frac{1}{M} \sum_{s=1}^S \sum_{i=1}^{N_s} (\hat{y}_{s,i} - y_{s,i})^2}$$

Ecuación I - 19

Donde M es el número total de muestras en el conjunto de calibración, S es el número total de segmentos de *cross-validation* y N_s el número de muestras de cada segmento. $\hat{y}_{s,i}$ es el valor calculado por el modelo con α componentes para la muestra i en cada segmento e $y_{s,i}$ el valor obtenido para la misma muestra con el método de referencia.

Al ir aumentando el número de factores PLS, el modelo se ajusta cada vez mejor a las muestras del conjunto de calibración, explicando cada vez mejor los datos y disminuyendo continuamente RMSEC. Esto puede llevar a la construcción de un modelo "a medida" de las muestras del conjunto de calibración, pero inefectivo para el cálculo de nuevas muestras que no estén presentes en el conjunto de calibración. Es el efecto conocido como sobreajuste del modelo (del inglés *overfitting*). El caso contrario, si no se utilizan suficientes factores PLS y, por lo tanto, no se incluye toda la variabilidad descrita en las variables X , la capacidad predictiva del modelo se resiente de la misma manera, es lo que conoce como falta de ajuste (del inglés *underfitting*).^{153,155}

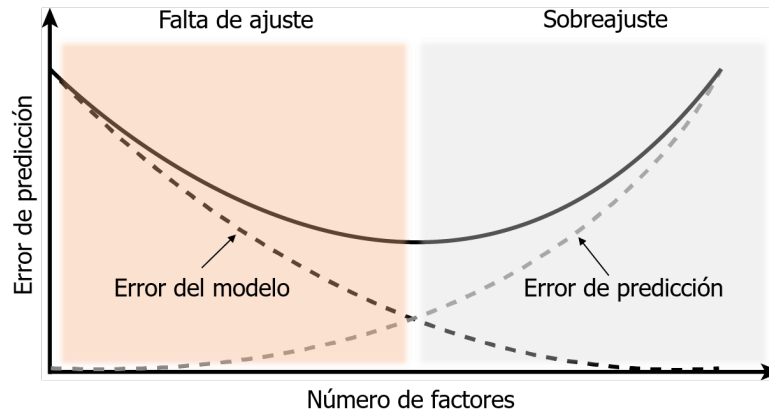


Figura I - 29. Variación del error de predicción al aumentar el número de factores.

En la Figura I - 29, se muestra un esquema del proceso de sobreajuste. El número óptimo de factores PLSR del modelo correspondería al mínimo de la curva, aunque no siempre el comportamiento es tan claro como el representado en la figura.

En el valor de RMSECV se incluye una fuente de error relacionada con el error aleatorio, el error estándar de remuestreo (SECV, siglas del inglés *Standard Error of Cross-Validation*), que sería equivalente a la desviación estándar del calibrado; y otro relacionado con el error sistemático, sesgo (en inglés *Bias*).

$$RMSECV^2 \approx SECV^2 + Sesgo^2 \quad \text{Ecuación I - 20}$$

Una vez decidido el número de factores PLSR del modelo, la desviación estándar de los resultados hallados por *cross-validation* es:

$$SECV = \sqrt{\frac{1}{M-1} \sum_{i=1}^M (\hat{y}_i - y_i - Sesgo)^2} \quad \text{Ecuación I - 21}$$

La cuantificación del error sistemático del modelo es:

$$Sesgo (Bias) = \frac{1}{M} \sum_{i=1}^M (\hat{y}_i - y_i) \quad \text{Ecuación I - 22}$$

La representación de los valores Y de las muestras del conjunto de calibración, calculados por el modelo, frente a los valores hallados por el método de referencia, ha de ser una recta que evaluada estadísticamente tenga una pendiente de valor 1, una ordenada en el origen de valor 0 y con un sesgo (*Bias*) de 0.

5.4.9. Validación del modelo quimiométrico

La validación del modelo quimiométrico es la evaluación de su capacidad predictiva, tanto si se trata de un modelo cualitativo o cuantitativo. Esta validación se debe realizar con muestras independientes de las empleadas durante el desarrollo, es decir, con las muestras del conjunto de predicción.

Además, para poder utilizar el modelo en el entorno regulado de la industria farmacéutica, su validación se debe realizar siguiendo las indicaciones de las agencias reguladoras. El proceso de validación de un método multivariante está descrito en la sección 6.3 de la presente tesis. En el caso de que la validación no sea satisfactoria, indicaría que el modelo es deficiente para cumplir con el objetivo y que, por lo tanto, debe rehacerse (escoger otro pretratamiento o combinación de estos, otro rango espectral, etc.).

5.4.10. Aplicación y mantenimiento del modelo

En el ámbito farmacéutico, una vez el modelo ha sido validado y se ha documentado el procedimiento analítico, detallando cómo realizar las medidas de las nuevas muestras y

cómo tratar los espectros de éstas, ya puede emplearse en rutina para determinar la propiedad estudiada en nuevas muestras.

Sin embargo, debido a que la técnica NIR es capaz de detectar tanto cambios físicos como químicos en las muestras, pequeñas modificaciones que se puedan producir durante el proceso podrían no estar contempladas en el modelo original. Por esta razón, los modelos NIR necesitan planificar un mantenimiento a lo largo del tiempo con el fin de asegurar su carácter predictivo.^{171,172}

6. VALIDACIÓN DE MÉTODOS ANALÍTICOS

Un método analítico se define como "la aplicación concreta de una técnica analítica para la determinación de un analito o mezcla en una muestra definida".¹⁵¹ El propósito final del mismo consiste en proporcionar resultados robustos, fiables y exactos de la propiedad estudiada. Por esta razón, es muy importante conocer su alcance y limitaciones antes de utilizarlo en rutina. Este conocimiento se obtiene mediante la validación del método.¹⁷³

Si nos centramos en la industria farmacéutica, la validación es una actividad fundamental para asegurar la calidad de la metodología empleada y se aplica no únicamente a los métodos analíticos sino también a sistemas, procesos, limpieza o servicios.^{91,174}

En este apartado se ofrece una descripción sobre el concepto de validación y como aplicar éste en los métodos analíticos de la industria farmacéutica.

6.1. Introducción

Debido a que los análisis que se realizan en los laboratorios farmacéuticos de todo el mundo están relacionados con la seguridad de la sanidad humana, es fundamental asegurar la validez de los resultados obtenidos para, a partir de ellos, poder tomar las decisiones pertinentes. Por esta razón, en 1971, las agencias regulatorias empezaron a publicar guías para definir un proceso que garantizara la calidad y autenticidad de los resultados analíticos obtenidos.^{173,175,176}

The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use conocida por sus siglas ICH, es una organización de la que forman parte tanto autoridades reguladoras como representantes de la industria farmacéutica. La principal función de la ICH consiste en realizar recomendaciones que ayudan a la interpretación y aplicación de guías técnicas y requisitos para el registro de productos farmacéuticos. Una de estas guías es la *Validation of analytical procedures: text and methodology. Q2 (R1)* dedicada a la validación de los métodos analíticos la cual se ha convertido en la referencia actual para el proceso de validación.¹⁷⁶

En esta guía se utiliza la palabra *procedures*, cuyo equivalente en español es procedimiento, para referirse a la "descripción pormenorizada del método analítico".¹⁵¹ Sin embargo, la FDA en este contexto define *methods* y *procedures* como sinónimos.¹⁷⁵ Por lo que en esta tesis se ha decidido utilizar la expresión validación de métodos

analíticos para referirse a este proceso de validación. Este término contempla todas las etapas que demuestran y documentan que un método analítico es apto para su uso, es decir, que los resultados obtenidos garantizan de forma inequívoca la calidad del producto.^{174,177,178}

En la industria farmacéutica, todos los métodos analíticos deben ser validados, independientemente de si van a usarse para analizar la materia prima, el producto terminado o la estabilidad de éstos.

El proceso de validación no incluye únicamente el apartado técnico, sino que debe estar documentado y planificado, detallando previamente qué pruebas se van a realizar, los criterios de aceptación para cada una de ellas y los resultados obtenidos, entre otros.¹⁷⁹

En la Figura I - 30 se presenta un esquema de la documentación implicada en el proceso de validación de un método analítico, juntamente con una breve descripción de que debe contener cada una.

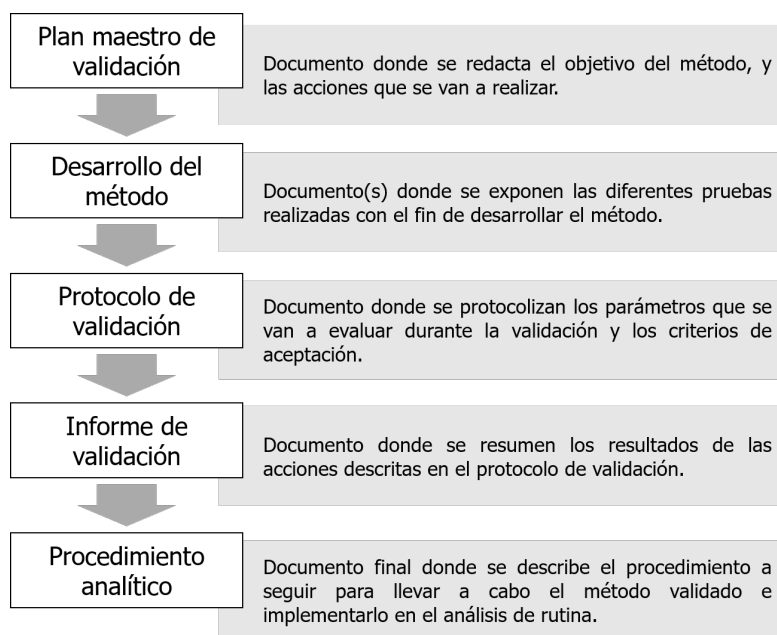


Figura I - 30. Esquema temporal de la documentación necesaria para desarrollar y validar un método analítico en la industria farmacéutica.

6.2. Propiedades analíticas estudiadas en la validación de métodos analíticos

Durante la validación, se deben determinar un conjunto de propiedades analíticas del método con el fin de aportar evidencias sobre la calidad y la consistencia de los resultados. Además, para su aplicación definitiva en rutina, se debe estudiar parámetros operacionales/económicos que aporten información sobre la viabilidad de uso del

método en rutina.^{162,173} Sin embargo, teniendo en cuenta los objetivos de la tesis, únicamente se estudiarán las propiedades analíticas.

Las propiedades analíticas estudiadas son: la exactitud, la precisión (repetibilidad y precisión intermedia), la especificidad, el límite de detección y de cuantificación, la linealidad, el intervalo de linealidad y la robustez.

Dependiendo del objetivo del método analítico que se quiera validar, es necesario evaluar distintas propiedades. Así, para la identificación del producto únicamente suele evaluarse la especificidad, mientras que, en otras circunstancias, como en la determinación de la concentración del API o de las posibles impurezas presentes, se deben estudiar muchas más propiedades. Las propiedades estudiadas en cada caso se muestran en la Tabla I - 5.

Tabla I - 5. Tabla resumen sobre las propiedades que se deben estudiar durante la validación dependiendo del tipo de método analítico que se evalué, información adaptada de las referencias.^{176,180}

Propiedades	Método analítico			
	Identificación	Impurezas		Valoración
		Cuantitativo	Límite	
Especificidad	Sí	Sí	Sí	Sí
Linealidad	No	Sí	No	Sí
Intervalo de linealidad	No	Sí	No	Sí
Exactitud	No	Sí	No	Sí
Precisión	No	Sí	No	Sí
- Repetibilidad	No	Sí	No	Sí
- Precisión intermedia	No	Sí	No	Sí
Límite de detección	No	No	Sí	No
Límite de cuantificación	No	Sí	No	No

En las siguientes secciones, se definen esas propiedades y se realiza un breve comentario sobre la metodología empleada en la validación.

6.2.1. Especificidad

La especificidad de un método analítico se define como "la habilidad de evaluar inequívocamente el analito en presencia de componentes que se puede esperar que

estén presentes en una muestra".¹⁷⁶ Estos componentes pueden acompañar la matriz de la muestra (excipientes), o ser impurezas.

Un método analítico, independientemente de si se emplea para obtener resultados cualitativos o cuantitativos, debe tener una buena especificidad. Debido a que, si falla esta propiedad, raramente se alcanzarán los niveles deseados en las demás propiedades analíticas de la validación. Una mala especificidad normalmente significa un mal desarrollo del método.^{173,177}

6.2.2. Linealidad

La linealidad se define como "la habilidad de obtener resultados directamente proporcionales a la concentración (o a la cantidad) del analito en la muestra dentro del intervalo estudiado".¹⁷⁶

La relación entre la señal obtenida en el instrumento y la concentración del analito, para la mayoría de las técnicas, sigue la función de una recta, conocida como recta de calibración (Ecuación I - 23).

$$Y = aX + b \qquad \text{Ecuación I - 23}$$

Donde Y es la señal del instrumento, X la concentración de la muestra, a la pendiente de la recta y b la ordenada en el origen de ésta.

Según las guías de las agencias reguladoras, la recta de calibración debe prepararse con un mínimo de cinco niveles repartidos en el intervalo de concentración que se quiera estudiar, con al menos tres replicados en cada nivel; y la linealidad de la misma debe ser evaluada estadísticamente para darla como conforme.^{173,177}

6.2.3. Intervalo de linealidad

El intervalo de linealidad se define como "el intervalo entre la máxima y la mínima concentración del analito en la muestra para las cuales se ha demostrado que el método analítico tiene un nivel adecuado de linealidad, exactitud y precisión".¹⁷⁶

Habitualmente, este parámetro se estudia a la vez que la linealidad, y su valor mínimo depende del objetivo del método analítico: concentración de producto acabado, uniformidad de contenido, pruebas de disolución, análisis de impurezas, etc.^{176,177}

6.2.4. Exactitud

La exactitud es la propiedad analítica más importante, y se define como “el grado de concordancia entre el valor obtenido (o la media de los valores obtenidos) en la medida experimental de una muestra y el valor de ésta aceptado como verdadero”.¹⁷⁶ La determinación de este parámetro permite detectar errores sistemáticos que afectan al resultado del método analítico.¹⁸¹

Las guías de las agencias reguladoras describen diferentes estrategias para la determinación de la exactitud, que están relacionadas con el tipo de análisis realizado. Por ejemplo, se puede aplicar el método analítico a un material de referencia que contenga el analito; o comparar los resultados obtenidos por el método que se está validando con los obtenidos utilizando un método de referencia; o utilizar un método de adición estándar.

Este parámetro se puede expresar como la relación entre la concentración encontrada experimentalmente y la concentración real (normalmente se expresa en porcentaje). Cuando más próximo sea este valor a 100 % más exacto será el método.¹⁷³

Independientemente de la estrategia que se escoja, se recomienda estudiar la exactitud con un mínimo de nueve determinaciones y, al menos, a tres niveles de concentración que abarquen todo el intervalo investigado. En el caso de la ICH se recomienda analizar tres replicados en cada punto, mientras que la FDA recomienda pentareplicados.^{175,176}

6.2.5. Precisión

La precisión se define como “el grado de concordancia entre una serie de mediciones, obtenidas de manera independiente, de la misma muestra homogénea en las condiciones del método analítico”.¹⁷⁶

Esta propiedad aporta la dispersión o la variabilidad de un resultado y, se acostumbra a cuantificar empleando la varianza (s^2), la desviación estándar (s) o la desviación estándar relativa (RSD, siglas del inglés *Relative Standard Deviation*) habitualmente expresada en porcentaje. Cuanto menor sea la desviación estándar o el RSD (%) mayor será la precisión que tendrá el método analítico en las condiciones estudiadas.¹⁷³

La precisión de un método se puede considerar a tres niveles:^{176,182}

- Repetibilidad: este nivel aporta la medida de la variabilidad en los resultados obtenidos cuando con el método analítico se analiza una misma muestra, varias

veces, en un corto periodo de tiempo. Se recomienda evaluar esta propiedad con un mínimo de 9 determinaciones distribuidas en todo el intervalo validado o realizando 6 determinaciones en el 100 % de la concentración del analito.

- Precisión intermedia: este nivel expresa los efectos aleatorios que pueden ocurrir cuando se usa el método analítico en rutina. Es decir, se modifican factores como el día de análisis, el equipo o el analista, para ver cómo varían los resultados al analizar una misma muestra.
- Reproducibilidad: este nivel aporta la variabilidad en los resultados obtenidos si el análisis se realiza en distintos laboratorios. La validación de esta propiedad exige un ejercicio colaborativo entre distintos laboratorios.

6.2.6. Límite de detección

El límite de detección (LOD, siglas del inglés *Limit of Detection*) es un valor estadístico fuera del intervalo de linealidad validado, que se define como "la mínima cantidad de analito en una muestra que puede ser detectada", pero no necesariamente cuantificada como un valor exacto.¹⁷⁶

Por lo tanto, las concentraciones que se encuentren por encima del LOD se podrán asociar al analito sin necesariamente llegar a cuantificarse, mientras que si se encuentran por debajo no se podrá afirmar la presencia del analito.¹⁷³

La ICH propone diferentes aproximaciones para determinar el LOD, como la basada en la evaluación visual, en la relación señal/ruido, en la desviación estándar de la respuesta y de la pendiente en una recta de calibración a concentraciones cercanas al LOD, o la que utiliza la desviación estándar de una serie de blancos.¹⁷⁶

6.2.7. Límite de cuantificación

El límite de cuantificación (LOQ, siglas del inglés *Limit of Quantification*) se define como "la mínima cantidad de analito en una muestra que puede determinarse cuantitativamente con una precisión y exactitud adecuadas".¹⁷⁶

Las estrategias que se proponen en las distintas guías son las mismas que las propuestas para la determinación del LOD, pero aumentando el nivel de concentración.¹⁸¹ Sin embargo, en el caso del LOD se requiere demostrar que a este nivel de concentración se trabaja con una precisión y exactitud aceptable, por lo que, aunque se determine estadísticamente, el LOD debe ser verificado de forma experimental.¹⁸²

6.2.8. Robustez

La robustez de un método analítico se puede definir como "la capacidad para no verse afectado por variaciones pequeñas pero deliberadas en los parámetros del método".¹⁷⁶

Esta propiedad proporciona una indicación de la fiabilidad del método durante su uso normal. No está incluida en la tabla resumen de la ICH (Tabla I - 5) debido a que se considera que esta propiedad debe ser estudiada durante el desarrollo del método analítico para así poder ajustar adecuadamente los parámetros experimentales del procedimiento.¹⁸²

Para el estudio de la robustez, primero se deben determinar los factores experimentales que afectan al método analítico y establecer el intervalo de trabajo en el que se quiera evaluar dicha propiedad. Esta evaluación se puede realizar estudiando cada parámetro independientemente o evaluándolos conjuntamente. Esta última estrategia, tiene en cuenta la interacción entre los diferentes parámetros y emplea un diseño de experimentos.¹⁷³ También puede realizarse un estudio a lo largo del tiempo de alguno de los parámetros del análisis para asegurar que se mantiene la validez del método.

6.3. Validación de métodos analíticos multivariantes

La validación de métodos analíticos en los que se ha utilizado un método multivariante para hallar el resultado, como podrían ser las técnicas espectroscópicas, deben comprobarse las mismas propiedades analíticas descritas en la Tabla I - 5. Sin embargo, dado el tratamiento matemático de los datos en estas técnicas, los parámetros de validación no pueden realizarse como se ha descrito anteriormente; sino que deben adaptarse, ya que la ICH Q2 (R1) actual no está pensada para contemplar estos métodos.¹¹³ Por ello, esta guía se encuentra actualmente en revisión para poder adaptar el contenido a los métodos analíticos multivariantes.¹⁸³

A pesar de esto, diferentes entidades reguladoras, como Ph. Eur., FDA, EMA (siglas del inglés *European Medicines Agency*), entre otras, han ido desarrollando diferentes guías dando pautas para la validación de los métodos analíticos multivariantes.^{113,114,163}

En la siguiente sección se describen las principales adaptaciones de los parámetros de la validación.

6.3.1. Adaptación de los parámetros de la validación de métodos analíticos multivariantes

La evaluación de los parámetros de la validación en métodos analíticos multivariantes debe realizarse con el conjunto de muestras validación, descrito en la sección 5.4.7.¹⁸⁴

- **Especificidad:** aunque la definición es la misma que la comentada en secciones anteriores, la evaluación de este parámetro en métodos analíticos multivariantes varía dependiendo de si se trata de un modelo cualitativo o cuantitativo. En el primer caso debe estudiarse el porcentaje de falsos positivos y negativos obtenidos tras predecir las muestras del conjunto de validación. En el caso de los modelos cuantitativos, se debe demostrar aportando evidencias que:
 - El modelo desarrollado va a predecir la propiedad bajo estudio.
 - Las muestras del conjunto de validación incluyen la variabilidad estudiada por el modelo.
 - El conjunto de validación no contiene muestras anómalas.
- **Linealidad:** este parámetro sólo debe ser evaluado en modelos cuantitativos. A diferencia de los métodos univariantes, en los que se evalúa la señal registrada frente a la concentración, en los multivariantes se estudia estadísticamente la linealidad de la recta obtenida al representar los resultados calculados con el modelo frente a los del método de referencia. Se debe evaluar estadísticamente que el modelo es lineal; la manera recomendada es evaluar la pendiente, ordenada de origen, el coeficiente de determinación (R^2) y estudiar los residuales, aunque se puede justificar cualquier estadística utilizada con el mismo fin.
- **Intervalo de linealidad:** la evaluación de este parámetro sólo se lleva a cabo en modelos cuantitativos y se realiza de igual manera que para los métodos univariantes descrita en la sección 6.2.3.
- **Exactitud:** este parámetro solamente se estudia en métodos cuantitativos y consiste en evaluar el grado de semejanza entre el valor calculado con el modelo y el valor del método de referencia a lo largo del intervalo de concentración del método. Para ello, puede hacerse una prueba *t* de muestras apareadas y estudiar si hay diferencias estadísticamente significativas entre ambos métodos. Además, se debe demostrar que el sesgo no es estadísticamente diferente de 0.
En el conjunto de guías de las agencias regulatorias ya mencionadas, también se recomienda evaluar que el error de predicción (RMSEP) sea comparable al error

de la calibración (RMSEC); así como que la relación entre el RMSEP y el error del método de laboratorio (SEL, siglas del inglés *Standard Error of Laboratory*) sea inferior a 1,4.

$$SEL = \sqrt{\frac{\sum_{i=1}^n (y_{1,i} - y_{2,i})^2}{n}}$$

Ecuación I - 24

Donde n es el número de muestras, e $y_{j,i}$ los valores obtenidos con el método de referencia para la muestra i y los distintos replicados $j = 1$ y $j = 2$.

- Precisión: al igual que para el intervalo de linealidad e independientemente del nivel (repetibilidad, precisión intermedia o reproducibilidad), este parámetro se realiza y evalúa igual que para los métodos univariantes.
- Límite de detección y cuantificación: para los métodos multivariantes, la evaluación de estos parámetros sólo se lleva a cabo en las mismas situaciones que en los métodos univariantes y en las guías regulatorias no se hace distinción entre ambos límites. El valor de ambos es el límite inferior del intervalo de linealidad, es decir, el último punto donde se asegura la linealidad, exactitud y precisión del método.
- Robustez: se han de presentar datos para la determinación y la garantía de la robustez sino se ha considerado esta propiedad en las etapas de desarrollo y optimización del modelo. Al igual que en los métodos univariantes, las guías no proponen una forma específica de hacerlo.

7. BIBLIOGRAFÍA

- (1) GROVES, M. J. Perspective on the Use of Parenteral Solutions. In *Parenteral Products*; Elsevier, 1973; Vol. i, pp 1–14.
- (2) Real Academia Española. Real Academia Española <https://www.rae.es/>.
- (3) Food and Drug Administration. Food and Drug Administration <https://www.fda.gov/drugs/drug-approvals-and-databases/drugsfda-glossary-terms>.
- (4) Salazar Macian, R. Aplicación Del Concepto de Calidad Total a La Industria Farmacéutica; Real Academia de Farmacia de Barcelona, 1991.
- (5) Eichhorn, P. General Introduction on Pharmaceuticals. In *Comprehensive Analytical Chemistry*; Elsevier B.V., 2013; Vol. 62, pp 1–36.
- (6) Dukes, G. The Pharmaceutical Industry and Its Products. In *The Law and Ethics of the Pharmaceutical Industry*; Elsevier, 2006; pp 3–15.
- (7) Speight, J. G. Pharmaceuticals. In *Handbook of Industrial Hydrocarbon Processes*; Elsevier, 2020; pp 553–595.
- (8) European Patients' Academy on Therapeutic Innovation. Fundamentos de la regulación de medicamentos <https://www.eupati.eu>.
- (9) Laboratorio Reig Jofre. Laboratorio Reig Jofre <https://www.reigjofre.com/es/>.
- (10) Nema, S.; Ludwig, J. D. *Pharmaceutical Dosage Forms - Parenteral Medications*, 3rd ed.; Nema, S., Ludwig, J. D., Eds.; CRC Press, 2016; Vol. 1.
- (11) Nireesha, G. R.; Divya, L.; Sowmya, C.; Venkateshan, N.; Niranjan Babu, M.; Lavakumar, V. Lyophilization/Freeze Drying - An Review. *Int. J. Nov. Trends Pharm. Sci.* 2013,3 (4), 87–98.
- (12) Franks, F. *Freeze-Drying of Pharmaceuticals and Biopharmaceuticals*; Royal Society of Chemistry: Cambridge, 2007.
- (13) Hua, T.-C.; Liu, B.-L.; Zhang, H. Fundamentals of Freeze Drying. In *Freeze-Drying of Pharmaceutical and Food Products*; Elsevier, 2010; pp 18–67.
- (14) Porfire, A.; Tomuta, I.; Iurian, S.; Casian, T. Quality by Design Considerations for the Development of Lyophilized Products. In *Pharmaceutical Quality by Design*; Elsevier, 2019; pp 193–207.
- (15) Cover, J. The Evolution of Freeze-Drying. *Innov. Pharm. Technol.* 2009, 66–70.
- (16) Varshney, D.; Singh, M. *Lyophilized Biologics and Vaccines*; Varshney, D., Singh, M., Eds.; Springer New York: New York, NY, 2015.
- (17) Freeze-Dry Foods Ltd. Freeze-Dry Foods <http://www.freeze-dry.com/2015/09/29/history-of-freeze-drying/> (accessed Dec 15, 2019).
- (18) Jennings, T. A. *Lyophilization*; CRC Press, 1999.
- (19) Shackell, L. F. AN IMPROVED METHOD OF DESICCATION, WITH SOME APPLICATIONS TO BIOLOGICAL PROBLEMS. *Am. J. Physiol. Content* 1909, 24(3), 325–340.
- (20) Hansen, L. J. J.; Daoussi, R.; Vervaet, C.; Remon, J.-P.; De Beer, T. R. M. Freeze-Drying of Live Virus Vaccines: A Review. *Vaccine* 2015, 33 (42), 5507–5519.

- (21) Goshima, H.; Do, G.; Nakagawa, K. Impact of Ice Morphology on Design Space of Pharmaceutical Freeze-Drying. *J. Pharm. Sci.* 2016, *105* (6), 1920–1933.
- (22) Berk, Z. Freeze Drying (Lyophilization) and Freeze Concentration. In *Food Process Engineering and Technology*; Elsevier, 2018; pp 567–581.
- (23) Tang, X. (Charlie); Pikal, M. J. Design of Freeze-Drying Processes for Pharmaceuticals: Practical Advice. *Pharm. Res.* 2004, *21* (2), 191–200.
- (24) Sharma, P.; Kessler, W. J.; Bogner, R.; Thakur, M.; Pikal, M. J. Applications of the Tunable Diode Laser Absorption Spectroscopy: In-Process Estimation of Primary Drying Heterogeneity and Product Temperature During Lyophilization. *J. Pharm. Sci.* 2019, *108* (1), 416–430.
- (25) Patel, S. M.; Chaudhuri, S.; Pikal, M. J. Choked Flow and Importance of Mach I in Freeze-Drying Process Design. *Chem. Eng. Sci.* 2010, *65* (21), 5716–5727.
- (26) Kawasaki, H.; Shimanouchi, T.; Kimura, Y. Recent Development of Optimization of Lyophilization Process. *J. Chem.* 2019, *2019*, 1–14.
- (27) Kremer, D. M.; Pikal, M. J.; Petre, W. J.; Shalaev, E. Y.; Gatlin, L. A.; Kramer, T. A Procedure to Optimize Scale-up for the Primary Drying Phase of Lyophilization. *J. Pharm. Sci.* 2009, *98* (1), 307–318.
- (28) Kasper, J. C.; Friess, W. The Freezing Step in Lyophilization: Physico-Chemical Fundamentals, Freezing Methods and Consequences on Process Performance and Quality Attributes of Biopharmaceuticals. *Eur. J. Pharm. Biopharm.* 2011, *78* (2), 248–263.
- (29) Liu, J. Physical Characterization of Pharmaceutical Formulations in Frozen and Freeze-Dried Solid States: Techniques and Applications in Freeze-Drying Development. *Pharm. Dev. Technol.* 2006, *11* (1), 3–28.
- (30) Searles, J. A.; Carpenter, J. F.; Randolph, T. W. Annealing to Optimize the Primary Drying Rate, Reduce Freezing-induced Drying Rate Heterogeneity, and Determine T_g in Pharmaceutical Lyophilization. *J. Pharm. Sci.* 2001, *90* (7), 872–887.
- (31) Parker, A.; Rigby-Singleton, S.; Perkins, M.; Bates, D.; Le Roux, D.; Roberts, C. J.; Madden-Smith, C.; Lewis, L.; Teagarden, D. L.; Johnson, R. E.; et al. Determination of the Influence of Primary Drying Rates on the Microscale Structural Attributes and Physicochemical Properties of Protein Containing Lyophilized Products. *J. Pharm. Sci.* 2010, *99* (11), 4616–4629.
- (32) Williams, N. A.; Polli, G. P. The Lyophilization of Pharmaceuticals: A Literature Review. *J. Parenter. Sci. Technol.* 1984, *38* (2), 48–60.
- (33) Pikal, M. J.; Shah, S. The Collapse Temperature in Freeze Drying: Dependence on Measurement Methodology and Rate of Water Removal from the Glassy Phase. *Int. J. Pharm.* 1990, *62* (2–3), 165–186.
- (34) Schersch, K.; Betz, O.; Garidel, P.; Muehlau, S.; Bassarab, S.; Winter, G. Systematic Investigation of the Effect of Lyophilizate Collapse on Pharmaceutically Relevant Proteins, Part 2: Stability During Storage at Elevated Temperatures. *J. Pharm. Sci.* 2012, *101* (7), 2288–2306.
- (35) Schersch, K.; Betz, O.; Garidel, P.; Muehlau, S.; Bassarab, S.; Winter, G. Systematic Investigation of the Effect of Lyophilizate Collapse on Pharmaceutically Relevant Proteins III: Collapse during Storage at Elevated Temperatures. *Eur. J. Pharm. Biopharm.* 2013, *85* (2), 240–252.

-
- (36) Luthra, S. A.; Hodge, I. M.; Pikal, M. J. Investigation of the Impact of Annealing on Global Molecular Mobility in Glasses: Optimization for Stabilization of Amorphous Pharmaceuticals. *J. Pharm. Sci.* 2008, *97*(9), 3865–3882.
- (37) Ohori, R.; Yamashita, C. Effects of Temperature Ramp Rate during the Primary Drying Process on the Properties of Amorphous-Based Lyophilized Cake, Part 1: Cake Characterization, Collapse Temperature and Drying Behavior. *J. Drug Deliv. Sci. Technol.* 2017, *39*, 131–139.
- (38) Patel, S. M.; Nail, S. L.; Pikal, M. J.; Geidobler, R.; Winter, G.; Hawe, A.; Davagnino, J.; Rambhatla Gupta, S. Lyophilized Drug Product Cake Appearance: What Is Acceptable? *J. Pharm. Sci.* 2017, *106*(7), 1706–1721.
- (39) WANG, W.; CHEN, M.; CHEN, G. Issues in Freeze Drying of Aqueous Solutions. *Chinese J. Chem. Eng.* 2012, *20*(3), 551–559.
- (40) Patel, S. M.; Pikal, M. J. Emerging Freeze-Drying Process Development and Scale-up Issues. *AAPS PharmSciTech* 2011, *12*(1), 372–378.
- (41) Shivanand, S. M. A Review on Lyophilization : A Technique to Improve Stability of Hygroscopic , Thermolabile Substances. *PharmaTutor* 2017, *5*(11), 28–39.
- (42) Pikal, M. J.; Shah, S. Intravial Distribution of Moisture during the Secondary Drying Stage of Freeze Drying. *PDA J. Pharm. Sci. Technol.* 1997, *51*(1), 17–24.
- (43) Fissore, D.; Pisano, R.; Barresi, A. A. Process Analytical Technology for Monitoring Pharmaceuticals Freeze-Drying – A Comprehensive Review. *Dry. Technol.* 2018, *36*(15), 1839–1865.
- (44) PIKAL, M.; SHAH, S.; ROY, M.; PUTMAN, R. The Secondary Drying Stage of Freeze Drying: Drying Kinetics as a Function of Temperature and Chamber Pressure☆. *Int. J. Pharm.* 1990, *60*(3), 203–207.
- (45) Fissore, D.; Pisano, R.; Barresi, A. A. On the Methods Based on the Pressure Rise Test for Monitoring a Freeze-Drying Process. *Dry. Technol.* 2010, *29*(1), 73–90.
- (46) Geoghegan-Quinn, M. European Commission Clarifies the Rules for Research Audits. *Nature* 2010, *466*(7306), 562–562.
- (47) Oetjen, G.-W.; Haseley, P. *Freeze-Drying*, 2nd ed.; Wiley-VCH GmbH & Co. KGaA, 2004.
- (48) Padivitage, N. L. T.; Smuts, J. P.; Armstrong, D. W. Water Determination. In *Specification of Drug Substances and Products*; Elsevier, 2014; pp 223–241.
- (49) Galletti, G. C.; Piccaglia, R. Water Determination in Silages by Karl Fischer Titration. *J. Sci. Food Agric.* 1988, *43*(1), 1–7.
- (50) Margolis, S.; Huang, P.; Hădărugă, N. G.; Hădărugă, D. I. Water Determination. In *Reference Module in Chemistry, Molecular Sciences and Chemical Engineering*; Elsevier, 2018; p 760.
- (51) Meadowcroft, D. B.; Verhappen, I. Chemical Analysis. In *Instrumentation Reference Book*; Verhappen, I., Ed.; Elsevier, 2003; Vol. c, pp 414–421.
- (52) Isengard, H.-D. Water Determination in Food. In *Ensuring Global Food Safety*; Elsevier, 2010; pp 109–123.
- (53) Scholz, E. *Karl Fischer Titration*, 1st ed.; Chemical Laboratory Practice; Springer Berlin Heidelberg: Berlin, Heidelberg, 1984.
- (54) Fischer, K. Neues Verfahren Zur Maßanalytischen Bestimmung Des
-

- Wassergehaltes von Flüssigkeiten Und Festen Körpern. *Angew. Chemie* 1935, 48 (26), 394–396.
- (55) USP. *Water Determination*; 921.
- (56) European Pharmacopoeia. 2.5.12. Water Semi-Micro Determination. 2018, 20512.
- (57) European Pharmacopoeia. 2.5.32. Water Micro Determination. 2019, 20532.
- (58) Cedergren, A. Reaction Rates between Water and the Karl Fischer Reagent. *Talanta* 1974, 21 (4), 265–271.
- (59) Smith, D. M.; Bryant, W. M. D.; Mitchell, J. Analytical Procedures Employing Karl Fischer Reagent. I. Nature of the Reagent. *J. Am. Chem. Soc.* 1939, 61 (9), 2407–2412.
- (60) Verhoef, J. C.; Barendrecht, E. Mechanism and Reaction Rate of the Karl-Fischer Titration Reaction. *J. Electroanal. Chem. Interfacial Electrochem.* 1976, 71 (2), 305–315.
- (61) Grünke, S. Main and Side Reactions in the Karl Fischer Solution. *Food Control* 2001, 12 (7), 419–426.
- (62) Skoog, D. A.; West, D. M.; Holler, F. J.; Crouch, S. R. *Fundamento de Química Analítica*, 9th ed.; Orozco, V., Abril, Eds.; Cengage Learning Editores, S.A., 2015.
- (63) Larsson, W. New Approaches to Moisture Determination in Complex Matrices Based on the Karl Fischer Reaction in Methanolic and Non-Alcoholic Media, Umea Universitet, 2008.
- (64) Johansson, A. Determination of Water by Titration. A Modified Karl Fischer Method. *Sven. Papperstidning* 1947, 50, 124–126.
- (65) Meyer, A. S.; Boyd, C. M. Determination of Water by Titration With Coulometrically Generated Karl Fischer Reagent. *Anal. Chem.* 1959, 31 (2), 215–219.
- (66) Scholz, E. Karl Fischer Coulometry? The Cathode Reaction. *Fresenius. J. Anal. Chem.* 1994, 348 (4), 269–271.
- (67) Nordmark, U.; Cedergren, A. Conditions for Accurate Karl Fischer Coulometry Using Diaphragm-Free Cells. *Anal. Chem.* 2000, 72 (1), 172–179.
- (68) Cedergren, A.; Jonsson, S. Progress in Karl Fischer Coulometry Using Diaphragm-Free Cells. *Anal. Chem.* 2001, 73 (22), 5611–5615.
- (69) Swarbrick, B. The Current State of near Infrared Spectroscopy Application in the Pharmaceutical Industry. *J. Near Infrared Spectrosc.* 2014, 22 (3), 153–156.
- (70) Burns, D. A.; Ciurczak, E. W.; Donald A., B.; Emil W., C.; Burns, D. A.; Ciurczak, E. W. *Handbook of Near-Infrared Analysis*, Third Edit.; Taylor & Francis Group, LLC, 2007.
- (71) Jamrógiewicz, M. Application of the Near-Infrared Spectroscopy in the Pharmaceutical Technology. *J. Pharm. Biomed. Anal.* 2012, 66, 1–10.
- (72) Bush, L.; Workman, J.; Chasse, J.; Delonas, C. Solutions for materials analysis Spectroscopy <http://www.spectroscopyonline.com/timeline-atomic-spectroscopy>.
- (73) Thomas, N. C. The Early History of Spectroscopy. *J. Chem. Educ.* 1991, 68 (8), 631.
- (74) Sun, D.-W. *Infrared Spectroscopy for Food Quality Analysis and Control*, 1st ed.;

- Elsevier, 2009; Vol. 53.
- (75) SZABADVÁRY, F.; ROBINSON, A. The History of Analytical Chemistry. In *Comprehensive Analytical Chemistry*; Elsevier, 1980; Vol. 10, pp 61–282.
- (76) Herschel, F. W. Experiments on the Refrangibility of the Invisible Rays of the Sun. *Abstr. Pap. Print. Philos. Trans. R. Soc. London* 1832, *1*, 22–23.
- (77) Ring, E. F. J. The Discovery of Infrared Radiation in 1800. *Imaging Sci. J.* 2000, *48* (1), 1–8.
- (78) Ciurczak, E. W.; Igne, B. *Pharmaceutical and Medical Applications of Near-Infrared Spectroscopy*, 1st ed.; CRC Press, 2014; Vol. 53.
- (79) Beć, K. B.; Huck, C. W. Breakthrough Potential in Near-Infrared Spectroscopy: Spectra Simulation. A Review of Recent Developments. *Front. Chem.* 2019, *7* (48), 1–22.
- (80) Pasquini, C. Near Infrared Spectroscopy: Fundamentals, Practical Aspects and Analytical Applications. *J. Braz. Chem. Soc.* 2003, *14* (2), 198–219.
- (81) Davies, T. Happy 90 Th Birthday to Karl Norris, Father of NIR Technology. *NIR news* 2011, *22* (4), 3–16.
- (82) McClure, W. F. Near-Infrared Spectroscopy. The Giant Is Running Strong. *Anal. Chem.* 1994, *66* (1), 43A-53A.
- (83) Blanco, M.; Villarroya, I. NIR Spectroscopy: A Rapid-Response Analytical Tool. *TrAC Trends Anal. Chem.* 2002, *21* (4), 240–250.
- (84) Ozaki, Y.; Huck, C. W.; Beć, K. B. Near-IR Spectroscopy and Its Applications. In *Molecular and Laser Spectroscopy*; Elsevier, 2018; pp 11–38.
- (85) Workman, J.; Weyer, L. A Programme for Men's Clubs on Citizenship *. *Relig. Educ.* 1908, *3* (1), 20–22.
- (86) Blanco, M.; Coello, J.; Iturriaga, H.; MasPOCH, S.; de la Pezuela, C. Near-Infrared Spectroscopy in the Pharmaceutical Industry. *Analyst* 1998, *123* (8), 135R-150R.
- (87) Smith, B. C. *Quantitative Spectroscopy: Theory and Practice*, 1st ed.; Elsevier, 2002.
- (88) Skoog, D. A.; Holler, F. J.; Nieman, T. A. *Principios de Análisis Instrumental*, 5th ed.; McGraw Hill: Madrid, 2001.
- (89) Kafle, B. P. Infrared (IR) Spectroscopy. In *Chemical Analysis and Material Characterization by Spectrophotometry*; Elsevier, 2020; pp 199–243.
- (90) Durig, J. R.; Panikar, S. S.; Klaassen, J. J. Far Infrared Spectroscopy Applications. In *Encyclopedia of Spectroscopy and Spectrometry*; Elsevier, 2017; pp 558–565.
- (91) Ríos, Á.; Moreno, M. C.; Simonet, B. M. *Técnicas Espectroscópicas En Química Analíticas. Volumen I. Aspectos Básicos y Espectrometría Molecular.*, 1st ed.; Editorial Síntesis, 2012.
- (92) Teixeira, J.; Vicente, A.; da Silva, F.; da Silva, J.; da Costa, R. Process Analytical Technology. In *Handbook of Pharmaceutical Granulation Technology, Second Edition*; Bakeev, K. A., Ed.; Wiley, 2013; pp 265–292.
- (93) Blanco, M.; Coello, J.; Iturriaga, H.; MasPOCH, S.; de la Pezuela, C. Near-Infrared Spectroscopy in the Pharmaceutical Industry. *Analyst* 1998, *123* (8), 135R-150R.

- (94) Siesler, H. W.; Ozaki, Y.; Kawata, H.; Heise, H. . *Near-Infrared Spectroscopy Principles, Instruments, Applications*; 2006.
- (95) Shaw, R. A.; Mantsch, H. H. Near-IR Spectrometers. In *Encyclopedia of Spectroscopy and Spectrometry*; Elsevier, 2017; pp 50–58.
- (96) Abbas, O.; Dardenne, P.; Baeten, V. Near-Infrared, Mid-Infrared, and Raman Spectroscopy. In *Chemical Analysis of Food: Techniques and Applications*; Elsevier, 2012; pp 59–89.
- (97) Manley, M.; Baeten, V. Spectroscopic Technique: Near Infrared (NIR) Spectroscopy. In *Modern Techniques for Food Authentication*; Elsevier, 2018; pp 51–102.
- (98) Agelet, L. E.; Hurburgh, C. R. A Tutorial on Near Infrared Spectroscopy and Its Calibration. *Crit. Rev. Anal. Chem.* 2010, *40* (4), 246–260.
- (99) OZAKI, Y. Near-Infrared Spectroscopy - Its Versatility in Analytical Chemistry. *Anal. Sci.* 2012, *28* (6), 545–563.
- (100) Deidda, R.; Sacre, P.-Y.; Clavaud, M.; Coïc, L.; Avohou, H.; Hubert, P.; Ziemons, E. Vibrational Spectroscopy in Analysis of Pharmaceuticals: Critical Review of Innovative Portable and Handheld NIR and Raman Spectrophotometers. *TrAC Trends Anal. Chem.* 2019, *114*, 251–259.
- (101) Luybaert, J.; Massart, D. L.; Vander Heyden, Y. Near-Infrared Spectroscopy Applications in Pharmaceutical Analysis. *Talanta* 2007, *72* (3), 865–883.
- (102) Fowle, F. E. The Spectroscopic Determination of Aqueous Vapor. *Astrophys. J.* 1912, *35*, 149.
- (103) Clavaud, M.; Lema-Martinez, C.; Roggo, Y.; Bigalke, M.; Guillemain, A.; Hubert, P.; Ziemons, E.; Allmendinger, A. Near-Infrared Spectroscopy to Determine Residual Moisture in Freeze-Dried Products: Model Generation by Statistical Design of Experiments. *J. Pharm. Sci.* 2020, *109* (1), 719–729.
- (104) Clua Palau, G. Application of Quality by Design and near Infrared Spectroscopy in Manufacturing and Control of Freeze-Dried Drug Products, Universitat Autònoma de Barcelona, 2018.
- (105) Blanco, M.; Romero, M. A. Near-Infrared Libraries in the Pharmaceutical Industry: A Solution for Identity Confirmation. *Analyst* 2001, *126* (12), 2212–2217.
- (106) Görög, S. Identification in Drug Quality Control and Drug Research. *TrAC Trends Anal. Chem.* 2015, *69*, 114–122.
- (107) Ali, H. R. H.; Edwards, H. G. M.; Scowen, I. J. Noninvasive in Situ Identification and Band Assignments of Some Pharmaceutical Excipients inside USP Vials with FT-near-Infrared Spectroscopy. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* 2009, *72* (4), 890–896.
- (108) Jones, J. A.; Last, I. R.; MacDonald, B. F.; Prebble, K. A. Development and Transferability of Near-Infrared Methods for Determination of Moisture in a Freeze-Dried Injection Product. *J. Pharm. Biomed. Anal.* 1993, *11* (11–12), 1227–1231.
- (109) Alvarenga, L.; Ferreira, D.; Altekruze, D.; Menezes, J. C.; Lochmann, D. Tablet Identification Using Near-Infrared Spectroscopy (NIRS) for Pharmaceutical Quality Control. *J. Pharm. Biomed. Anal.* 2008, *48* (1), 62–69.
- (110) Rose, J. R. Qualitative and Quantitative Analysis of Pharmaceuticals with NIRA. In

-
- Proceedings of the 2nd Annual Symposium on Near-Infrared Reflectance Analysis (NIRA)*; New York, 1983.
- (111) Stark, E.; Luchter, K.; Margoshes, M. Near-Infrared Analysis (NIRA): A Technology for Quantitative and Qualitative Analysis. *Appl. Spectrosc. Rev.* 1986, 22 (4), 335–399.
- (112) Miró Vera, A. Y.; Alcalà Bernàrdez, M. Near-Infrared Spectroscopy in Identification of Pharmaceutical Raw Materials. In *Encyclopedia of Analytical Chemistry*; John Wiley & Sons, Ltd: Chichester, UK, 2017; pp 1–19.
- (113) European Pharmacopoeia. 5.21. Chemometric Methods. 2016, 5641–5658.
- (114) European Medicines Agency. Guideline on the Use of near Infrared Spectroscopy by the Pharmaceutical Industry and the Data Requirements for New Submissions and Variations. 2014, p 28.
- (115) Li, J.; Yu, B.; Zhao, W.; Chen, W. A Review of Signal Enhancement and Noise Reduction Techniques for Tunable Diode Laser Absorption Spectroscopy. *Appl. Spectrosc. Rev.* 2014, 49 (8), 666–691.
- (116) Tittel, F. K.; Lewicki, R. Tunable Mid-Infrared Laser Absorption Spectroscopy. In *Semiconductor Lasers*; Elsevier, 2013; pp 579–629.
- (117) Duncan, D.; Cundell, T.; Levac, L.; Veale, J.; Kuiper, S.; Rao, R. The Application of Noninvasive Headspace Analysis to Media Fill Inspection. *PDA J. Pharm. Sci. Technol.* 2016, 70 (3), 230–247.
- (118) Werle, P.; Slemr, F.; Maurer, K.; Kormann, R.; Mücke, R.; Jänker, B. Near- and Mid-Infrared Laser-Optical Sensors for Gas Analysis. *Opt. Lasers Eng.* 2002, 37 (2–3), 101–114.
- (119) Butler, J. F.; Calawa, A. R.; Phelan, R. J.; Harman, T. C.; Strauss, A. J.; Rediker, R. H. PbTe DIODE LASER. *Appl. Phys. Lett.* 1964, 5 (4), 75–77.
- (120) Hinkley, E. D. HIGH-RESOLUTION INFRARED SPECTROSCOPY WITH A TUNABLE DIODE LASER. *Appl. Phys. Lett.* 1970, 16 (9), 351–354.
- (121) Roggo, Y.; Chalus, P.; Maurer, L.; Lema-Martinez, C.; Edmond, A.; Jent, N. A Review of near Infrared Spectroscopy and Chemometrics in Pharmaceutical Technologies. *J. Pharm. Biomed. Anal.* 2007, 44 (3), 683–700.
- (122) Bolshov, M. A.; Kuritsyn, Y. A.; Romanovskii, Y. V. Tunable Diode Laser Spectroscopy as a Technique for Combustion Diagnostics. *Spectrochim. Acta Part B At. Spectrosc.* 2015, 106 (July), 45–66.
- (123) Lundin, P.; Cocola, L.; Lewander, M.; Olsson, A.; Svanberg, S. Non-Intrusive Headspace Gas Measurements by Laser Spectroscopy – Performance Validation by a Reference Sensor. *J. Food Eng.* 2012, 111 (4), 612–617.
- (124) Lackner, M. TUNABLE DIODE LASER ABSORPTION SPECTROSCOPY (TDLAS) IN THE PROCESS INDUSTRIES – A REVIEW. *Rev. Chem. Eng.* 2007, 23 (2), 65–147.
- (125) Brückner, D. Innovating Quality Control Mechanisms in Aseptic Drug Manufacturing by Means of Isothermal Microcalorimetry and Tunable Diode Laser Absorption Spectroscopy, University Basel, 2017.
- (126) Cocola, L.; Allermann, H.; Fedel, M.; Sønderby, S.; Tondello, G.; Bardenstein, A.; Poletto, L. Validation of an In-Line Non-Destructive Headspace Oxygen Sensor. *Food Packag. Shelf Life* 2016, 9, 38–44.
-

- (127) Werle, P. A Review of Recent Advances in Semiconductor Laser Based Gas Monitors. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* 1998, *54* (2), 197–236.
- (128) Gordon, I. E.; Rothman, L. S. High-resolution transmission molecular absorption database (HITRAN) <https://hitran.org/>.
- (129) Sharma, P.; Kessler, W. J.; Bogner, R.; Thakur, M.; Pikal, M. J. Applications of the Tunable Diode Laser Absorption Spectroscopy: In-Process Estimation of Primary Drying Heterogeneity and Product Temperature During Lyophilization. *J. Pharm. Sci.* 2019, *108* (1), 416–430.
- (130) Wang, F.; Jia, S.; Wang, Y.; Tang, Z. Recent Developments in Modulation Spectroscopy for Methane Detection Based on Tunable Diode Laser. *Appl. Sci.* 2019, *9* (14), 2816.
- (131) Verma, R.; Neethu, S.; Kamble, S. S.; Radhakrishnan, J. K.; Krishnapur, P. P.; Padaki, V. C. Tunable Diode Laser Absorption Spectroscopy Based Oxygen Sensor. In *Sixth International Conference on Sensing Technology (ICST) Tunable*; 2012; pp 130–135.
- (132) Frish, M. B.; Laderer, M. C.; Wainner, R. T.; Wright, A. O.; Patel, A. H.; Stafford-Evans, J.; Morency, J. R.; Allen, M. G.; Green, B. D. The next Generation of TDLAS Analyzers. In *Next-Generation Spectroscopic Technologies*; Brown, C. D., Druy, M. A., Coates, J. P., Eds.; 2007; Vol. 6765.
- (133) Frish, M. B.; Wainner, R. T.; Laderer, M. C.; Green, B. D.; Allen, M. G. Standoff and Miniature Chemical Vapor Detectors Based on Tunable Diode Laser Absorption Spectroscopy. *IEEE Sens. J.* 2010, *10* (3), 639–646.
- (134) Cocola, L.; Fedel, M.; Allermann, H.; Landa, S.; Tondello, G.; Bardenstein, A.; Poletto, L. Validation and Calibration of a TDLAS Oxygen Sensor for In-Line Measurement on Flow-Packed Products. In *Promoting Objective and Measurable Food Quality & Safety*; Druy, M. A., Crocombe, R. A., Eds.; 2016.
- (135) Cai, T.; Wang, G.; Cao, Z.; Zhang, W.; Gao, X. Sensor for Headspace Pressure and H₂O Concentration Measurements in Closed Vials by Tunable Diode Laser Absorption Spectroscopy. *Opt. Lasers Eng.* 2014, *58*, 48–53.
- (136) Neethu, S.; Verma, R.; Kamble, S. S.; Radhakrishnan, J. K.; Krishnapur, P. P.; Padaki, V. C. Validation of Wavelength Modulation Spectroscopy Techniques for Oxygen Concentration Measurement. *Sensors Actuators B Chem.* 2014, *192*, 70–76.
- (137) Persson, L.; Lewander, M.; Andersson, M.; Svanberg, K.; Svanberg, S. Simultaneous Detection of Molecular Oxygen and Water Vapor in the Tissue Optical Window Using Tunable Diode Laser Spectroscopy. *Appl. Opt.* 2008, *47* (12), 2028.
- (138) Brown, H.; Mahler, H.-C.; Mellman, J.; Nieto, A.; Wagner, D.; Schaar, M.; Mathaes, R.; Kossinna, J.; Schmitting, F.; Dreher, S.; et al. Container Closure Integrity Testing—Practical Aspects and Approaches in the Pharmaceutical Industry. *PDA J. Pharm. Sci. Technol.* 2017, *71* (2), 147–162.
- (139) Brueckner, D.; Roesti, D.; Zuber, U. G.; Schmidt, R.; Kraehenbuehl, S.; Bonkat, G.; Braissant, O. Comparison of Tunable Diode Laser Absorption Spectroscopy and Isothermal Micro-Calorimetry for Non-Invasive Detection of Microbial Growth in Media Fills. *Sci. Rep.* 2016, *6* (27894), 1-9.

-
- (140) Duncan, D. Container Closure Integrity Testing of Finished Sterile Injectable Product. *Pharm. Technol.* 2019, *43* (9), 38–41.
- (141) Kanamori, H.; Momona, M.; Sakurai, K. Diode Laser Spectroscopy of the Atmospheric Oxygen Band. *Can. J. Phys.* 1990, *68* (3), 313–316.
- (142) Anderson, B.; Brecha, R. J. Tunable Diode Laser Absorption Measurement of Oxygen A-Band Line Strengths. *Appl. Phys. B* 2007, *87* (2), 379–385.
- (143) Laing, M. The Three Forms of Molecular Oxygen. *J. Chem. Educ.* 1989, *66* (6), 453.
- (144) Lou, X.; Somesfalean, G.; Chen, B.; Zhang, Z. Oxygen Measurement by Multimode Diode Lasers Employing Gas Correlation Spectroscopy. *Appl. Opt.* 2009, *48* (5), 990.
- (145) Weldon, V.; O’Gorman, J.; Pérez-Camacho, J. J.; Hegarty, J. Oxygen Sensing Using Single Frequency GaAs-AlGaAs DFB Laser Diodes and VCSELs. *Electron. Lett.* 1996, *32* (3), 219.
- (146) Brereton, R. G. *Applied Chemometrics for Scientists*, 1st Editio.; John Wiley & Sons, Ltd: Chichester, UK, 2007.
- (147) Luypaert, J.; Heuerding, S.; Massart, D. L.; Heyden, Y. Vander. Direct Orthogonal Signal Correction as Data Pretreatment in the Classification of Clinical Lots of Creams from near Infrared Spectroscopy Data. *Anal. Chim. Acta* 2007, *582* (1), 181–189.
- (148) Esbensen, K.; Geladi, P. The Start and Early History of Chemometrics: Selected Interviews. Part 2. *J. Chemom.* 1990, *4* (6), 389–412.
- (149) Shepherd, P. T. Retrospective. *J. Chemom.* 1987, *1* (1), 3–6.
- (150) Wold, S. Chemometrics; What Do We Mean with It, and What Do We Want from It? *Chemom. Intell. Lab. Syst.* 1995, *30* (1), 109–115.
- (151) Valcárcel, M. *Principios de Química Analítica*, 1st ed.; Springer-Verlag Ibérica S.A., Ed.; Springer-Verlag Ibérica, S.A., 1999.
- (152) Varmuza, K.; Filzmoser, P. *Introduction to Multivariate Statistical Analysis in Chemometrics*; CRC Press, 2016; Vol. 64.
- (153) Naes, T.; Isaksoon, T.; Fearn, T.; Davies, T. *A User-Friendly Guide to Multivariate Calibration and Classification*; IM Publications LLP, Ed.; NIR Publications, 2004.
- (154) Louangrath, P. I. Correlation Coefficient According to Data Classification. *SSRN Electron. J.* 2014, 1–28.
- (155) Gemperline, P. *Practical Guide To Chemometrics*, 2nd ed.; Gemperline, P., Ed.; CRC Press, 2006.
- (156) Lee, L. C.; Liong, C. Y.; Jemain, A. A. Partial Least Squares-Discriminant Analysis (PLS-DA) for Classification of High-Dimensional (HD) Data: A Review of Contemporary Practice Strategies and Knowledge Gaps. *Analyst* 2018, *143* (15), 3526–3539.
- (157) Bona, E.; Marçó, P. H.; Valderrama, P. Chemometrics Applied to Food Control. In *Food Control and Biosecurity*; Elsevier, 2018; Vol. 16, pp 105–133.
- (158) Prichard, G.; MacKay, G. M.; Points, J. Chapter 1. Achieving Valid Trace Analysis. In *Trace Analysis*; Prichard, E., MacKay, G. M., Points, J., Eds.; Royal Society of Chemistry: Cambridge, 1996; pp 1–39.
-

- (159) Beebe, K. R.; Pell, R. J.; Seasholtz, M. B. *Chemometrics: A Practical Guide*; John Wiley & Sons, Ltd, 1998.
- (160) Burgard, D. R.; Kuznicki, J. T. *Chemometrics. Chemical and Sensory Data*, 1st Editio.; CRC Press, 2017.
- (161) Brereton, R. G. *Chemometrics*; John Wiley & Sons, Ltd: Chichester, UK, 2003.
- (162) Blanco, M.; Cerdà, V. V.; Universitat de les Illes Balears. Servei de Publicacions i Intercanvi Científic. *Temas Avanzados de Quimiometría*, 1st ed.; Edicions UIB, 2007.
- (163) Food and Drug Administration. *Development and Submission of Near Infrared Analytical Procedures Guidance for Industry*; 2015.
- (164) Savitzky, A.; Golay, M. J. E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.* 1964, *36* (8), 1627–1639.
- (165) Barnes, R. J.; Dhanoa, M. S.; Lister, S. J. Standard Normal Variate Transformation and De-Trending of Near-Infrared Diffuse Reflectance Spectra. *Appl. Spectrosc.* 1989, *43* (5), 772–777.
- (166) Blanco, M.; Coello, J.; Montoliu, I.; Romero, M. . Orthogonal Signal Correction in near Infrared Calibration. *Anal. Chim. Acta* 2001, *434* (1), 125–132.
- (167) Forina, M.; Lanteri, S.; Casale, M. Multivariate Calibration. *J. Chromatogr. A* 2007, *1158* (1–2), 61–93.
- (168) Hotelling, H. The Generalization of Student's Ratio. *Ann. Math. Stat.* 1931, *2* (3), 360–378.
- (169) Kennard, R. W.; Stone, L. A. Computer Aided Design of Experiments. *Technometrics* 1969, *11* (1), 137–148.
- (170) Stone, M. Cross-Validatory Choice and Assessment of Statistical Predictions. *J. R. Stat. Soc. Ser. B* 1974, *36* (2), 111–133.
- (171) Wise, B. M.; Roginski, R. T. A Calibration Model Maintenance Roadmap. *IFAC-PapersOnLine* 2015, *48* (8), 260–265.
- (172) Mercader, M. B.; Puigdomènech, A. R. Near Infrared Multivariate Model Maintenance: The Cornerstone of Success. *NIR news*. 2014, *25* (7), 7–9.
- (173) Peris-Vicente, J.; Esteve-Romero, J.; Carda-Broch, S. Validation of Analytical Methods Based on Chromatographic Techniques: An Overview. In *Analytical Separation Science*; Anderson, J. L., Berthod, A., Pino-Estévez, V., Stalcup, A. M., Eds.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2015; pp 1757–1808.
- (174) Hoffmann, A.; Kähny-Simonius, J.; Plattner, M.; Schmidli-Vckovski, V.; Kronseder, C. Computer System Validation: An Overview of Official Requirements and Standards. *Pharm. Acta Helv.* 1998, *72* (6), 317–325.
- (175) Food and Drug Administration. *Analytical Procedures and Methods Validation for Drugs and Biologics Guidance for Industry*; 2015.
- (176) The Internacional Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. *ICH Q2 (R1) Validation of Analytical Procedures: Text and Methodology*; 2011; Vol. 20.

-
- (177) Bretnall, A. E.; Clarke, G. S. Validation of Analytical Test Methods. In *Separation Science and Technology*; Elsevier Inc., 2011; Vol. 10, pp 429–457.
- (178) Indrayanto, G. Validation of Analytical Methods—Update 2011. In *Profiles of Drug Substances, Excipients and Related Methodology*; Elsevier Inc., 2012; Vol. 37, pp 439–465.
- (179) Mohammad, A. S.; Devidi, S.; Fatima, N.; Badar, H.; Sulthana, S. S.; Sulthana, M. A.; Rasheed, N. An Overview of Validation and Basic Concepts of Process Validation: Quality Assurance View Point. *Asian J. Pharm. Technol.* 2016, 6 (3), 169.
- (180) Lal, B.; Kapoor, D.; Jaimini, M. A Review on Analytical Method Validation and Its Regulatory Perspectives. *J. Drug Deliv. Ther.* 2019, 9(2), 501–506.
- (181) Araujo, P. Key Aspects of Analytical Method Validation and Linearity Evaluation. *J. Chromatogr. B* 2009, 877(23), 2224–2234.
- (182) McPolin, O. *Method Validation in Pharmaceutical Analysis*, 1st ed.; Ermer, J., McB. Miller, J. H., Eds.; Wiley, 2005.
- (183) The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. *Final Concept Paper ICH Q14: Analytical Procedure Development and Revision of Q2(R1) Analytical Validation Dated 14 November 2018*; 2018; Vol. 2.
- (184) Broad, N.; Graham, P.; Hailey, P.; Hardy, A.; Holland, S.; Hughes, S.; Lee, D.; Prebble, K.; Salton, N.; Warren, P.; et al. Guidelines for the Development and Validation of Near-Infrared Spectroscopic Methods in the Pharmaceutical Industry. In *Handbook of Vibrational Spectroscopy*; Chalmers, J. M., Ed.; John Wiley & Sons, Ltd: Chichester, UK, 2006; pp 3590–3610.

II. OBJETIVOS

1. **OBJETIVOS**

El objetivo principal de esta tesis consiste en ofrecer soluciones analíticas que puedan mejorar la calidad del proceso de fabricación o resolver problemas que puedan aparecer a lo largo de la vida útil de distintos productos farmacéuticos liofilizados.

Cada uno de los temas estudiados se ha abordado de una manera diferente dependiendo de la técnica analítica empleada (destrucciona o no destructiva), del tipo de respuesta (univariante o multivariante), y del uso de la respuesta final (cuantitativa o cualitativa).

Los objetivos concretos que se pretende alcanzar con los estudios realizados son los siguientes:

- Estudio del impacto de la humedad contenida en los tapones, sobre la degradación de un producto liofilizado. Para ello, se desarrollará y validará un método analítico capaz de determinar el contenido de agua de los tapones de clorobutilo utilizados como envase primario. Este método se aplicará para monitorizar el impacto de esta humedad en un producto liofilizado hidrolizable, almacenado durante 6 meses en condiciones de estabilidad acelerada.
- Cuantificación de la humedad residual, a liberación, de un producto liofilizado utilizando la espectroscopia NIR. Se ha elegido utilizar la espectroscopia NIR con el fin de realizar un análisis no destructivo, por lo que será necesario realizar un modelo de calibración multivariable. Previamente a la realización del modelo de calibración, se desarrollará y validará un método de referencia para la determinación de trazas de humedad en el producto. También se realizará un estudio de viabilidad utilizando el análisis modal de fallos y efectos. Finalmente, el método multivariable será validado según las recomendaciones de las agencias reguladoras.
- Determinación de oxígeno en el *headspace* a partir del espectro visible-NIR de un producto liofilizado que puede desarrollar coloración debido a la presencia de oxígeno. Desarrollo de un método de análisis basado en el espectro visible-NIR adquirido para la determinación de la humedad residual del producto, que permita identificar aquellas muestras que puedan presentar una mayor coloración con el tiempo. Una vez desarrollado el método deberá ser validado siguiendo las recomendaciones de las agencias reguladoras.

OBJETIVOS

- Prueba piloto del desarrollo de un método cualitativo para la identificación, en el propio envase, de los productos farmacéuticos liofilizados antes de su liberación al mercado. Se creará y validará una biblioteca de espectros NIR de los distintos productos que permita su identificación mediante un método multivariante cualitativo y no destructivo. Dependiendo de los resultados obtenidos se decidirá la viabilidad de la aplicación de este método en la industria.

III. REACTIVOS Y MATERIALES

1. REACTIVOS

En esta tesis, los reactivos que se han utilizado se centran en la determinación del agua por el método de Karl Fischer.

- Valoración volumétrica Karl Fischer: se ha utilizado como solvente el Aquagent® Solvent, sin piridina, de Scharlab (Barcelona, España). Como reactivo valorante se ha utilizado Aquagent® Titrant 5 o el Aquagent® Titran 2, ambos sin piridina, y del mismo proveedor que el solvente.
- Valoración coulombimétrica Karl Fischer: para este tipo de valoraciones solamente hace falta un solo reactivo, en este caso se ha utilizado el HYDRANAL™ Coulomat AG para células con y sin diafragma de Honeywell Fluka™ (Carolina del Norte, EE. UU.).

2. PATRONES

Tanto para realizar las verificaciones de los equipos (semanales o diarias), como para validar los métodos desarrollados durante esta tesis, se han utilizado tres patrones de agua distintos:

- HYDRANAL™ – patrón de agua 1.0 (Honeywell Fluka™, Carolina del Norte, EE. UU.), patrón líquido que contiene 1 mg/g de agua, se utiliza en solución y su uso a temperaturas elevadas no está recomendado.
- Patrón de agua para horno 1 % Aquastar® (Merck, Darmstadt, Alemania), patrón sólido para su uso en horno en el intervalo de temperaturas de 140°C a 400°C.
- HYDRANAL™ – patrón de agua para horno (Honeywell Fluka™, Carolina del Norte, EE. UU.). Patrón sólido apto para su uso con horno en un intervalo de temperaturas desde 140°C hasta 160°C.

3. INSTRUMENTOS

En las siguientes secciones se describen los instrumentos, tanto de escala laboratorio como industrial, utilizados durante la realización de la tesis.

3.1. Espectrofotómetro NIR

Todos los espectros en el infrarrojo cercano registrados durante el desarrollo de la tesis se han realizado utilizando un espectrofotómetro LabSpec 5000 VIS/NIR de ASD Inc® (Colorado, EE. UU.) equipado con el accesorio para muestras High Intensity Muglight y sonda de fibra óptica.



Figura III - 1. Imagen del espectrofotómetro LabSpec 5000 VIS/NIR equipado con el accesorio para muestras High Intensity Muglight.

Este equipo, junto con el accesorio, permite el registro de los espectros en modo de reflectancia en el intervalo de longitudes de onda entre 350 nm y 2500 nm. El detector en la región 350 – 1000 nm consiste en una matriz de fotodiodos de silicio de 512 canales; en las regiones de 1000 – 1800 nm y entre 1800 – 2500 nm, los detectores son de diodos de InGaAs con refrigeración termo-eléctrica.

La resolución espectral es de 3 nm en la región del visible y de 10 nm en la región del NIR; cada espectro es el promedio de 32 barridos registrados cada 100 ms.

Teniendo en cuenta el ámbito de aplicación de las medidas, el equipo está cualificado siguiendo las normas de correcta fabricación (GMP, siglas del inglés *Good Manufacturing Practice*) y antes de proceder con el registro de las muestras el equipo se verifica según la farmacopea europea *Ph.Eur. 2.2.40 Near-Infrared Spectroscopy*.¹

3.2. Liofilizadores

Las muestras de rutina analizadas en el desarrollo de los distintos capítulos experimentales pueden proceder de seis liofilizadores distintos, cinco distribuidos en dos zonas de producción, denominadas zona estéril 1 y 2, y otro en el laboratorio piloto. De estos cinco liofilizadores, tres tienen una mayor capacidad del condensador, por lo que pueden liofilizar lotes de mayor tamaño y son del proveedor IMA Life (Italia). De este proveedor se tienen tres modelos distintos resumidos en la Tabla III - 1. Los otros dos

liofilizadores restantes, tienen una menor capacidad del condensador y son del proveedor Leybold (Alemania) (Tabla III - 1).

Tabla III - 1. Resumen de los distintos modelos de los liofilizadores utilizados indicando su capacidad del condensador y la zona de producción donde se encuentran.

Modelo	Abreviación	Capacidad del condensador	Zona de producción
Lyofast 15	Edw 3	263 kg	1
Lyofast 25	Edw 2	525 kg	2
Lyomax 22	Edw 1	600 kg	2
GT-100	-	100 kg	1
GT-150	-	150 kg	1
Lyobeta 20	LB20	30 kg	Laboratorio

Para poder obtener muestras con distintas humedades y producir lotes a pequeña escala con características diferentes a las obtenidas en rutina, se ha empleado el liofilizador piloto Lyobeta 20 de Azbil Telstar (Terrassa, España). Este equipo cuenta con un extractor de muestras (traducción del inglés *sample thief*) acoplado a la puerta, que permite extraer muestras durante el ciclo de liofilización sin alterar el ciclo de las demás muestras del liofilizador.



Figura III - 2. Imagen del extractor de muestras acoplado al liofilizador Lyobeta 20, en las instalaciones de los Laboratorios Reig Jofre de Barcelona.

3.3. Valorador Karl Fischer

Todos los valores de referencia de humedad residual se han obtenido utilizando el valorador Karl Fischer 852 Titrande (Metrohm AG, Suiza) y, en caso de realizar las

medidas con horno, se ha acoplado al instrumento el procesador de muestras 874 Oven Sample Processor (Metrohm AG, Suiza).



Figura III - 3. Imagen del instrumento 852 Titrando y del 874 Oven Sample Processor. Imágenes extraídas de la página web del proveedor.²

El instrumento de valoración Karl Fischer 852 Titrando contiene dos interfaces de medida, una dedicada a la valoración volumétrica con un electrodo polarizable y la otra dedicada a la valoración coulombimétrica con un electrodo generador sin membrana y un electrodo de platino doble. Además, cuenta con dos módulos de agitación, un módulo de dosificación, el cual dosifica la solución valorante a la interfaz de medida de la valoración volumétrica y el 874 Oven Sample Processor, el cual permite precalentar las muestras para que éstas desprendan la humedad y así poder valorarla con la interfaz de medida de la titulación coulombimétrica.

El procesador de muestras 874 Oven Sample Processor contiene los siguientes componentes:

- Horno: módulo de horno de aluminio para calentar el recipiente donde se halla la muestra; la temperatura se regula a partir del software asociado.
- Ventilador de hélice: permite refrigerar el módulo de horno después de su uso.
- Bomba de aire: este componente genera el flujo de gas (en este caso nitrógeno) que permite transportar el agua liberada por las muestras hasta la celda de valoración.
- Aguja: para poder insertar el flujo de gas dentro de la muestra, este procesador cuenta con una aguja, un dispositivo de guía y tuberías para el flujo de entrada y salida del gas.
- Plato giratorio: con el fin de automatizar este procedimiento, este procesador cuenta con una gradilla de muestras estándar con una capacidad de 36 viales, que siguiendo la secuencia programada con el software analiza las muestras de la gradilla.

Teniendo en cuenta el ámbito de aplicación de los resultados, el equipo está cualificado siguiendo las GMP, la idoneidad del sistema se verifica semanalmente para la valoración volumétrica según la *Ph.Eur. 2.5.12. Water: semi-micro determination* y diariamente para la coulombimétrica la *Ph.Eur. 2.5.32. Water: micro determination*.^{3,4}

3.4. Espectrofotómetro de absorción láser de diodo sintonizable para la determinación de oxígeno

Todos los valores de referencia del contenido de oxígeno del *headspace* en los productos estudiados durante la realización de la tesis, se han realizado utilizando un espectrofotómetro de absorción láser de diodo sintonizable Headspace analyzer FMS-Oxygen de Lighthouse Instruments, LLC. (Virginia, EE. UU.).

Este equipo permite registrar, en modo de transmitancia, el *headspace* de las muestras, obteniendo un valor de la concentración de oxígeno en %. La configuración del espectrofotómetro consiste en un láser de diodo sintonizable centrado a la longitud de onda de 760 nm, aproximadamente.

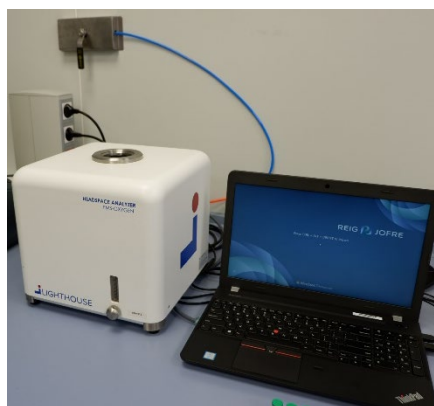


Figura III - 4. Imagen del espectrofotómetro de absorción láser de diodo sintonizable Headspace analyzer FMS Oxygen, acoplado al ordenador por donde se controla y donde se puede ver la toma de nitrógeno.

El equipo está cualificado siguiendo las GMP y antes de proceder con el registro de las muestras se debe centrar la señal del láser y establecer la recta de calibración del equipo. Para poder realizar ambas acciones, se requiere el uso de patrones certificados para cada formato de vial producidos y certificados por el mismo proveedor del equipo Lighthouse Instruments, LLC. (Virginia, EE. UU.).



Figura III - 5. Imagen de los patrones de tres formatos de viales distintos de las 6 concentraciones necesarias para realizar el centrado de la señal y establecer la recta de calibración (0 %, 1 %, 2 %, 4 %, 8 % y 25 %).

3.5. Espectrofotómetro UV-Visible

Los resultados de la absorción de la solución del producto, para el control de su coloración, se han obtenido utilizando un espectrofotómetro UV-Visible LAMBDA 25 (PerkinElmer Inc., Massachusetts, EE. UU.).

Este equipo permite el registro en modo de transmitancia, obteniendo un espectro comprendido en el intervalo de longitudes de onda entre 200 – 900 nm. El espectrofotómetro tiene una lámpara de deuterio y otra de halógeno, con una rejilla holográfica cóncava y un detector de fotodiodos de silicio.

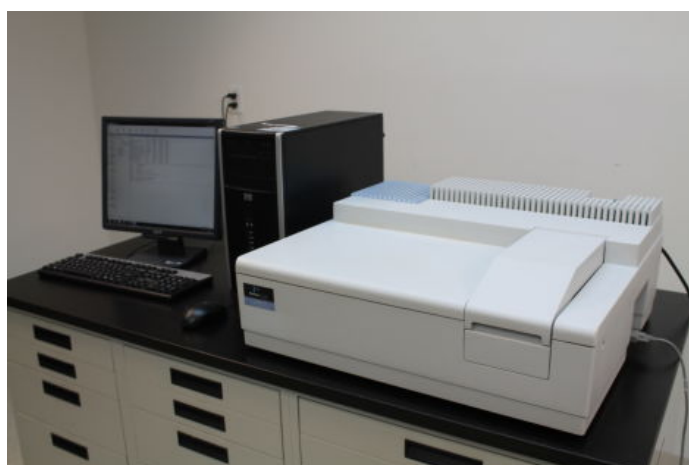


Figura III - 6. Imagen del UV-Visible LAMBDA 25 utilizado.

3.6. Otros instrumentos

Otros instrumentos utilizados para el desarrollo de esta tesis han sido:

- Difractómetro de polvo X'Pert Powder de Malvern Panalytical (Almelo, Países Bajos y Malvern, Reino Unido), con geometría theta-theta, ánodo de cobre y detector PIXcel^{1D}. Los análisis realizados durante el desarrollo de esta tesis para determinar la estructura de algunas muestras, se ha realizado en colaboración con el servicio de rayos X de la Universitat Autònoma de Barcelona (Cerdanyola del Vallès, España).⁵
- Cabina de humedad controlada del proveedor Enext (Barcelona, España). Teniendo en cuenta el alto grado de higroscopicidad de los productos liofilizados, en muchas ocasiones debe trabajarse en un espacio con una atmósfera de humedad controlada. Durante el desarrollo de esta tesis se ha utilizado la cabina que se muestra en la Figura III - 7. Esta cabina está equipada con tres guantes que permiten trabajar con el producto sin interferir en las condiciones de humedad, además de con una balanza y un sinóptico que permite marcar la consigna de porcentaje de humedad a la cual se quiere trabajar.



Figura III - 7. Imagen de la cabina de humedad controlada utilizada en el desarrollo de la tesis.

- Autoclave modelo 700 de Azbil Telstar (Terrassa, España). Este equipo industrial, situado en la zona de producción 1 de la compañía, ha permitido esterilizar los tapones utilizados para el estudio del impacto de la humedad contenida en los tapones de clorobutilo en el producto liofilizado.
- Cámara climática 40°C-75 % de humedad relativa de Ingeniería técnica de mantenimiento, S.A. (Barcelona, España). Esta cámara climática ofrece la posibilidad de mantener los productos una vez liofilizados a unas condiciones

climáticas controladas. La variación de temperatura puede ser de hasta $\pm 2^{\circ}\text{C}$ mientras que para la humedad relativa se acepta un margen de trabajo de $\pm 5\%$.

4. PROGRAMAS PARA LA ADQUISICIÓN Y EL TRATAMIENTO DE LOS DATOS

En las siguientes secciones se indican los programas empleados para la adquisición y el tratamiento de los datos durante el desarrollo de esta tesis.

4.1. Programas para la adquisición de los datos

- IndicoPro™ versión 6 (ASD Inc[®], Colorado, EE. UU.), programa utilizado en el espectrofotómetro LabSpec 5000 VIS/NIR de ASD Inc[®] para el registro y la visualización de los espectros.
- Lyosyutelab2 (Telstar, España), programa utilizado en el liofilizador piloto Lyobeta 20 de Telstar para el control y el desarrollo de ciclos de liofilización a escala piloto.
- Tiamo™ versión 2.4 (Metrohm AG, Suiza), programa utilizado para gestionar el registro, uso y almacenamiento de los valores de humedad de las muestras analizadas con los instrumentos 852 Titrand y 874 Oven Sample Processor.
- Lighthouse FMS-Oxygen Analyser (Lighthouse Instruments, Virginia, EE. UU.), programa utilizado en el Headspace analyzer FMS-Oxygen para registrar y almacenar los valores de oxígeno del *headspace* de los productos analizados.
- UV WinLab versión 6 (PerkinElmer Inc., Massachusetts, EE. UU.), programa utilizado en el espectrofotómetro UV-Visible LAMBDA 25 de PerkinElmer Inc., para el registro y la visualización de los espectros.

4.2. Programas para el tratamiento de los datos

- Unscrambler[®] 10.5 (CAMO Analytics, Noruega), este programa incorpora un amplio abanico de herramientas y pretratamientos quimiométricos útiles para el desarrollo de modelos multivariantes cuantitativos y cualitativos.
- OPUS versión 6.5 (Bruker Optics Inc., Massachusetts, EE. UU.), este programa incorpora una serie de herramientas quimiométricas empleadas para el desarrollo de modelos multivariantes cualitativos durante esta tesis.

- MATLAB® R2014a (MathWorks®, Massachusetts, EE. UU.), este programa incorpora la posibilidad de programar distintas aplicaciones y se ha utilizado para el cálculo del límite de detección.
- Microsoft Excel 2016 (Microsoft Corporation, Washington, EE. UU.), este programa se ha utilizado para realizar cálculos de estadística.
- Minitab® versión 17 (Minitab Ltd., Pensilvania, EE. UU.), este programa se ha utilizado para realizar cálculos estadísticos.

5. BIBLIOGRAFÍA

- (1) European Pharmacopoeia. *2.2.40. Near-Infrared Spectroscopy*; 2019.
- (2) Metrom AG. Metrohm AG <https://www.metrohm.com/en/products/karl-fischer-titration/>.
- (3) European Pharmacopoeia. 2.5.32. Water Micro Determination. 2019, 20532.
- (4) European Pharmacopoeia. 2.5.12. Water Semi-Micro Determination. 2018, 20512.
- (5) Universitat Autònoma de Barcelona. Servei Difracció Raigs X <http://sct.uab.cat/sdrx/es>.

**IV. IMPACTO DE LA HUMEDAD DEL
TAPÓN DEL VIAL EN EL PRODUCTO
TERMINADO**

1. INTRODUCCIÓN

Como se ha comentado anteriormente, el contenido de agua en un producto farmacéutico liofilizado es un atributo de calidad, que se debe mantener bajo control durante toda la vida útil del producto. El valor inicial de este CQA depende de que se desarrolle un correcto ciclo de liofilización que consiga obtener valores de humedad bajos y homogéneos, tanto en la pastilla del producto como en el *headspace* de éste. Para mantener este valor invariable a lo largo del tiempo, se debe escoger el envase y las condiciones de almacenaje adecuadas, siendo el primer factor el que tiene un mayor impacto.¹

El envase de un producto parenteral liofilizado consiste en un vial de vidrio cerrado con un tapón de goma, que tiene la función de proteger el producto de agentes externos (microorganismos, aire, humedad, luz, ...) que puedan alterar su calidad antes de su uso.²

Según la información publicada, las tres principales fuentes de humedad que pueden afectar a la variabilidad de esta durante el tiempo de vida útil del producto son:^{3,4}

1. Fugas en la interfaz del cierre entre el vial y el tapón.
2. Permeabilidad de la humedad ambiental a través del tapón.
3. Transferencia de la humedad contenida en el tapón al producto.

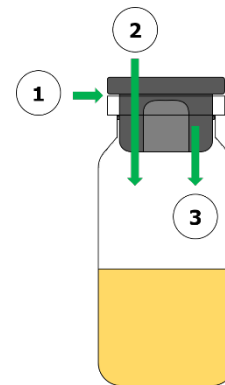


Figura IV - 1. Esquema de las principales fuentes de variabilidad de humedad.

Sin embargo, como se comenta en varios estudios, las dos primeras causas (1 y 2 de la Figura IV - 1) carecen de importancia frente a la última.⁵

El agua contenida en los tapones se debe al proceso de esterilización al que son sometidos antes de su uso. Aunque estos también se pueden esterilizar por radiación, es mucho más habitual tratarlos con vapor de agua en autoclaves, los cuales utilizan ciclos en los que se varía la temperatura y la presión, para conseguir esterilizar el

material. Este proceso conlleva una etapa posterior de secado para eliminar o reducir la humedad introducida en el tapón.^{6,7}

En este capítulo de la tesis, el objetivo consiste en estudiar el impacto de la humedad contenida en los tapones sobre un producto liofilizado hidrolizable, el cual ha sido almacenado durante 6 meses en condiciones de estudio de estabilidad acelerado. Para ello, ha sido necesario desarrollar y validar, con la rigurosidad exigida en el entorno farmacéutico, un método analítico apto para la determinación de la humedad residual en los tapones.

2. MATERIALES Y MÉTODOS EXPERIMENTALES

2.1. Muestras

2.1.1. Tapones

Los tapones utilizados en productos parenterales liofilizados tienen una forma que les permite encajar con en el cuello del vial, cerrándolo completamente. A la vez, contienen unos soportes laterales que les permiten estabilizarse en una posición semicerrada (pretaponada) y, que permite la evacuación del vapor de agua del vial a través de una abertura durante el ciclo de liofilización debida a la sublimación.

En este estudio, se han empleado tapones grises de clorobutilo de 13 mm de diámetro, del proveedor West Pharmaceutical Services, Inc. (Pensilvania, EE. UU.).

Estos tapones han sido esterilizados utilizando el autoclave 500, descrito anteriormente. El ciclo empleado consiste en una primera etapa de esterilización, a una temperatura de 121,5°C, seguida de una etapa de secado. En ésta, se realizan 20 pulsos de presión de 5 min/pulso, en los que se varia la presión desde 100 mbar hasta 900 mbar. Posteriormente, se dejan los tapones, ya fuera del autoclave, bajo flujo laminar durante un mínimo de 8 h.

2.1.2. Muestras del producto H

El producto H consiste en un producto parenteral liofilizado que contiene 2 mg de API. También incluye un excipiente empleado para aportar forma al producto (excipiente δ) y otro para ajustar el pH (excipiente ϵ).

Este producto se caracteriza por tener un contenido en sólidos muy bajo, lo que se traduce en una pastilla muy porosa y frágil. Se encuentra dosificado en un vial de tubo incoloro de 3 ml, cerrado con el tapón anteriormente descrito y capsulado con una cápsula de aluminio con tapa. La selección de este producto se debe a que, además de tener una alta higroscopicidad, como la mayoría de liofilizados, presenta dos reacciones de degradación por hidrólisis (más información en la sección 3.3). Estas degradaciones ayudarán a estudiar el impacto de la humedad del tapón en el producto.

Para este estudio se ha utilizado solución de fábrica (escala industrial), aunque ha sido liofilizada en la planta piloto en el liofilizador Lyobeta 20, con el fin de poder modificar los tapones de los productos sin afectar a la producción comercial.

Una vez liofilizadas, las muestras se han guardado en una cámara de temperatura y humedad controladas, a 40°C ($\pm 2^\circ\text{C}$) y 75 % (± 5 %), respectivamente.

2.2. Métodos experimentales

2.2.1. Determinación de la humedad en tapones de clorobutilo de 13 mm

La determinación de la humedad se realiza con una valoración coulombimétrica acoplada a un horno. Todos los detalles del método se encuentran descritos en la Tabla IV - 1.

Después de verificar la celda de valoración, se analizan tres blancos. El primero sirve para acondicionar todo el sistema, eliminando trazas de agua. Los otros dos permiten determinar la humedad propia de la preparación de muestras y del equipo; el promedio de su valor se sustrae de las siguientes determinaciones. A continuación, se analizan las muestras intercalando un patrón, de aproximadamente 1500 μg de agua, cada tres muestras, para verificar la fiabilidad de los resultados. El porcentaje de recuperación de estos patrones se debe encontrar dentro del intervalo de 97,5-102,5 %.⁸

La preparación de las muestras consiste en cortar el tapón en 12 trozos, de aproximadamente el mismo tamaño, e introducirlos en un vial de 6 ml. Dado el alto contenido de humedad en los tapones, se debe llenar la celda de valoración con, aproximadamente, 200 ml del medio de la titulación (Columat) y debe cambiarse después de unas 30 determinaciones para evitar colapsarlo.

2.2.2. Determinación de la humedad residual del producto H

La determinación de la humedad residual del producto H se realiza con la técnica de valoración Karl Fischer volumétrica, empleando un reactivo valorante de una concentración menor (Titrant 2) que el reactivo habitual (Titrant 5), consiguiendo así una mayor precisión en el análisis. Para ello, se elabora una mezcla de 6 viales y se realizan dos análisis, pesando 100 mg en cada caso, dando el valor medio de ambos como valor final.

2.2.3. Determinación de la concentración y las impurezas del producto H

La valoración cuantitativa del API y la determinación de impurezas del producto H se realiza por HPLC, con un método desarrollado y validado en la empresa por el departamento de investigación y desarrollo.

Todos los análisis se han realizado en el departamento de control de calidad de la empresa Reig Jofre, siguiendo lo establecido en el procedimiento analítico.⁹

Los análisis se han realizado por duplicado, mezclando 8 viales para cada determinación. El valor medio de ambos análisis es el presentado como valor final.

3. RESULTADOS Y DISCUSIÓN

3.1. Desarrollo del método

Tal y como se ha descrito en secciones anteriores, uno de los objetivos de este capítulo es desarrollar un método analítico capaz de determinar el agua libre que se encuentra en el tapón y que puede ser transferida al producto liofilizado, afectando a su calidad durante su vida útil. En las siguientes secciones se detallan los pasos del método desarrollado.

3.1.1. Selección de la técnica

La selección de la técnica analítica se realizó teniendo en cuenta las características de la muestra y la disponibilidad instrumental. Al ser una muestra insoluble, el método volumétrico de Karl Fischer se descartó. La alternativa, utilizar un método de secado, era imposible debido a la carencia de los equipos adecuados, sin entrar en

consideraciones sobre su falta de especificidad en el caso de muestras que contienen aditivos orgánicos, como los tapones. Por consiguiente, se decidió utilizar el equipo de valoración coulombimétrica de Karl Fischer acoplado a un horno.

Aunque la técnica está diseñada para análisis de bajo contenido en agua, la utilización de un tercio más de reactivo y su sustitución cada, aproximadamente, 30 análisis, permite su utilización en la determinación de una cantidad de agua no habitual para la técnica.

3.1.2. Preparativa de la muestra

Para determinar el contenido de humedad, se corta el tapón de 13 mm en 12 partes, siguiendo el esquema descrito en la Figura IV - 2. Este procedimiento facilita la liberación del agua, reduciendo la temperatura y el tiempo de análisis.^{1,10} La preparativa de la muestra puede realizarse en el mismo laboratorio sin necesidad de usar una cabina de humedad controlada.

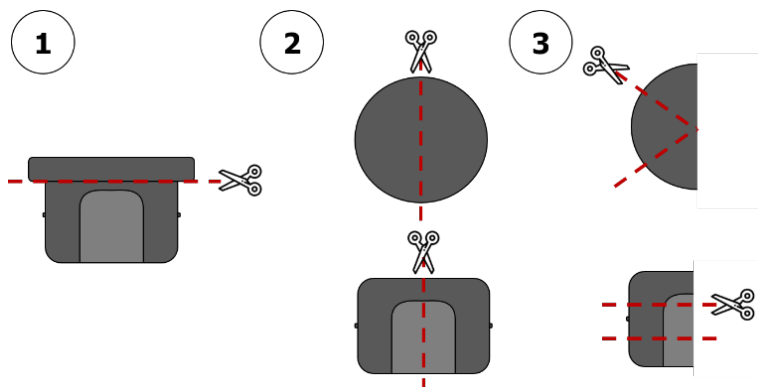


Figura IV - 2. Esquema del procedimiento de como cortar un tapón de 13 mm en 12 partes.

3.1.3. Selección de los parámetros del método

Para la mayoría de los parámetros del método se han mantenido los valores recomendados por el proveedor del equipo; en la Tabla IV - 1 se pueden ver con más detalle.

Tabla IV - 1. Resumen de los principales parámetros del método para determinar la humedad en los tapones de clorobutilo 13 mm.

Técnica	Karl Fischer culombimétrico acoplado a un horno	Temperatura del horno	145°C
---------	---	--------------------------	-------

IMPACTO DE LA HUMEDAD DEL TAPÓN DEL VIAL EN EL PRODUCTO TERMINADO

Tiempo de estabilización	120 segundos	Intervalo de recuperación del patrón	97,5-102,5 %
Deriva inicial	< 20 µg/min	Intensidad del generador	400 mA
Criterio de parada	Deriva relativa (20 µg/min)	Corriente de polarización	10 µA
Tiempo de extracción de la muestra	1800 segundos	Flujo de nitrógeno	85 ml/min

Sin embargo, considerando las características del producto a analizar, se modificaron los parámetros de temperatura, flujo de nitrógeno y tiempo de extracción.

- Flujo de nitrógeno: este parámetro define el caudal del flujo del gas portador. En este caso, teniendo en cuenta la experiencia previa de la compañía, se utiliza un flujo de 85 ml/min.
- Tiempo de extracción: este parámetro define el tiempo mínimo que la muestra debe estar analizándose, independientemente de que se alcance el criterio de parada. Se utiliza para muestras que no liberan la humedad inmediatamente, sino que requieren un cierto tiempo. En este caso, dada la naturaleza de la muestra, se propone como valor inicial 1800 segundos.
- Temperatura: el objetivo de este método es determinar el agua que puede captar la muestra liofilizada proveniente del tapón, no el agua total contenida en el mismo. Tal y como se detalla en la bibliografía, para determinar la totalidad del agua en un tapón se debe seleccionar una temperatura superior a 300°C. Sin embargo, para determinar el agua libre contenida en él se deben seleccionar temperaturas entre 100°C y 150°C.^{1,7,10}

En este caso, se ha seleccionado la temperatura de 145°C, debido a que es la temperatura mínima de trabajo aconsejada para el patrón de agua de que se dispone. Tener una temperatura similar entre las muestras y el patrón reduce el tiempo de análisis, puesto que no se requiere un tiempo de enfriamiento del horno.

Este método es una propuesta, por lo que para darlo por bueno se debe validar. En el caso de que los resultados de la validación no cumplan con los requisitos descritos en el protocolo de validación, se debe volver a desarrollar el método para mejorarlo.

3.2. Validación del método

Antes de empezar con la validación, y con la finalidad de determinar el intervalo de concentración de agua que se va a validar, se realiza el procedimiento descrito en las secciones anteriores con muestras de distintas concentraciones de humedad. Para determinar el contenido habitual de humedad en rutina (valor central), se han analizado 18 muestras que han seguido el ciclo de esterilización descrito en la sección (2.1.1). Para el contenido máximo de humedad (límite superior) se utilizaron 18 muestras sólo con el ciclo de esterilización, sin realizar la posterior exposición bajo el flujo laminar, y para el límite inferior de humedad se analizaron 18 muestras que, después del ciclo de esterilización, se han secado durante 48 h en una estufa a 105°C.

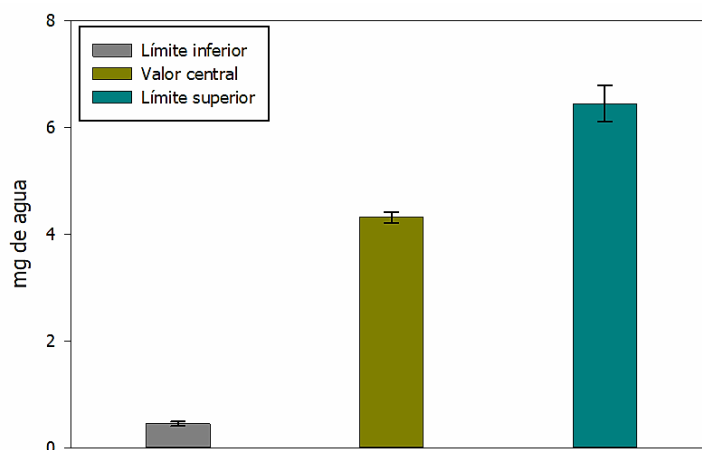


Figura IV - 3. Representación de los resultados de los mg de agua encontrados en las 18 muestras analizadas para cada uno de los tres límites estudiados.

Tal y como se puede observar en la Figura IV - 3, el contenido de agua en los tapones se encuentra entre 7,5 mg y 0,3 mg, con un valor en rutina antes de liofilizar de alrededor de 4,3 mg.

3.2.1. Especificidad

Para definir la especificidad del método, en este caso, se debe asegurar que no hay reacciones secundarias. Teniendo en cuenta que la muestra no entra en contacto con el medio de valoración, se evitan las posibles reacciones secundarias entre ambos. Además, a 145°C, como ya se ha comentado, no se produce la degradación de los materiales que forman parte del tapón, por lo que a la celda de valoración sólo llega agua. Esto significa que el método es específico para determinar este compuesto en los tapones.¹⁰

3.2.2. Exactitud

Para estudiar la exactitud del método, se han utilizado tapones del límite inferior, debido a que éstos presentan un contenido de humedad más homogéneo entre muestras (Figura IV - 3). A continuación, siguiendo el procedimiento descrito en las secciones anteriores, se han cortado los tapones en 12 partes (Figura IV - 2), mezclando los trozos para realizar una mezcla homogénea. De esta mezcla se han separado aleatoriamente 6 trozos de la parte superior y de la parte inferior para confeccionar lo que llamaremos muestra. Estas muestras han sido dopadas con una cantidad de patrón de agua conocida para tener tres niveles distintos de concentración. Estos niveles se han basado en los análisis realizados durante el desarrollo del método.

Tabla IV - 2. Resultados del porcentaje de recuperación (R) de los tres replicados (Rep.) de los niveles estudiados. Donde μ_{gT} son los microgramos de agua teóricos y μ_{gE} los microgramos de agua encontrados experimentalmente.

Rep.	Límite inferior			Valor central			Límite superior		
	μ_{gT}	μ_{gE}	R (%)	μ_{gT}	μ_{gE}	R (%)	μ_{gT}	μ_{gE}	R (%)
1	470,0	444,2	94,5	4599,0	4687,5	101,9	6593,4	6673,7	101,2
2	585,6	610,9	104,3	4442,6	4481,9	100,9	6574,5	6617,2	100,6
3	483,3	504,1	104,3	4534,6	4602,7	101,5	6586,3	6683,4	101,5

Tal y como se puede observar en Tabla IV - 2, las recuperaciones para todos los niveles cumplen con los criterios establecidos en la Ph.Eur., 90-110 % para el nivel inferior, y de 97,5-102,5 % para los niveles con un mayor contenido en agua.⁸ Con esto se confirma que el método es exacto en todo el intervalo de concentraciones estudiado.

3.2.3. Linealidad

Para estudiar la linealidad del método se ha seguido el mismo procedimiento comentado en el estudio de la exactitud, dopando las muestras con una cantidad de estándar conocida, cubriendo todo el intervalo de concentraciones que se quiere validar.

Los tres niveles estudiados en la exactitud se han usado también para el estudio de la linealidad del método. Además, se han añadido otros dos niveles, realizados también por triplicado, entre el límite inferior y el valor central (800 μg y 1000 μg de agua).

Tabla IV - 3. Resultados de los niveles para el estudio de la linealidad. μg_T son los microgramos de agua teóricos y μg_E los microgramos de agua encontrados experimentalmente.

Replicados	800 μg de agua		1000 μg de agua	
	μg_T	μg_E	μg_T	μg_E
1	831,4	869,7	1156,7	1188,1
2	878,9	886,2	1072,0	1127,7
3	827,4	843,9	1051,1	1097,7

Con estas muestras (Tabla IV - 2 y Tabla IV - 3) se ha calculado la recta de linealidad descrita en la Tabla IV - 4 y Figura IV - 4. Como se puede observar, el coeficiente de regresión es próximo a 1 y la pendiente y la ordenada de origen incluye en sus intervalos de confianza (R^2) el 1 y el 0, respectivamente. Esto indica que el método es lineal en el intervalo de concentraciones evaluado.

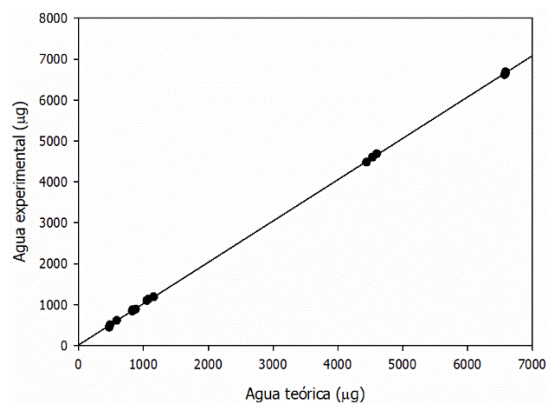


Figura IV - 4. Recta de regresión del estudio de la linealidad del método para la determinación de la humedad en los tapones de clorobutilo 13 mm.

Tabla IV - 4. Resumen de los parámetros evaluados para determinar la linealidad del método.

Parámetro	Valor
Pendiente	1,01±0,01
Ordenada en el origen	16,3±19,3
R^2	0,999

3.2.4. Intervalo de linealidad

El intervalo de linealidad en el cual se ha demostrado que el método es exacto (sección 3.2.2), lineal (sección 3.2.3) y preciso (sección 3.2.5) incluye los valores de agua entre 350 μg y 7500 μg . Estos datos en porcentaje equivalen a 0,05 % y 0,99 %, tomando como valor promedio del peso de los tapones 0,76 g. Todas las muestras que se encuentren fuera de este margen se deberán expresar como <0,05 % o >1 %.

3.2.5. Precisión

Considerando que el objetivo del método no contempla usarlo en otro laboratorio, se han evaluado solamente los niveles de precisión intermedia y repetibilidad.

- **Precisión intermedia:** para evaluar este parámetro y teniendo en cuenta que las muestras no se mantienen estables en el tiempo, se ha analizado sólo la variable analista en dos niveles de trabajo distintos. Para ello, cada analista ha analizado 6 muestras en el límite superior e inferior.

Considerando la dispersión de los resultados entre muestras, para realizar una muestra homogénea se ha seguido el mismo procedimiento que para la exactitud y la linealidad.

Tabla IV - 5. Resultados del porcentaje de humedad al analizar 6 muestras, en el límite superior e inferior, por dos analistas (A1 y A2).

Muestras	Límite superior		Límite inferior	
	A1	A2	A1	A2
1	0,93	0,87	0,05	0,05
2	0,94	0,93	0,05	0,05
3	0,95	0,89	0,05	0,05
4	0,92	0,87	0,05	0,05
5	0,95	0,93	0,05	0,06
6	0,94	0,96	0,06	0,06
Promedio	0,94	0,91	0,05	0,05
Desviación estándar.	0,012	0,037	0,004	0,005
RSD (%)	1,25	4,08	7,90	9,68

Para saber si hay diferencias entre analistas a los dos niveles estudiados, se realizó un análisis de varianzas (ANOVA). En los dos límites estudiados, el valor de F_{exp} es inferior al valor de $F_{crítica}$ (para $\alpha = 0,05$), por lo que se puede decir que no hay diferencias significativas entre analistas.

- **Repetibilidad:** para evaluar este parámetro se ha utilizado el peor valor de RSD (%) obtenido durante la precisión intermedia (Tabla IV - 5). En este caso, es el valor obtenido en el límite inferior por el segundo analista, con un RSD (%) de 9,68 %.

3.2.6. Límite de cuantificación

En este caso, el límite de cuantificación (LOQ) se ha calculado siguiendo el procedimiento de la ICH Q2. Éste se basa en aplicar la siguiente ecuación:

$$LOQ = \frac{10\sigma}{S} \quad \text{Ecuación IV - 1}$$

Donde S es la pendiente de la recta de linealidad, y σ es la desviación estándar de la respuesta. Para su cálculo se ha utilizado la desviación estándar de la ordenada en el origen de la recta de la sección 3.2.3, la cual tiene un valor de 10,44. Como resultado se obtiene como valor de LOQ de 103,3 μg de agua, que representaría un 0,01 %.

3.3. Impacto de la humedad inicial de los tapones en un producto

Una vez validado el método analítico para determinar la humedad contenida en los tapones, se ha procedido al estudio del impacto de esta humedad en el producto H.

Durante los estudios de estabildades de este producto se generan distintas impurezas; dos de ellas están directamente relacionadas con la hidrólisis. El producto H contiene dos grupos funcionales ésteres los cuales se pueden hidrolizar fácilmente con la presencia de agua formando grupos ácido. En la Figura IV - 5 se muestra la estructura del producto, juntamente con ambas impurezas de hidrólisis.

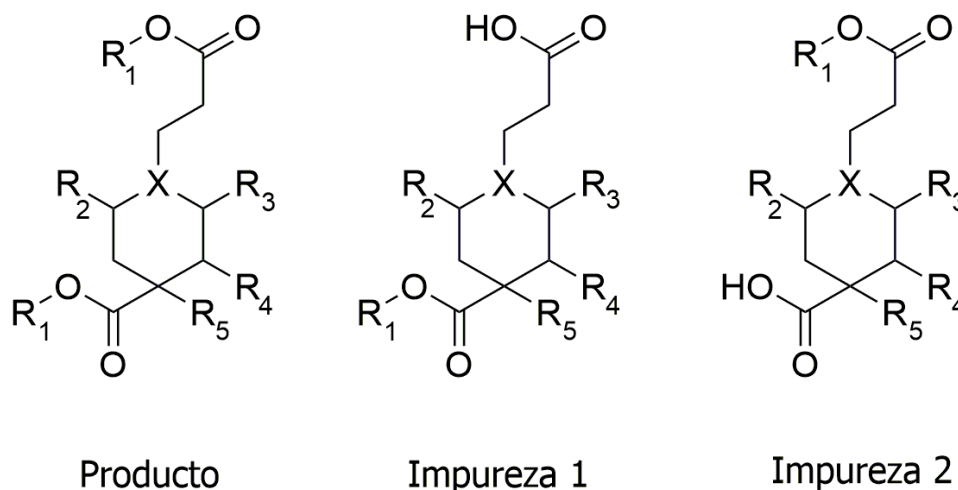


Figura IV - 5. Estructura de Markush del producto H juntamente con ambas impurezas de hidrólisis, dónde X representa un grupo -N- o -CH-; y R₁-R₅ representan uno de los siguientes grupos: un átomo de H, un átomo de halógeno, un grupo alquilo pequeño o un grupo alcoxi pequeño.

Aunque ambas impurezas son producto del mismo tipo de reacción (hidrólisis del grupo éster), la obtención de la impureza 1 se encuentra favorecida frente a la impureza 2,

debido a un menor impedimento estérico del grupo éster. Esto se ha visto confirmado observando los resultados de estabildades del producto y se ve reflejado en las especificaciones de ambas impurezas. Para la impureza 1 se tiene una especificación de no más de 2,5 %, en cambio para la impureza 2 la especificación es de no más del 0,2 %.

3.3.1. Planteamiento del experimento

Para realizar este estudio se han recogido 270 muestras del producto H fabricadas y dosificadas en la planta industrial de Reig Jofre (antes de que fueran liofilizadas) y, se han liofilizado en el liofilizador piloto Lyobeta20, siguiendo el mismo ciclo de liofilización que en rutina.

Antes de iniciar el ciclo de liofilización, a la mitad de las muestras se les cambió los tapones esterilizados por el método de rutina (en lo sucesivo se denominarán tapones de "rutina") por tapones que, después de la esterilización habitual, se han secado durante 48 h en una estufa a 105°C (en lo sucesivo se denominarán tapones "secos"). De esta manera, se dispone de 135 muestras para cada tipo de tapón.

Una vez terminado el ciclo, todas las muestras se colocaron en condiciones de estudio de estabilidad acelerado (40°C y 75 % de humedad relativa) durante 6 meses, analizando el producto cada 3 meses. En la Tabla IV - 6 se resumen los análisis realizados en cada punto de estabilidad.

Tabla IV - 6. Resumen de los análisis realizados al producto en cada punto de estabilidad.

Análisis	Antes de liofilizar	0 meses	3 meses	6 meses
Humedad de los tapones	X	X	X	X
Humedad del producto	-	X	X	X
Valoración	-	X	X	X
Sustancias relacionadas	-	X	X	X

Los resultados de estos análisis, así como su interpretación y discusión, se encuentran detallados en las siguientes secciones.

3.3.2. Humedad de los tapones

En la Figura IV - 6 se pueden observar los resultados obtenidos después de analizar la humedad de los tapones de tres muestras en los tiempos descritos en la Tabla IV - 6.

Como se puede observar en la Figura IV - 6, los tapones de rutina presentan un porcentaje de humedad del 0,5 % antes de la liofilización, valor que destaca frente al $<0,05$ % de los tapones secos.

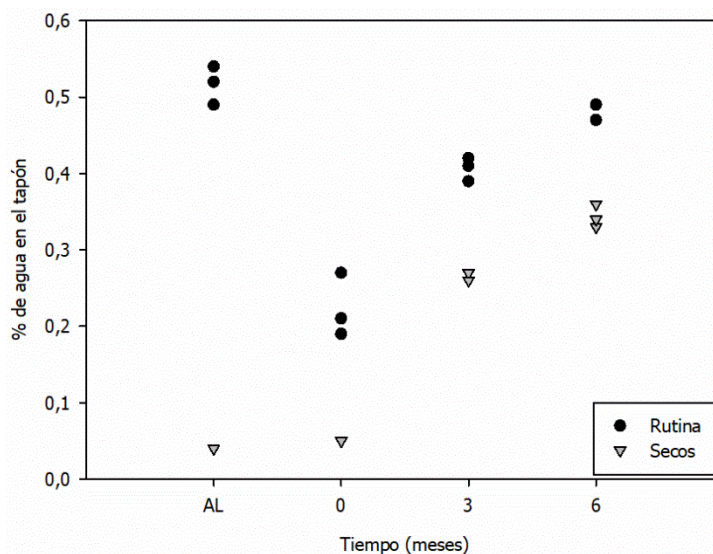


Figura IV - 6. Porcentaje de humedad que contienen los dos grupos de tapones estudiados en los puntos de estabilidad. AL indica "antes de liofilizar".

Durante el ciclo de liofilización, el agua contenida en los tapones de rutina disminuye hasta llegar a valores entre 0,2 % - 0,3 %, mientras que los tapones secos se mantienen sin apenas cambios. Esto indica que las condiciones del ciclo de liofilización tienen un efecto crítico en el valor final del contenido de humedad de los tapones.

Respecto a la evolución del contenido de agua con el tiempo, se observa que el aumento absoluto de humedad para ambos tipos de tapones es el mismo, un 0,3 %, aproximadamente, 2 mg de agua. En el caso de los tapones de rutina, en 6 meses su humedad prácticamente se iguala con la que tenían antes de la liofilización, mientras que en el caso de los tapones secos el aumento respecto al valor inicial aún queda por debajo de la humedad total de los tapones de rutina a los tres meses en condiciones del estudio de estabilidad acelerado.

3.3.3. Degradación del producto

Dado que el producto H presenta reacciones de hidrólisis, una forma de evaluar el impacto de la humedad de los tapones en el producto terminado es analizando sus impurezas de hidrólisis.

En la Figura IV - 7 se muestra la evolución con el tiempo del contenido de la impureza 1 en el producto, después de realizar su análisis por duplicado en los puntos descritos en la Tabla IV - 6.

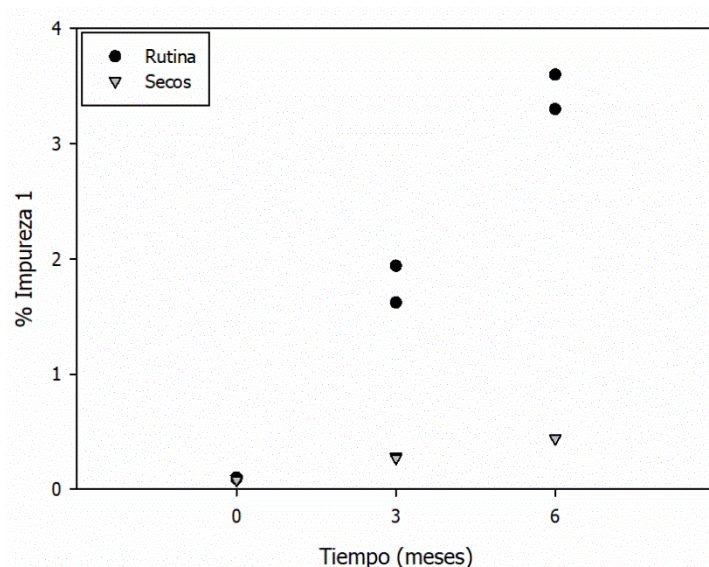


Figura IV - 7. Evolución del porcentaje de impureza 1 con el tiempo que contiene el producto para los dos grupos de tapones estudiados.

Como se puede observar, hay una clara relación entre la impureza 1 y el contenido de humedad en los tapones. Las muestras con tapones de rutina presentan un incremento de la impureza 1 que supera la especificación a los 6 meses. Sin embargo, para los tapones secos, a los 6 meses sólo llega al 0,5 %.

No obstante, el comportamiento de la impureza 2 es el opuesto al de la impureza 1. Tal y como se puede apreciar en la Figura IV - 8, con los tapones de rutina esta impureza llega al límite de la especificación (0,2 %) a los 6 meses; en cambio, al utilizar los tapones secos, esta impureza alcanza el límite de especificación a los 3 meses y a los 6 se encuentra fuera de especificaciones.

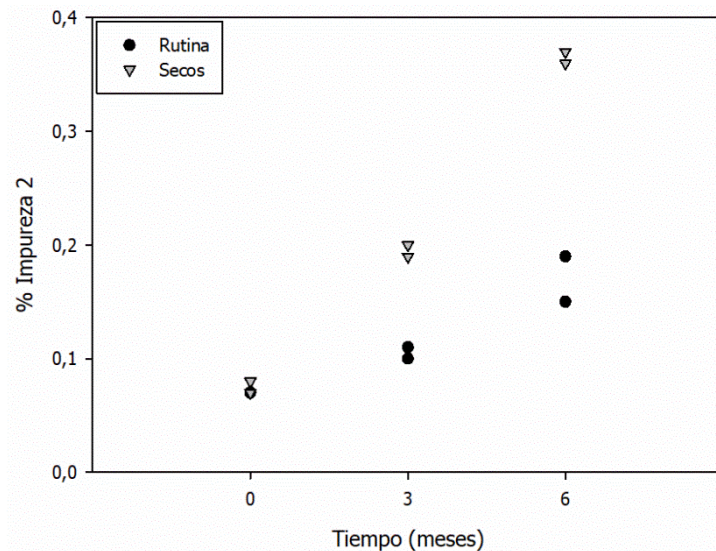


Figura IV - 8. Evolución del porcentaje de impureza 2 con el tiempo que contiene el producto para los dos grupos de tapones estudiados.

El aumento en el contenido de impurezas se refleja en la pérdida de la concentración del API en el producto a lo largo del tiempo, que se muestra en la Figura IV - 9.

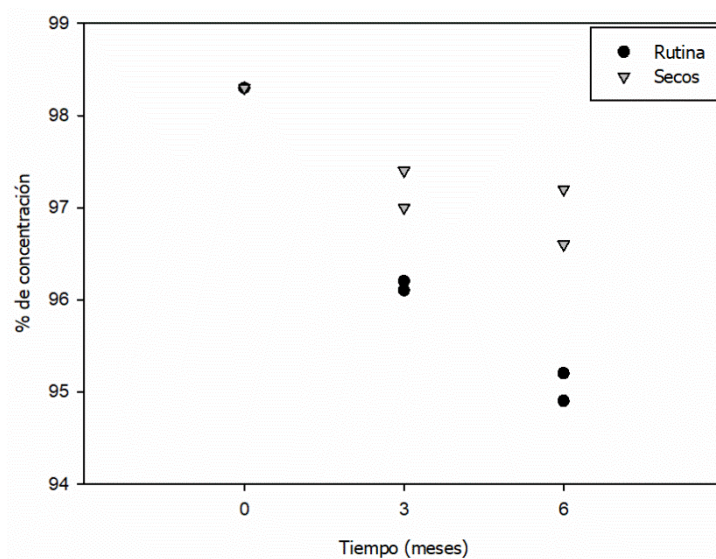


Figura IV - 9. Evolución del porcentaje de concentración del API con el tiempo, para los dos grupos de tapones estudiados.

Durante los 6 meses que se ha estudiado el producto, se observa una clara degradación de éste para ambos tipos de tapones debido a la alta temperatura de la cámara. Sin embargo, las muestras con los tapones de rutina presentan una mayor degradación que las de los tapones secos, lo que concuerda con la diferencia de concentraciones observada principalmente en la impureza 1.

La producción de las dos impurezas por hidrólisis necesita una fuente de agua, que puede proceder del propio producto o de la humedad del tapón. Puesto que el aumento

de la principal impureza es mayor en los viales con tapones de rutina, en caso de consumirse agua del producto se debería ver una disminución mayor de la humedad del producto en este tipo de muestras que en las muestras con los tapones secos.

Por este motivo, se ha estudiado la evolución del contenido de humedad en el producto H en los puntos descritos en la Tabla IV - 6.

La humedad inicial del producto es inferior al 1 % (~0,2 mg de agua) y disminuye con el tiempo de manera similar para ambos tipos de tapones.

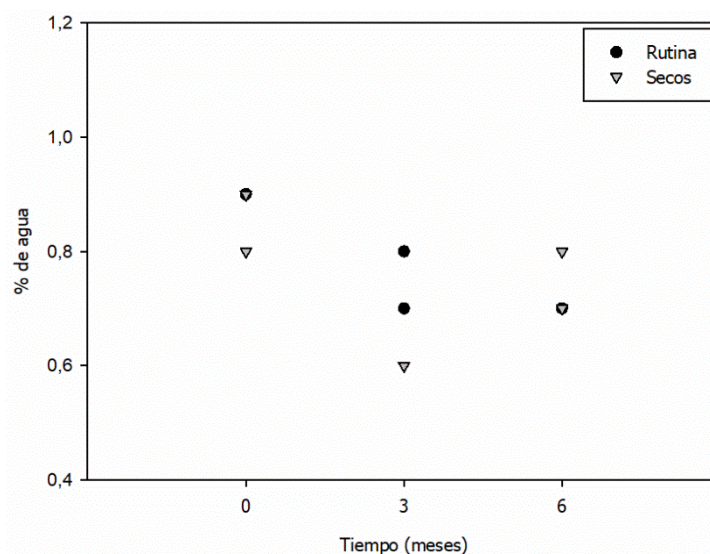


Figura IV - 10. Evolución del porcentaje de humedad con el tiempo que contiene el producto, para los dos grupos de tapones estudiados.

La baja incidencia de la impureza 1 en los viales con tapones secos frente a la que se obtiene en los viales con los tapones de rutina, y el comportamiento similar de la humedad del producto para ambos tipos de tapones, apoyan la hipótesis de que los tapones son la principal fuente de agua para la producción de esta impureza.

La impureza 2 aparece con un contenido 10 veces menor que la impureza 1 debido al impedimento estérico del grupo hidrolizado. Por lo tanto, que el aumento en las muestras con tapones secos sea mayor que en las muestras con tapones de rutina, indica que el mecanismo de incorporación del agua necesaria para la reacción es distinto, y podría proceder de la humedad residual de la propia muestra. La relación entre ambos parámetros no se ha observado claramente en los resultados obtenidos. Una explicación podría ser que, teniendo en cuenta que el incremento de la impureza 2 (en % en peso) es muy pequeño comparado con la cantidad de agua que tiene el producto, las pequeñas variaciones causadas por esta reacción en el contenido de humedad residual serían insignificantes y casi indetectables.

4. **CONCLUSIONES**

En este trabajo, se ha desarrollado y validado de forma satisfactoria un método analítico basado en la valoración coulombimétrica Karl Fischer acoplada a un horno, para la determinación de la humedad contenida en un tapón de clorobutilo de 13 mm.

En este caso, el desarrollo del método y del procedimiento de análisis se han basado en la bibliografía consultada y en las pruebas previas a la validación del método. El método ha sido validado en un intervalo de concentraciones de 0,05 % hasta el 1 % de humedad en tapones de 13 mm, demostrando que en tal intervalo es un método lineal, exacto y preciso.

Una vez desarrollado el método, se ha evaluado el impacto de la humedad contenida en los tapones esterilizados siguiendo el procedimiento de rutina, y la de tapones con la misma esterilización más una etapa adicional de secado posterior, de 48 h a 105°C. Tal y como se ha expuesto en los resultados, se ha observado que la humedad contenida en los tapones de rutina varía durante el ciclo de liofilización, disminuyendo su valor inicial; el proceso de liofilización seca el producto y también los tapones.

Ambos tapones, independientemente de la humedad inicial que presentaran, han mostrado un comportamiento similar con el tiempo, aumentando su contenido de humedad hasta, aproximadamente, 2 mg de agua.

La aparición de la impureza 1 está relacionada con la humedad procedente del tapón, mientras que la de la impureza 2 se produciría a partir de la humedad de la muestra.

5. **BIBLIOGRAFÍA**

- (1) Templeton, A. C.; Placek, J.; Xu, H.; Mahajan, R.; Hunke, W. A.; Reed, R. A. Determination of the Moisture Content of Bromobutyl Rubber Stoppers as a Function of Processing: Implications for the Stability of Lyophilized Products. *PDA J. Pharm. Sci. Technol.* 2003, 57(2), 75–87.
- (2) Wagner, T.; Van Mullekom, J.; Severin, J.; Scholla, M. H. Package/Container Closures. In *Assurance of Sterility for Sensitive Combination Products and Materials*; Elsevier, 2020; pp 79–134.
- (3) Le Meste, M.; Simatos, D.; Préaud, J. M.; Precausta, P. M. Factors Influencing Changes in Moisture Content during Storage of Freeze-Dried Vaccines in Vials. *J. Biol. Stand.* 1985, 13(3), 177–184.
- (4) Earle, J. P.; Bennett, P. S.; Larson, K. A.; Shaw, R. The Effects of Stopper Drying on Moisture Levels of Haemophilus Influenzae Conjugate Vaccine. *Dev. Biol.*

- Stand.* 1992, 74, 203–210.
- (5) Pikal, M. J.; Shah, S. Moisture Transfer from Stopper to Product and Resulting Stability Implications. *Dev. Biol. Stand.* 1992, 74, 165–177; discussion 177-9.
 - (6) Nema, S.; Ludwig, J. D. *Pharmaceutical Dosage Forms - Parenteral Medications*, 3rd ed.; Nema, S., Ludwig, J. D., Eds.; CRC Press, 2016; Vol. 1.
 - (7) Oetjen, G.-W.; Haseley, P. *Freeze-Drying*, 2nd ed.; Wiley-VCH GmbH & Co. KGaA, 2004.
 - (8) European Pharmacopoeia. 2.5.32. Water Micro Determination. 2019, 20532.
 - (9) Laboratorio Reig Jofre. *Procedimiento Analítico de Producto Terminado*, AP0396_07; 07; 2017.
 - (10) Wang, Z.; Frankel, B. A.; Lambert, W. J. Determination of Moisture in Rubber Stoppers: Effect of Karl Fischer Oven Temperatures. *PDA J. Pharm. Sci. Technol.* 2001, 55 (3), 162–170.

V.DETERMINACIÓN DE LA HUMEDAD
RESIDUAL EN UN PRODUCTO
LIOFILIZADO

1. INTRODUCCIÓN

Como se ha comentado en la introducción de la tesis, la determinación del contenido de humedad residual (RMC) en un producto liofilizado se considera un atributo crítico de la calidad (CQA). La cuantificación de este atributo aporta información sobre cómo se ha desarrollado el ciclo de liofilización y sobre la evolución de la calidad del producto durante su vida útil.

En la determinación de este CQA en los lotes de rutina sólo se analizan unas pocas muestras para obtener un único resultado que se extrapola al resto del lote. Para poder realizar esta aproximación, se debe estudiar la homogeneidad de este atributo durante la validación del proceso de fabricación (PPQ, siglas del inglés *Process Performance Qualification*), tanto en un lote (dentro de la cámara del liofilizador) como entre distintos lotes y liofilizadores.^{1,2}

Para el estudio de la homogeneidad de RMC se realiza un muestreo extensivo del lote, por lo que, si se utilizan métodos clásicos, como la valoración por Karl Fischer, se requiere mucho tiempo y aumenta el impacto económico y medioambiental del proceso de validación. Teniendo en cuenta estos factores, la espectroscopia NIR y el desarrollo de un modelo cuantitativo para la determinación de este CQA se presenta como una alternativa interesante.³

En este apartado, se detallan las diferentes etapas realizadas (Figura V - 1) para el desarrollo de un modelo multivariante, basado en la espectroscopia NIR, con el objetivo de cuantificar el RMC de un producto liofilizado, así como su aplicación en el estudio de homogeneidad de este CQA en los lotes de PPQ.



Figura V - 1. Esquema de la estructura seguida durante el desarrollo de este capítulo.

2. MATERIALES Y MÉTODOS EXPERIMENTALES

2.1. Muestras

Las muestras estudiadas en este apartado consisten en un producto parenteral liofilizado, por temas de confidencialidad el nombre de éste debe permanecer en el anonimato, por lo que será denominado producto N.

El producto N consiste en un medicamento que se administra por vía intravenosa y una vez liofilizado tiene un aspecto de polvo blanco o casi blanco. Este producto lleva fabricándose en la compañía desde hace más de 10 años. Sin embargo, en 2018 se decidió realizar varios cambios en el proceso de fabricación (cambio de vial y modificación de los parámetros del ciclo de liofilización, entre otros) que llevaron a realizar una nueva validación de proceso. El nuevo producto se presenta envasado en un vial incoloro de tubo de 10 ml, con un tapón de clorobutilo de 20 mm de diámetro y una cápsula de aluminio con tapa. Tabla V - 1 muestra la composición cualitativa de este producto.

Tabla V - 1. Composición del Producto N.

Compuesto	Función^{4,5}
API	Compuesto activo que proporciona eficacia al medicamento.
Manitol	Compuesto con propiedades de carga, que tiene la función de incrementar el contenido de sólidos totales de la formulación.
Excipiente 1	Compuesto con propiedades quelantes, que tiene la función de estabilizar el producto en solución.
Excipiente 2	Compuesto empleado para ajustar el pH de la solución.
Agua	Solvente.
Nitrógeno	Gas neutro empleado para desplazar el oxígeno presente en la solución y utilizado para la composición del <i>headspace</i> del producto final.

Debido a los cambios en el proceso de fabricación y al posterior registro y aprobación del dossier del producto N, su producción en rutina no ha estado disponible durante la realización de esta tesis. Esto ha implicado que no se dispusieran de lotes industriales de rutina durante el desarrollo del modelo NIR. Con el fin de poder incorporar la variabilidad de los lotes industriales al modelo NIR, se han incluido muestras procedentes del lote previo a la validación del proceso (lote de ingeniería); tanto en su desarrollo como en su validación.

Para cubrir el intervalo de concentraciones de RMC necesario, se ha utilizado solución de lotes de planta del producto precedente y se han liofilizado en el liofilizador piloto LB20. Este liofilizador dispone de un extractor de muestras, que permite extraerlas durante el ciclo de liofilización (final del secado primario y durante el secado secundario) para obtener muestras con distintos RMC.

2.2. Método de referencia

El método analítico empleado para la cuantificación del agua del producto N, anterior a la validación del proceso del 2018, consistía en realizar una mezcla de 6 muestras bajo una lámpara infrarroja y realizar dos determinaciones de dicha mezcla. La técnica empleada para este análisis era la valoración volumétrica de Karl Fischer.

Este procedimiento requería una gran manipulación de una muestra muy higroscópica, por lo que se elevaba la dispersión de los resultados. Por ello, se realizó el desarrollo y la validación de un nuevo método analítico para cuantificar el RMC que permitiera reducir el impacto de la manipulación. Este método se utilizó como método de referencia para el desarrollo del modelo NIR. En la Tabla V - 2 se encuentran resumidas las principales características y parámetros del método.

Tabla V - 2. Parámetros y características del método de referencia del producto N para la cuantificación del RMC.

Técnica	Karl Fischer Culombimétrico acoplado a un horno	Temperatura horno para muestras y blancos	80°C
Tiempo estabilización	30 segundos	Temperatura horno para estándares	145°C
Deriva inicial	< 20 µg/min	Intensidad del generador	400 mA
Criterio de parada	Deriva relativa (5 µg/min)	Corriente de polarización	10 µA
Tiempo de extracción de la muestra	400 segundos	Flujo del nitrógeno	90 ml/min

El procedimiento se inicia con el análisis de tres blancos. El primero de ellos permite acondicionar todo el sistema, eliminando las trazas de agua que pueda haber. Los otros dos, permiten determinar la humedad propia de la preparación de muestras y del equipo; el promedio de su valor se sustrae a las siguientes determinaciones. A continuación, se procede al análisis de las muestras, intercalando un patrón, de aproximadamente 560 µg

DETERMINACIÓN DE LA HUMEDAD RESIDUAL EN UN PRODUCTO LIOFILIZADO

de agua, cada tres muestras para verificar la fiabilidad de los resultados. El porcentaje de recuperación de estos patrones, se debe encontrar dentro del intervalo 90 – 110 %.⁶

Todas las muestras (blancos y estándares inclusive) se deben preparar en una cabina de humedad controlada (humedad relativa inferior al 10 %) para reducir la manipulación de la muestra.

Este método se validó siguiendo el procedimiento descrito en la ICH Q2(R1), y en la Tabla V - 3 se pueden ver los resultados obtenidos.⁷

Tabla V - 3. Resultados de la validación del método analítico de referencia para la determinación de agua en el producto N.

Parámetro	Descripción del procedimiento
Especificidad	Con los resultados obtenidos de realizar una rampa de temperatura y un análisis termogravimétrico, se ha concluido que a 80°C el producto no se degrada y se puede extraer toda el agua de la muestra.
Linealidad	La linealidad se ha evaluado con 8 puntos distribuidos en el intervalo de concentración. Los resultados demuestran que el método es lineal con un R^2 del 0,999 y un pendiente y ordenada de origen próximas a 1 y 0, respectivamente.
Intervalo de linealidad	El intervalo en que cual se ha demostrado que el método es lineal, exacto y preciso es de 240 µg a 5260 µg.
Exactitud	La exactitud se ha evaluado en 5 puntos de concentración, comparando el valor obtenido con el método en validación con el resultado de la valoración volumétrica de Karl Fischer. Los resultados demuestran que el método es exacto en el intervalo de concentraciones estudiado.
Precisión intermedia	La precisión intermedia se ha evaluado realizando un ANOVA con los factores día y analista, obteniendo que ninguno de ellos era significativo para $\alpha = 0,05$.
Repetibilidad	La repetibilidad se ha evaluado estudiando los 6 replicados analizados por analista y día de la precisión intermedia. El RSD (%) obtenido es del 15 %, por lo que teniendo en cuenta que el producto tiene menos de 1 mg de agua se acepta el valor y se puede concluir que el método es repetible.
Límite de cuantificación	El límite cuantificación, calculado con la recta obtenida en la linealidad, es de 99,6 µg.
Robustez	Para evaluar la robustez del método se han realizado pequeñas variaciones en los parámetros de temperatura y de flujo de nitrógeno, sin observar un impacto significativo en los resultados de las muestras.

Durante la validación del método, se calculó el error estándar de laboratorio (SEL). Para ello se utilizaron los 24 valores de la precisión intermedia y se calculó su desviación estándar, obteniendo un valor de SEL igual a 0,22 %.

2.3. Adquisición de los espectros

Los espectros de las muestras utilizadas en este apartado se han registrado empleando el espectrofotómetro NIR Labspec5000 descrito anteriormente, siguiendo las conclusiones extraídas del estudio de viabilidad del modelo NIR que se describe en la sección 1.3.1.

El procedimiento consiste en colocar la parte inferior del vial encima de la entrada de luz del accesorio ASD Muglight, intentando cubrir el máximo de superficie del haz de luz. Para cada muestra, se debe registrar su espectro por triplicado, realizando una rotación de 120° entre cada uno, sin necesidad de apartar la muestra de la fuente de luz.

Previamente al desarrollo o aplicación del modelo, los tres espectros registrados por cada muestra se pasan a unidades de absorbancia. Después de aplicar el pretratamiento los tres espectros de cada muestra se promedian con el fin de tener sólo un espectro por muestra.

3. RESULTADOS Y DISCUSIÓN

3.1. Análisis de viabilidad del método

Previamente al desarrollo del modelo cuantitativo NIR, siguiendo las recomendaciones de la EMA, se ha realizado un análisis de viabilidad de la aplicación del método NIR.⁸

Una de las opciones que se proponen en la guía de la EMA es ejecutar un análisis de riesgos que permita identificar, evaluar y controlar aquellos parámetros que puedan tener influencia, tanto en el desarrollo, como en la validación del modelo NIR. En esta tesis se ha utilizado el análisis modal de fallos y efectos (FMEA, siglas del inglés *Failure Mode and Effects Analysis*) para realizar el análisis de riesgos.

El FMEA consiste en realizar una evaluación inicial analizando las posibles formas en que pueden fallar los diferentes parámetros estudiados. Seguidamente, se estudian diferentes acciones para reducir o mitigar el impacto de estos fallos potenciales en el desarrollo y validación del modelo NIR.⁹

Para determinar que parámetro se clasifica como un riesgo y debe ser evaluado, se calcula el número de prioridad de riesgo (RPN, siglas del inglés *Risk Priority Number*). Este valor se obtiene de multiplicar la severidad del posible fallo, por la probabilidad de que ocurra, por la facilidad de detectarlo. Si el valor de esta operación se encuentra por

DETERMINACIÓN DE LA HUMEDAD RESIDUAL EN UN PRODUCTO LIOFILIZADO

debajo de un umbral, se considera que ese parámetro no constituye un riesgo, si por el contrario se encuentra por encima debe evaluarse para tratar de mitigar o reducir su impacto. El valor máximo que se puede obtener en el RPN es de 1000, por lo que el umbral elegido ha sido del 2 %, es decir, 20 puntos.

- **Severidad.** La severidad o gravedad de un fallo, abreviada como S, se ha clasificado en 3 niveles, descritos en la Tabla V - 4.

Tabla V - 4. Clasificación de los niveles de la severidad utilizados.

Definición de la severidad del fallo	Consecuencia del fallo a este nivel de severidad	Puntuación
Ninguno	No tiene ningún efecto relevante en el desarrollo del modelo.	1
Bajo	Tiene un efecto relevante en el desarrollo del modelo.	5
Alto	Tiene un efecto muy relevante en el desarrollo del modelo.	10

- **Probabilidad.** La probabilidad de que se produzca un fallo, abreviada como P, se puede puntuar con valores del 1 al 10, definidos en la Tabla V - 5.
- **Detección.** La facilidad con la que se puede detectar si el fallo se ha producido o no, abreviada como D, se puede puntuar con valores del 1 al 10, definidos en la Tabla V - 5.

Tabla V - 5. Definición y clasificación de la probabilidad y detección de un fallo.

Probabilidad del fallo	Detección del fallo	Puntuación
Improbable	Seguro	1
Muy baja	Muy alta	2
Baja	Alta	3
Ocasional	Altamente moderada	4
Posible	Moderada	5
Frecuente	Baja	6
Muy frecuente	Muy baja	7
Común	Remota	8
Alta	Muy remota	9
Muy alta	Casi imposible	10

La selección previa de qué parámetros se debían introducir en el análisis de riesgos se realizó empleando un diagrama de Ishikawa, también conocido como diagrama de espina

de pez (Figura V - 2). Cada una de estas espinas representan las etapas que intervienen en el proceso de registro del espectro, las cuales se desglosan en los diferentes parámetros que pueden tener un efecto sobre la señal.

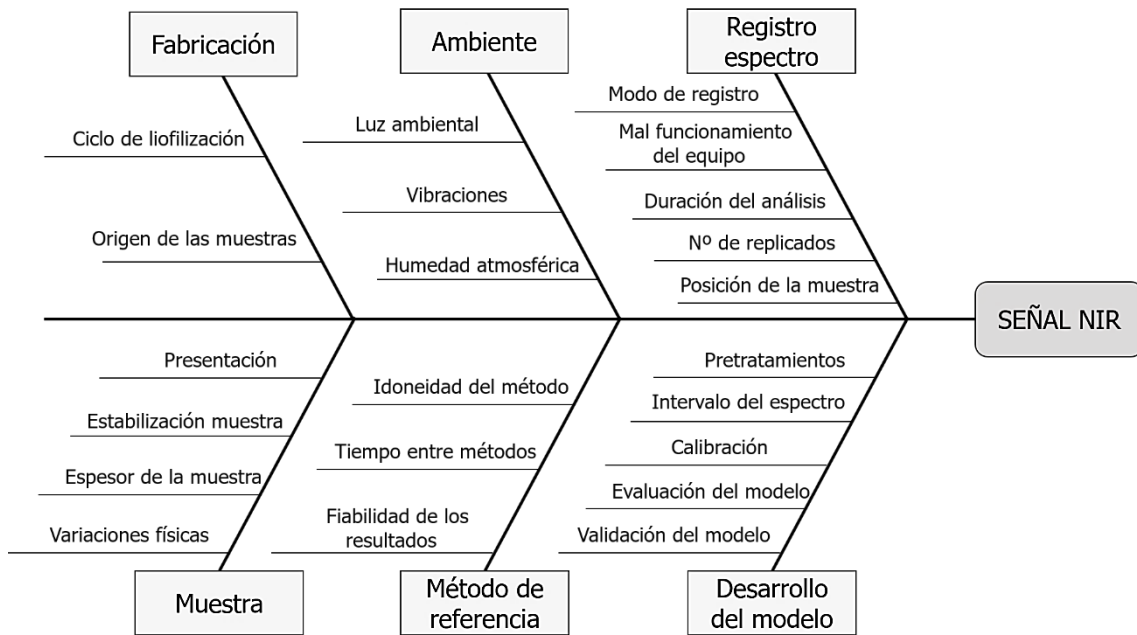


Figura V - 2. Diagrama de Ishikawa empleado para identificar los principales parámetros que podrían afectar al desarrollo del modelo NIR alterando la señal del espectro.

Teniendo en cuenta esta organización se ha realizado el FMEA de esos parámetros, dividiendo cada una de las etapas en una tabla individual (Tabla V - 6 a Tabla V - 11).

DETERMINACIÓN DE LA HUMEDAD RESIDUAL EN UN PRODUCTO LIOFILIZADO

Tabla V - 6. Análisis de riesgo asociado a la fabricación de la muestra.

Fabricación											
Parámetros	Modo de fallo potencial	Efecto del potencial fallo	S	P	D	Riesgo inicial	Acciones de mitigación	S	P	D	Riesgo residual
Ciclo de liofilización	El ciclo no es adecuado para liofilizar el producto N.	Se produciría el colapso de las muestras variando la estructura de éstas. La pérdida de estructura aumentaría la variabilidad del espectro entre muestras.	10	7	3	210	El ciclo de liofilización usado para la producción de las muestras debe encontrarse dentro del espacio de diseño de éste. Sólo se variará la duración de las etapas de secado con el fin de obtener muestras con distintos valores de RMC.	10	1	1	10
Origen de las muestras	Las muestras producidas con el liofilizador piloto presentan una variabilidad en el espectro diferente al de las muestras industriales.	El modelo NIR no predeciría correctamente las muestras industriales debido a que el modelo no contiene su variabilidad.	10	8	2	160	Dado que el producto no se encuentra en producción. La variabilidad industrial se incorporará usando muestras del lote de ingeniería del proceso de validación.	10	1	1	10

Tabla V - 7. Análisis de riesgos asociados a las condiciones ambientales durante el registro del espectro.

Ambiente											
Parámetros	Modo de fallo potencial	Efecto del potencial fallo	S	P	D	Riesgo inicial	Acciones de mitigación	S	P	D	Riesgo residual
Luz ambiental	Cambios en la intensidad o el tipo de luz en el lugar de registro de los espectros.	El espectro NIR podría mostrar cambios debido a las diferencias entre las condiciones de la luz.	10	3	1	30	Los espectros se deben registrarse con la luz encendida para proporcionar una luz ambiental constante.	10	1	1	10

Ambiente											
Parámetros	Modo de fallo potencial	Efecto del potencial fallo	S	P	D	Riesgo inicial	Acciones de mitigación	S	P	D	Riesgo residual
Vibraciones	Vibraciones no controladas durante el registro de los espectros.	El ruido del espectro podría incrementar a causa de las vibraciones.	10	5	1	50	El espectrofotómetro está colocados encima de una poyata fijada para reducir las vibraciones al mínimo.	10	1	1	10
Humedad atmosférica	Trabajar con una humedad atmosférica elevada y no controlada.	La humedad de la muestra podría verse afectada observando cambios en el espectro debido a la absorción de trazas de humedad.	10	3	1	30	La humedad relativa del laboratorio se mantiene constante y baja durante el registro de los espectros. Además, las muestras se encuentran capsuladas para minimizar la entrada de humedad atmosférica.	10	1	1	10

Tabla V - 8. Análisis de riesgos asociados al registro del espectro.

Registro espectro											
Parámetros	Modo de fallo potencial	Efecto del potencial fallo	S	P	D	Riesgo inicial	Acciones de mitigación	S	P	D	Riesgo residual
Modo de registro: reflectancia	El modo de registro por reflectancia no es adecuado para la muestra.	Los espectros no serían representativos, lo que afectaría a la exactitud y precisión del modelo.	10	5	5	250	Según la literatura, el modo de reflectancia es idóneo para muestras sólidas. Además, considerando que la luz del NIR puede profundizar unos 2-3 mm en la muestra, la reflectancia permite el registro del espectro directo. ¹⁰	10	1	1	10

DETERMINACIÓN DE LA HUMEDAD RESIDUAL EN UN PRODUCTO LIOFILIZADO

Registro espectro											
Parámetros	Modo de fallo potencial	Efecto del potencial fallo	S	P	D	Riesgo inicial	Acciones de mitigación	S	P	D	Riesgo residual
Mal funcionamiento del equipo	Mal funcionamiento del espectrofotómetro NIR debido a factores externos o internos.	Los espectros realizados tendrían valores erróneos en su totalidad o en diferentes regiones dependiendo del tipo de fallo.	10	3	2	60	Diariamente se debe verificar el correcto funcionamiento del espectrofotómetro NIR siguiendo las directrices regulatorias. ^{11,12}	10	1	1	10
Duración del análisis	La exposición prolongada de la muestra sobre el haz de luz calienta la muestra.	La superficie de la muestra escaneada podría variar su contenido de humedad o incrementar sus impurezas observando variaciones en el espectro.	10	7	1	70	Se necesita más información para determinar el efecto de la exposición prolongada de la muestra frente al haz de luz para definir las acciones de mitigación y calcular el riesgo residual.	-	-	-	Por definir
Número de replicados	Algunos de los replicados no son representativos de la muestra.	El espectro promedio estaría sesgado.	10	7	2	140	Para poder detectar espectros anómalos entre replicados, se deben registrar tres espectros por cada muestra.	10	1	1	10
Posición de la muestra	La colocación de las muestras sobre el accesorio no es repetitiva y no se cubre la máxima superficie del haz de luz.	Los niveles de error del modelo aumentarían debido a las variaciones de la línea base y/o aumentaría el ruido del espectro.	10	4	4	160	Se necesita más información para determinar el efecto de la posición de la muestra frente al haz de luz para definir las acciones de mitigación y calcular el riesgo residual.	-	-	-	Por definir

Tabla V - 9. Análisis de riesgos asociados a la muestra.

Muestra												
Parámetros	Modo de fallo potencial	Efecto del potencial fallo	S	P	D	Riesgo inicial	Acciones de mitigación	S	P	D	Riesgo residual	
Presentación	Aspecto de la muestra	La pastilla del producto presenta una superficie irregular (rotura en múltiples trozos o polvo en la parte inferior del vial).	Los replicados de los espectros no serían repetitivos o aumentaría el <i>scattering</i> de los espectros aportando una variabilidad extra a estas muestras.	10	4	4	160	Esta variabilidad en el aspecto de la muestra es algo propio del proceso de fabricación, por lo que se debe introducir en el modelo.	10	1	1	10
	Calidad óptica del cristal	El cristal del vial no es transparente a la luz NIR, mostrando un espectro de absorción.	Las bandas de absorción del cristal podrían interferir con el espectro del producto N.	10	4	6	240	Siempre se utilizará el mismo tipo de vial. Por lo que se debe incluir esta variabilidad dentro del modelo.	10	1	1	10
	Temperatura de la muestra durante el análisis	La temperatura de la muestra durante el análisis no es constante o se encuentra por encima de la temperatura recomendada de almacenamiento (25°C).	Esta variación podría modificar la homogeneidad del agua en la muestra o aumentar la migración de la humedad de los tapones a la muestra. Estos flujos podrían afectar a la intensidad de las bandas del agua o aumentar el error entre el modelo y el método de referencia.	10	7	5	350	Se necesita más información para determinar el efecto de la temperatura de la muestra durante el análisis en el espectro NIR para definir las acciones de mitigación y calcular el riesgo residual.	-	-	-	Por definir
Estabilización de la muestra	Las muestras son inestables y no se mantienen constantes con el tiempo.	Se introduciría una variabilidad en el espectro, que causaría un incremento en el error del modelo. Por ejemplo, el cambio de la	10	7	1	70	Se necesita más información para determinar el efecto de la estabilidad de la muestra durante el tiempo para definir las	-	-	-	Por definir	

DETERMINACIÓN DE LA HUMEDAD RESIDUAL EN UN PRODUCTO LIOFILIZADO

Muestra											
Parámetros	Modo de fallo potencial	Efecto del potencial fallo	S	P	D	Riesgo inicial	Acciones de mitigación	S	P	D	Riesgo residual
		concentración de la humedad entre el espectro y el análisis del método de referencia.					acciones de mitigación y calcular el riesgo residual.				
Espesor de la muestra	La pastilla del producto no es lo suficiente compacta o el espesor no es el adecuado para evitar que la luz no sea transmitida a través de la muestra.	Aumentaría la variabilidad entre espectros de diferentes puntos de una misma muestra y entre distintas muestras, debido a una pérdida de señal.	10	4	3	120	Se necesita más información para determinar el efecto del espesor de la muestra para definir las acciones de mitigación y calcular el riesgo residual.	-	-	-	Por definir
Variaciones físicas	Polimorfismo	Presencia de polimorfismo en el producto liofilizado.	10	3	3	90	Se necesita más información para determinar si hay presencia de polimorfos en la muestra, para definir las acciones de mitigación y calcular el riesgo residual.	-	-	-	Por definir
	Tamaño partículas	Diferencias en el tamaño de las partículas en el producto liofilizado.	10	5	3	150	Teniendo en cuenta que la formación de los cristales durante la congelación es la causa de la diferencia entre el tamaño de partículas y que éste es un fenómeno espontaneo e incontrolable, la mitigación del riesgo se basa en incorporar muestras del máximo número de lotes posible.	10	1	1	10

Muestra											
Parámetros	Modo de fallo potencial	Efecto del potencial fallo	S	P	D	Riesgo inicial	Acciones de mitigación	S	P	D	Riesgo residual
Homogeneidad del agua	El agua no se encuentra distribuida de forma homogénea en la muestra.	Incoherencias entre los valores obtenidos con los espectros NIR y el método de referencia, causando una falta de exactitud del modelo.	10	5	1	50	Se necesita más información para determinar si el agua se encuentra distribuida homogéneamente en la muestra para definir las acciones de mitigación y calcular el riesgo residual.	-	-	-	Por definir

Tabla V - 10. Análisis de riesgos asociados al método de referencia.

Método de referencia											
Parámetros	Modo de fallo potencial	Efecto del potencial fallo	S	P	D	Riesgo inicial	Acciones de mitigación	S	P	D	Riesgo residual
Idoneidad del método	El método de referencia no es el idóneo para cuantificar la humedad del producto.	Los valores de referencia no serán los correctos causando un error en la exactitud del modelo.	10	4	4	160	Previamente al desarrollo del modelo NIR se ha estudiado y modificado el método de referencia para asegurar su idoneidad (2.2).	10	1	1	10
Tiempo entre el registro del espectro NIR y el análisis con el método de referencia	El tiempo transcurrido entre el registro del espectro y el análisis con el método de referencia no es representativo.	Los valores de referencia no serían relacionables con el espectro tomado de la muestra, lo que causaría un error en la exactitud del modelo.	10	5	6	300	Se necesita más información para determinar el efecto del tiempo entre el registro del espectro y el análisis con el método de referencia para definir las acciones de mitigación y calcular el riesgo residual.	-	-	-	Por definir

DETERMINACIÓN DE LA HUMEDAD RESIDUAL EN UN PRODUCTO LIOFILIZADO

Método de referencia											
Parámetros	Modo de fallo potencial	Efecto del potencial fallo	S	P	D	Riesgo inicial	Acciones de mitigación	S	P	D	Riesgo residual
Fiabilidad de los resultados	Los resultados del método de referencia no son fiables.	Los valores de referencia podrían no ser correctos, lo que causaría un error en la exactitud y precisión del modelo NIR.	10	6	4	240	Para asegurar la fiabilidad de los resultados previamente al desarrollo del modelo NIR se lleva a cabo la validación del método de referencia (Tabla V - 3).	10	1	1	10

Tabla V - 11. Análisis de riesgos asociados al desarrollo del modelo.

Desarrollo del modelo											
Parámetros	Modo de fallo potencial	Efecto del potencial fallo	S	P	D	Riesgo inicial	Acciones de mitigación	S	P	D	Riesgo residual
Pretratamientos	Al aplicar los pretratamientos, se produce un aumento del ruido o modificación de las diferencias entre muestras con distinto RMC.	El modelo NIR final presentaría una baja capacidad predictiva.	10	5	1	50	Se estudiarán distintos pretratamientos y combinaciones de éstos con los espectros del producto N, comparándolos entre ellos. Este proceso es una etapa del proceso de desarrollo de un modelo multivariante.	10	1	1	10
Intervalo de longitudes de onda	El intervalo de longitudes de onda considerado para desarrollar el modelo no incluye las bandas principales del agua o no es adecuado.	El modelo NIR final presentaría una mala exactitud, robustez y una baja capacidad predictiva.	10	5	1	50	Se debe incluir como mínimo la banda de mayor intensidad del agua en el intervalo de longitudes de onda del modelo NIR y debe justificarse el escogido.	10	1	1	10

Desarrollo del modelo											
Parámetros	Modo de fallo potencial	Efecto del potencial fallo	S	P	D	Riesgo inicial	Acciones de mitigación	S	P	D	Riesgo residual
Calibración	Las muestras del conjunto de calibración no incluyen la variabilidad de los lotes industriales.	Durante la validación del modelo NIR se observarían unos altos residuales en muestras industriales y una baja capacidad predictiva.	10	5	1	50	Es necesario incluir muestras que cubran todo el intervalo de RMC estudiado, así como muestras industriales para incorporar esta variabilidad al modelo final.	10	1	1	10
Evaluación del modelo	El conjunto de evaluación no cubre toda la variabilidad ni el intervalo de RMC estudiado.	Durante la validación del modelo se observarían varianzas residuales altas al igual que el RMSEP.	10	5	1	50	Se debe usar un conjunto de muestras que cubra toda la variabilidad estudiada. Si no se disponen de suficientes muestras con estas características, se debe plantear la opción de utilizar técnicas de validación cruzada.	10	1	1	10
Validación del modelo	Las muestras del conjunto de validación se han usado durante la calibración o no incluyen toda la variabilidad posible.	El modelo NIR final tendría una validación poco robusta y el modelo presentaría una mala capacidad predictiva con muestras de rutina.	10	5	1	50	Las muestras del conjunto de validación deben ser representativas a las muestras que se van a encontrar en rutina e incluir la máxima variabilidad posible.	10	1	1	10

De acuerdo con los resultados obtenidos con el FMEA, hay ocho parámetros que requieren de información experimental para poder definir las acciones de mitigación pertinentes y calcular su riesgo residual final. A continuación, se describe cómo se ha obtenido esta información, tanto el planteamiento como los resultados y conclusiones.

- Duración del análisis: dependiendo de cómo se realicen los replicados de la muestra: consecutivos o retirando la muestra después de registrar cada espectro, el tiempo de exposición frente a la fuente de luz puede variar.

El impacto de esta exposición prolongada se ha estudiado comparando los espectros de dos muestras registradas con ambos procedimientos. Primero, extrayendo la muestra de la luz y luego prolongando la exposición. Este último procedimiento se ha repetido 5 veces, retirando la muestra de la fuente de luz después de realizar los triplicados, para ver si el impacto de la exposición prolongada se intensificaba al reanalizar la muestra.

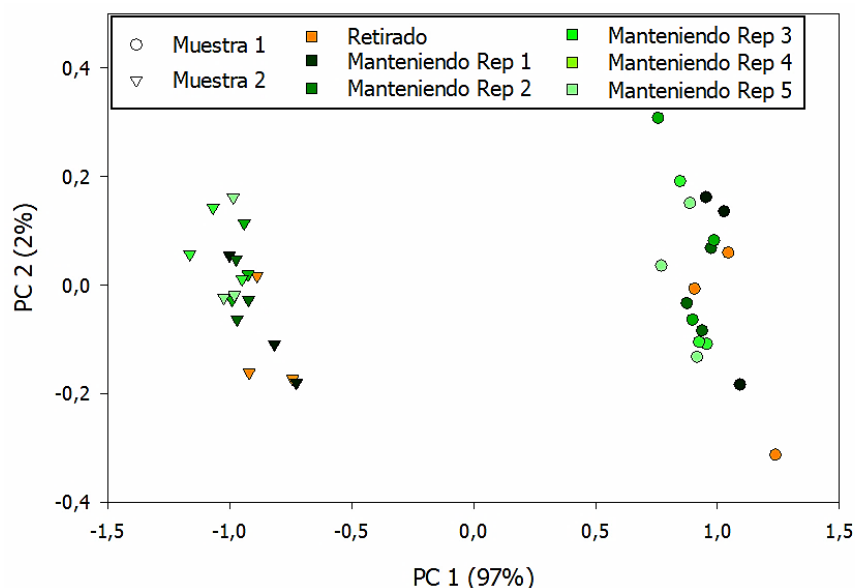


Figura V - 3. Representación de las dos primeras componentes principales (PC) (99 % de la variabilidad total) de los espectros NIR de las dos muestras estudiadas. Los espectros se han pretratado mediante SNV en el intervalo de longitudes de onda de 1100 nm a 2500 nm.

Como se puede observar en la Figura V - 3, los espectros quedan agrupados en dos grupos, correspondientes a cada una de las muestras. Dentro de estas agrupaciones, los espectros realizados con ambos procedimientos no muestran ninguna agrupación, ni se observa ninguna tendencia entre los análisis repetidos con exposición prolongada.

Gracias a esta información, se puede concluir que la exposición prolongada de la muestra durante el registro de los espectros no tiene un impacto en la señal NIR.

- Posición de la muestra durante el registro del espectro: dado que el diámetro del vial de la muestra es más pequeño que el de la ventana de la fuente de luz, la posición de la muestra puede variar entre replicados y entre muestras.

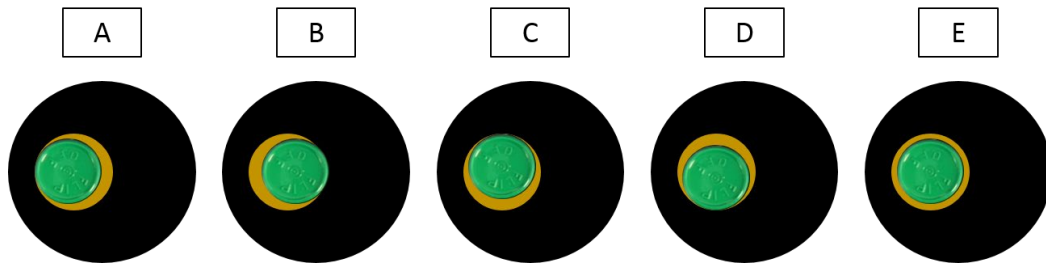


Figura V - 4. Representación de las cinco posiciones de medida estudiadas. Visto desde arriba, el círculo negro representa el accesorio Muglight, el círculo naranja representa la ventana de la fuente de luz y el círculo verde el vial de la muestra.

El impacto de esta variación en la posición de la muestra en el registro del espectro se ha estudiado analizando 3 muestras en 5 posiciones distintas (Figura V - 4) y comparando los espectros obtenidos.

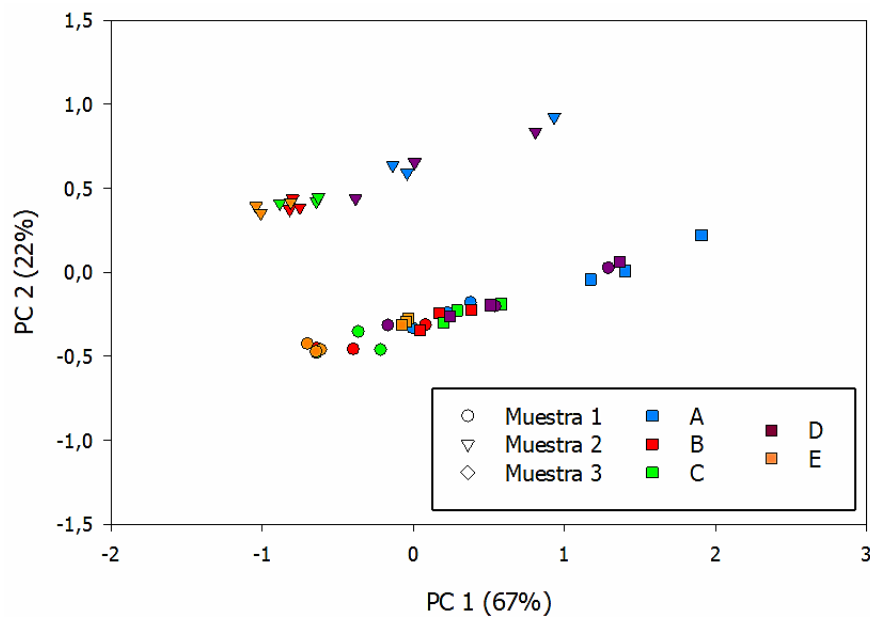


Figura V - 5. Representación de las dos primeras PC de los espectros de las tres muestras estudiadas (89 % de la variabilidad total). Pretratamiento previo de los espectros con SNV, en el intervalo de longitudes de onda de 1100 nm a 2500 nm.

Como se puede observar en la Figura V - 5, los espectros quedan agrupados en dos grupos, debido a que dos muestras presentan una concentración de agua similar. Dentro de estos grupos se observa como las posiciones A y D tienden a diferenciarse de las otras. Sin embargo, dado que las diferencias entre los replicados de una misma posición son mayores que las diferencias observadas entre estas dos posiciones, se puede concluir que no hay diferencias entre posiciones.

Gracias a esta información, se puede deducir que, siempre que la muestra se encuentre dentro de la ventana del haz de luz, pequeñas variaciones en la posición no tienen ningún efecto apreciable en el espectro.

- Temperatura de la muestra: dependiendo de la temperatura del producto durante el análisis de la muestra la migración del agua podría variar afectando al registro del espectro final. Para estudiar si este factor tiene un impacto en la señal NIR, se han registrado los espectros de 3 muestras, después de almacenar cada una de ellas durante 1 hora, a diferentes temperaturas (20°C, 25°C y 30°C).

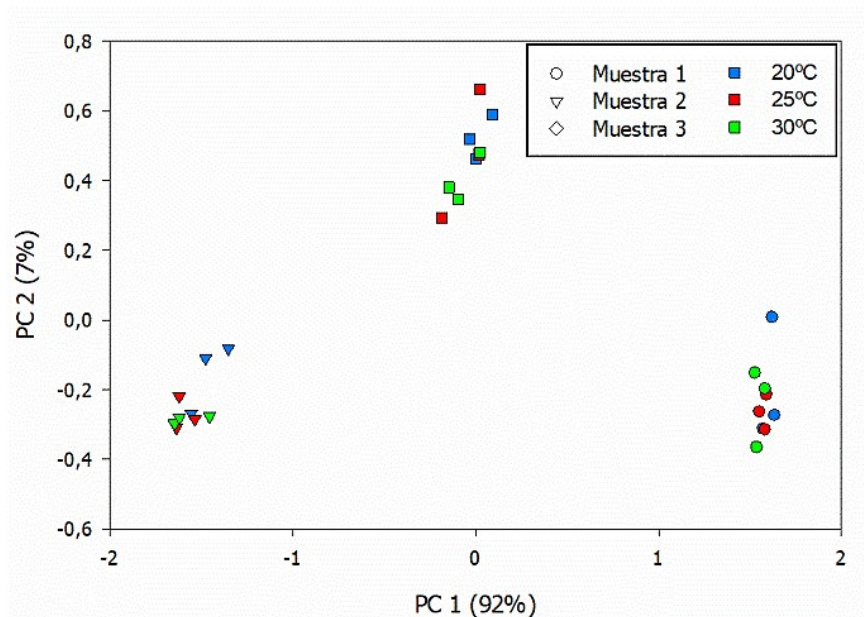


Figura V - 6. Representación de las dos primeras PC de los espectros de las 3 muestras estudiadas (99 % de la variabilidad total). Los espectros se han pretratado con SNV en el intervalo de longitudes de onda de 1100 nm a 2500 nm.

Tal y como se puede observar en la Figura V - 6, los espectros realizados se agrupan en 3 grupos, los cuales están relacionados con las 3 muestras. Dentro de cada grupo no se observa ninguna distribución debida a la temperatura.

Con esta información se puede concluir que cambios en la temperatura de las muestras, debidos a la variación de la temperatura que pueda ocurrir en el ambiente de trabajo, de 20°C a 30°C, no tiene un impacto en el espectro final de las mismas.

- Estabilidad de la muestra: se ha estudiado si los espectros NIR registrados mostraban variaciones a lo largo del tiempo. Para ello, se ha registrado el espectro de 3 muestras, desde el final del ciclo de liofilización hasta un mes después, recogiendo 23 lecturas de éstas.

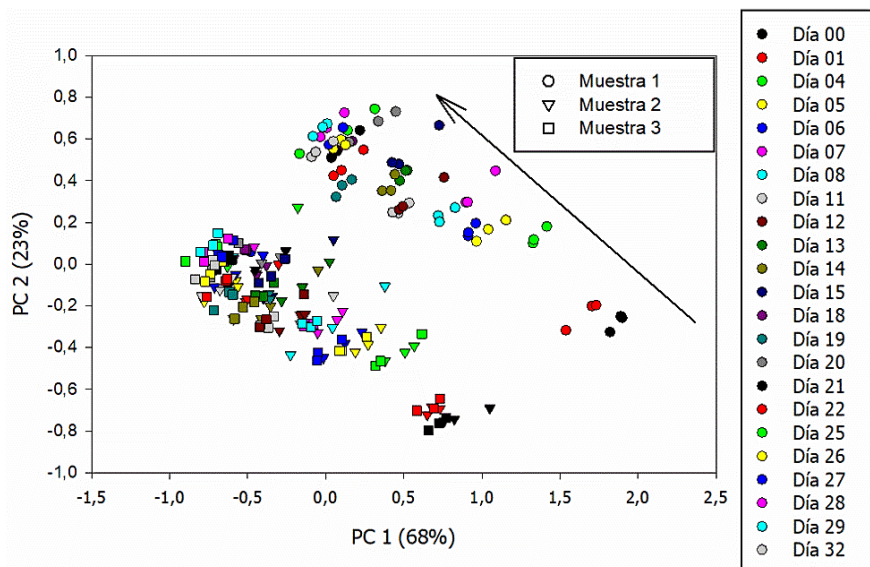


Figura V - 7. Representación de las dos primeras PC de las 3 muestras estudiadas (91 % de la variabilidad total). Los espectros se han pretratado con un SNV, en el intervalo de longitudes de onda de 1100 nm a 2500 nm.

Al realizar un PCA de los espectros (Figura V - 7), los *loadings* de la PC 1 la relacionan con el contenido de agua (banda de 1940 nm). Dos de las muestras tienen espectros prácticamente idénticos y la tercera algo más distinto debido a un contenido de RMC superior. Sin embargo, la evolución a lo largo del tiempo de las tres muestras sigue el mismo patrón, un aumento no constante del contenido de humedad, que se estabiliza a los 25 días después de terminar el ciclo de liofilización. A partir de este tiempo desaparece la tendencia y la dispersión entre los espectros es aleatoria.

Este incremento en la humedad residual del producto puede ser debido a una redistribución de la humedad en el interior del vial después del ciclo de liofilización (humedad proveniente del tapón o redistribución interior de la pastilla).

- **Transmisión de la luz:** dependiendo de la porosidad y el espesor de la muestra, la luz NIR puede transmitirse a través de la pastilla y no reflejarse, causando una pérdida de señal. Para estudiar este efecto en el producto N, se ha escaneado el espectro de una muestra y posteriormente se ha vuelto a registrar el espectro de ésta, pero añadiendo una fina capa de poliestireno en la parte superior de la pastilla.

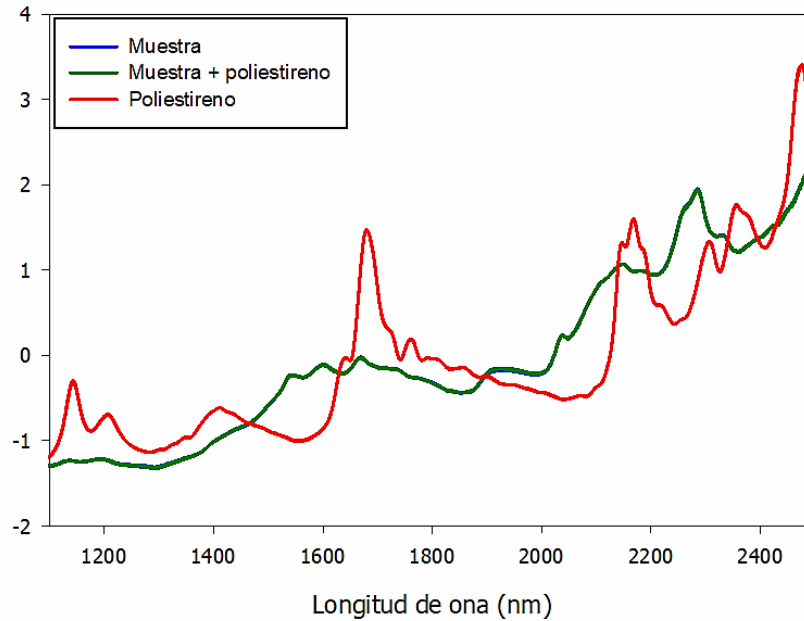


Figura V - 8. Espectros obtenidos de analizar la muestra, el poliestireno y la muestra con poliestireno. Los espectros se han pretratado mediante SNV en el intervalo de longitudes de onda de 1100 nm a 2500 nm.

Como se puede observar en la Figura V - 8, el espectro de la muestra con la fina capa de poliestireno en la pastilla no muestra ninguno de los picos característicos del poliestireno. Esta ausencia de picos del poliestireno indica que la luz NIR no ha llegado a ser transmitida a través de la pastilla y que, por lo tanto, la porosidad y espesor de la muestra es suficiente para que no la atravesase el haz de luz y no haya pérdida de la señal.

- Polimorfismo: la presencia de polimorfos en una muestra puede variar la señal de los espectros NIR. Como se ha descrito en la Tabla V - 1, el producto N contiene manitol, el cual presenta una forma polimórfica llamada hemihidrato. Este polimorfo es metaestable y desproporciona a δ -manitol, por lo que la presencia de éste podría afectar y variar la señal del espectro NIR.¹³⁻¹⁵ Para determinar la presencia de esta forma, se han registrado los difractogramas de dos muestras, una justo el día después de terminar el ciclo y la otra después de un mes de terminar el ciclo.

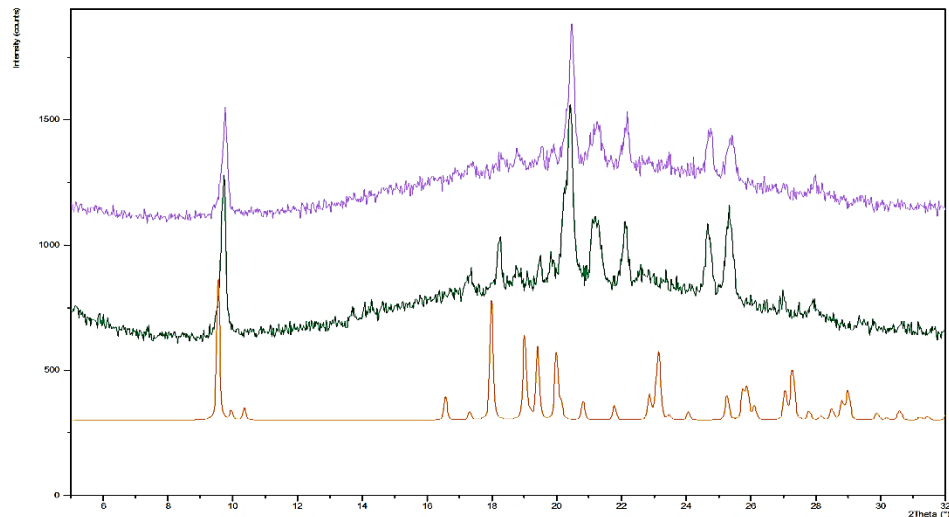


Figura V - 9. Difractogramas de una muestra recién fabricada (en verde) y de una fabricada un mes antes (en morado). En naranja se puede ver el difractograma de un patrón de manitol hemihidrato.

Como se puede observar en la Figura V - 9, ambas muestras tienen el mismo patrón de picos intensos y éstos no coinciden con los picos característicos del hemihidrato. Esto indica que las muestras tienen una parte amorfa y otra cristalina, pero que no hay presencia de hemihidrato, por lo que no hay efecto en el espectro NIR.

- Distribución del agua en la muestra: dependiendo de la distribución del agua en la pastilla, el espectro podría verse afectado. Para estudiar cómo se distribuye el agua en la pastilla del producto N, se han analizado 3 muestras con diferente contenido de agua y del mismo lote. El procedimiento ha consistido en registrar el espectro de las muestras en pastilla y, posteriormente, el de la pastilla desmoronada, es decir, del polvo final.

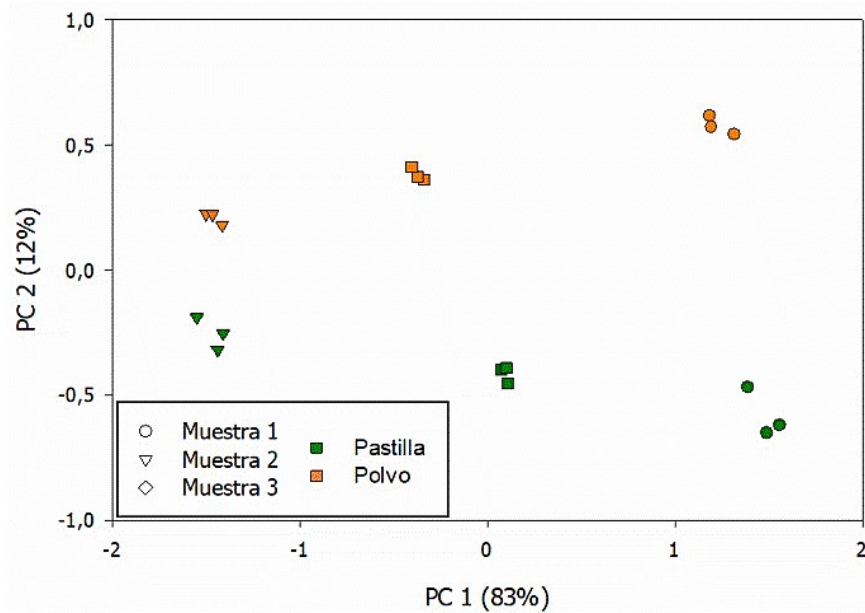


Figura V - 10. Representación de las dos primeras PC (95 % de la variabilidad total) de los espectros de las 3 muestras estudiadas. El pretratamiento realizado consiste en un SNV en el intervalo de longitudes de onda de 1100 nm a 2500 nm.

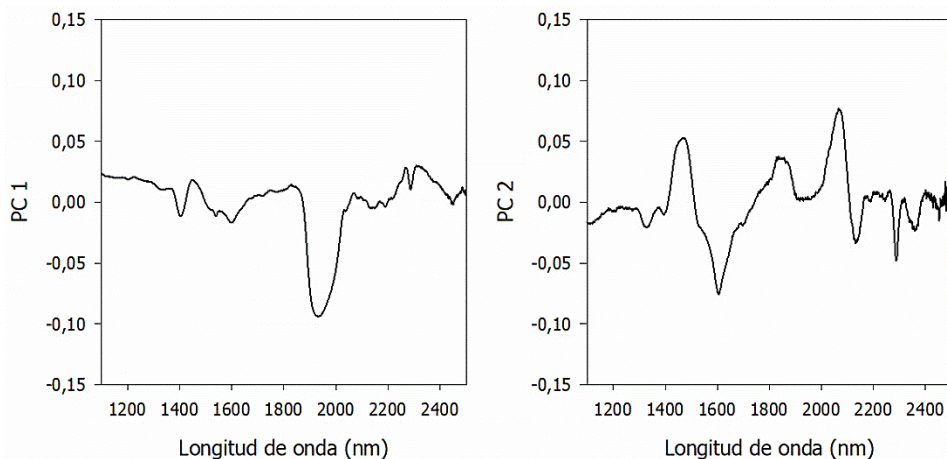


Figura V - 11. Representación de los *loadings* de PC 1 y PC 2.

Como se puede observar en las Figura V - 10, la PC 1 agrupa las muestras por su contenido de agua (Figura V - 11), mientras que la PC 2 agrupa las muestras por el aspecto físico de la pastilla. Los *loadings* de PC 1 muestran que están claramente relacionados con la banda principal del agua a 1940 nm, mientras que los de PC 2 no están relacionados con las bandas de agua (Figura V - 11). Se puede concluir que el agua está distribuida de forma homogénea dentro de la pastilla del producto N.

- Tiempo entre el registro del espectro NIR y el análisis de la muestra con el método de referencia: como se ha visto anteriormente, el espectro de la muestra varía durante el primer mes, sobre todo la banda de absorción asociada al agua.

Por ello, las muestras a las que se les registre el espectro durante ese período deberán ser analizadas el mismo día con el método de referencia. Para las muestras con un tiempo de vida mayor al anterior, se ha estudiado el tiempo máximo que puede transcurrir entre la obtención del espectro y la determinación de agua por el método de referencia. Para ello, se ha registrado el espectro de 4 muestras, fabricadas dos meses atrás, durante 15 días consecutivos.

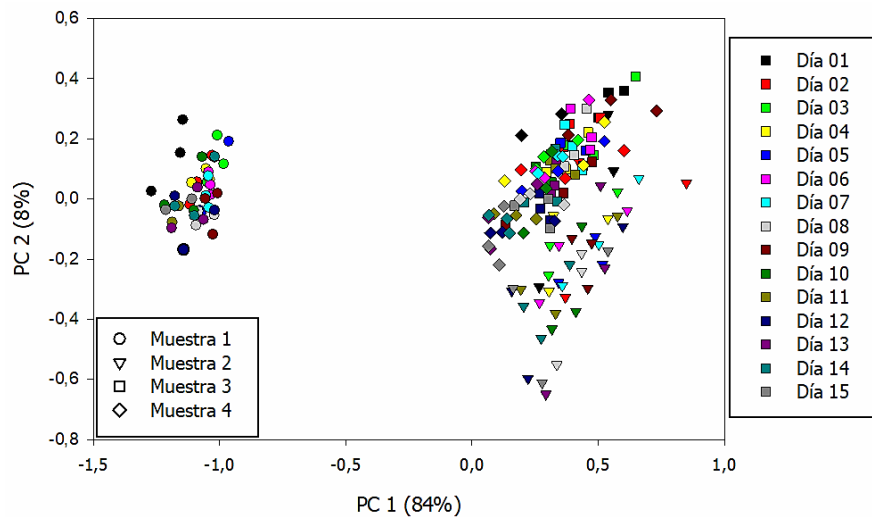


Figura V - 12. Representación de las dos primeras PC (92 % de la variabilidad total) de los espectros de las 4 muestras estudiadas. El pretratamiento realizado consiste en un SNV en el intervalo de longitudes de onda de 1100 nm a 2500 nm.

Los espectros de las muestras analizadas no muestran cambios sistemáticos con el tiempo durante los 15 días del estudio (Figura V - 12). Por lo que, una vez ha pasado el período de los 25 días desde la producción de la muestra, el tiempo entre el registro del espectro y el análisis con el método de referencia puede llegar a ser hasta 15 días como máximo.

Con esta información se recalcula el valor obtenido de estos parámetros en el FMEA (Tabla V - 12).

DETERMINACIÓN DE LA HUMEDAD RESIDUAL EN UN PRODUCTO LIOFILIZADO

Tabla V - 12. Análisis de riesgo de los parámetros cuyo riesgo residual no se había podido definir inicialmente.

Parámetros		Riesgo inicial	Acciones de mitigación	S	P	D	Riesgo residual
Duración del análisis		70	Con la información obtenida, se puede decir que el procedimiento de adquisición del espectro (retirando la muestra después de cada registro o no) no tiene un impacto en éste. Por lo que la exposición prolongada durante 5 análisis no supone un riesgo.	10	1	1	10
Posición de la muestra		160	Con la información obtenida, se ha observado que la posición de la muestra (siempre que ésta se encuentre dentro de la ventana de luz) no tiene un impacto en el espectro.	10	1	1	10
Temperatura de la muestra		350	Con la información obtenida, se ha observado que la muestra puede permanecer hasta 1 h a una temperatura entre 20°C a 30°C sin observarse ningún efecto en el espectro.	10	1	1	10
Estabilización de la muestra		70	Con la información obtenida, se ha observado que durante los 25 primeros días el espectro de la muestra no es estable, y va aumentando su humedad. Para mitigar este riesgo se propone analizar todas las muestras una vez han transcurrido estos primeros 25 días, para así obtener una menor variabilidad en el RMC.	10	1	1	10
Espesor de la muestra		120	Con la información obtenida, se ha observado que el espesor y porosidad del producto N es suficiente para evitar que la luz no sea transmitida a través de la muestra. Por lo que no hay ningún efecto en el espectro.	10	1	1	10
Variaciones físicas	Polimorfismo	90	Con la información obtenida, se ha observado que el polimorfismo no es un riesgo para el espectro del producto N.	10	1	1	10
	Homogeneidad agua	50	Con la información obtenida, se ha observado que el agua se encuentra distribuida de forma homogénea dentro de la pastilla del producto N, por lo que no hay ningún efecto en el espectro.	10	1	1	10
Tiempo entre el registro del espectro NIR y el análisis con el método de referencia		300	Con la información obtenida, se puede decir que el tiempo entre el registro del espectro NIR y el análisis de la muestra con el método de referencia, se deberá hacer el mismo día si ha transcurrido menos de 25 días desde su fabricación, mientras que, si la muestra tiene más de 25 días, la diferencia de tiempo entre los métodos puede ser de 15 días, como máximo.	10	1	1	10

Teniendo en cuenta estos resultados, se puede concluir que se puede desarrollar un modelo cuantitativo NIR para determinar el valor de humedad residual del producto N. Sin embargo, se tienen que considerar las mitigaciones propuestas, con especial atención a los siguientes parámetros:

- Si el tiempo transcurrido entre la finalización del ciclo y el registro del espectro es menor a 25 días, el análisis por el método de referencia deberá hacerse el mismo día que se ha obtenido el espectro.
- Si el tiempo transcurrido entre la finalización del ciclo y el registro del espectro es mayor a 25 días, el análisis por el método de la referencia puede hacerse en periodo inferior a 15 días.

3.2. Objetivo del modelo

Teniendo en cuenta la información aportada en la sección anterior, y como se ha comentado en la introducción, el primer paso para desarrollar un modelo multivariante es definir su objetivo.

El modelo en cuestión se desarrollará para determinar el contenido de agua residual en el producto N descrito en la sección 2.1. Dado que en un producto liofilizado el agua se trata como una impureza, según las guías ICH el intervalo de concentración a cubrir va desde el mínimo posible de RMC hasta el 120 % de la especificación de este CQA, que en este caso es un 5 %. El uso de aplicación del modelo se limita a la determinación de este parámetro en muestras de liberación.

3.3. Desarrollo del modelo cuantitativo NIR

El conjunto de muestras utilizado para la calibración del modelo (conjunto de muestras de calibración) consta de un total de 142 muestras, provenientes de 10 lotes diferentes preparados en la planta piloto y de un lote industrial.

Tal y como se puede observar en la Figura V - 13, las 142 muestras se encuentran distribuidas en el intervalo de RMC comprendido entre 0,36 % y 3,93 %. Esto no cubre las expectativas del objetivo de llegar hasta el 120 % de la especificación, debido a que no se consiguieron producir muestras con esta cantidad de agua, pero el límite superior de RMC cubre las necesidades para el histórico del producto a liberación.

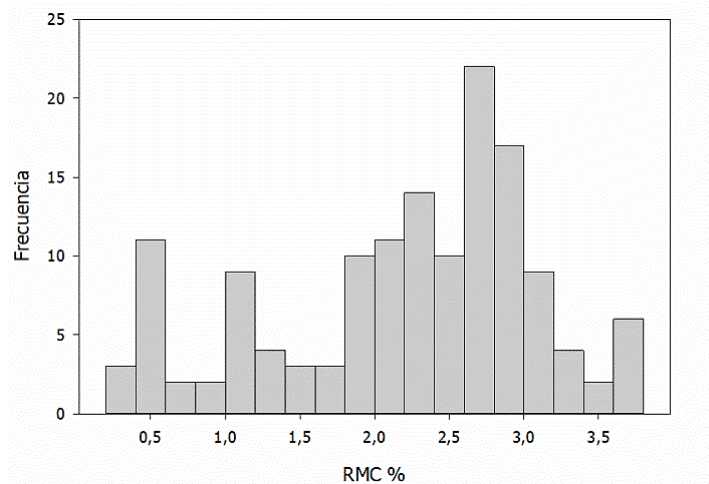


Figura V - 13. Histograma del RMC (%) de las 142 muestras utilizadas para desarrollar el modelo NIR.

El modelo desarrollado se ha construido utilizando el método de PLSR, utilizando validación cruzada como método de evaluación del modelo, con 7 segmentos de 20 muestras cada uno, tomados de forma aleatoria. La selección de los pretratamientos e intervalos de longitudes de onda se ha ido realizando de manera iterativa.

En primer lugar, utilizando la totalidad del intervalo del espectro (1100 nm a 2500 nm) se han evaluado diferentes pretratamientos y combinaciones de éstos para realizar una primera selección de los que consiguen un modelo más simple y con un menor valor de RMSECV. Las derivadas se han calculado con el método de Savitzky-Golay, con un polinomio de orden dos y 11 puntos de ventana móvil. Los resultados se muestran en la Tabla V - 13.

Tabla V - 13. Evaluación de los diferentes pretratamientos y sus combinaciones en el intervalo de 1100 nm a 2500 nm. 1Der y 2Der son las abreviaciones de primera y segunda derivada. R^2 , pendiente y ordenada de origen se refieren a la recta de la concentración de humedad calculada para cada muestra por validación cruzada, frente a la concentración encontrada por el método de referencia.

Pretra- tamientos	Nº de factores	RMSEC	RMSECV	R^2	Pendiente	Sesgo	Ordenada de origen
SNV	4	0,076	0,084	0,991	0,990	9,27E-05	0,022
1Der	4	0,120	0,129	0,980	0,981	6,82E-04	0,043
2Der	6	0,107	0,162	0,968	0,953	-1,81E-03	0,100
SNV+1Der	4	0,073	0,082	0,992	0,991	4,60E-04	0,021
SNV+2Der	6	0,097	0,155	0,971	0,959	9,66E-04	0,092
1Der+SNV	4	0,070	0,077	0,993	0,992	2,13E-04	0,018
2Der+SNV	6	0,099	0,150	0,973	0,963	-4,05E-03	0,077

Como se puede observar en la Tabla V - 13, el uso de únicamente la primera o segunda derivada de Savitzky-Golay proporciona peores resultados que utilizar únicamente SNV o la combinación de SNV y 1ª derivada. El uso de la segunda derivada en combinación con SNV proporciona los modelos más complejos y con peores resultados.

Los mejores pretratamientos se han evaluado en tres intervalos de longitudes de onda diferentes, comentados a continuación:

- Intervalo A: este intervalo es el estudiado en la Tabla V - 13, desde 1100 nm hasta 2500 nm, es decir, todo el espectro NIR de la muestra. Permite detectar más fácilmente muestras anómalas, pero también incluye más ruido en el modelo.
- Intervalo B: va desde 1350 nm hasta 2350 nm eliminando los extremos del espectro, ya que son las zonas que presentan más ruido. Incluye las dos bandas más intensas del agua situadas a 1450 nm y 1940 nm.
- Intervalo C: este intervalo de longitudes de onda recorta el anterior eliminando la zona del espectro en la que se realiza el cambio de detector. Consiste en dos tramos que cubren desde 1350 nm hasta 1750 nm y desde 1850 nm hasta 2350 nm, respectivamente.

Los resultados obtenidos de aplicar los pretratamientos seleccionados a los intervalos B y C, junto con el de todo el espectro (intervalo A) se resumen en la Tabla V - 14.

Tabla V - 14. Evaluación de la aplicación de los pretratamientos seleccionados a los diferentes intervalos de longitudes de onda (I). R², pendiente y ordenada de origen se refieren a la recta de la concentración de humedad calculada para cada muestra por validación cruzada, frente a la concentración encontrada por el método de referencia.

I	Pretra- tamiento	Nº de factores	RMSEC	RMSECV	R ²	Pendiente	Sesgo	Ordenada de origen
A	SNV	4	0,076	0,084	0,991	0,990	9,27E-05	0,022
	SNV+1Der	4	0,073	0,082	0,992	0,991	4,60E-04	0,021
	1Der+SNV	4	0,070	0,077	0,993	0,992	2,13E-04	0,018
B	SNV	3	0,079	0,086	0,991	0,990	-8,16E-04	0,021
	SNV+1Der	3	0,090	0,094	0,989	0,987	-4,65E-04	0,027
	1Der+SNV	3	0,097	0,103	0,987	0,988	-5,47E-04	0,026
C	SNV	3	0,085	0,097	0,988	0,992	-2,35E-03	0,016
	SNV+1Der	3	0,090	0,095	0,989	0,987	4,80E-04	0,029

I	Pretra- tamiento	Nº de factores	RMSEC	RMSECV	R ²	Pendiente	Sesgo	Ordenada de origen
	1Der+SNV	3	0,100	0,106	0,986	0,987	-1,42E-03	0,028

Como se puede observar en la Tabla V - 14, todos los modelos desarrollados en el intervalo A requieren un factor más que los intervalos B y C. Respecto a los intervalos B y C, las diferencias entre pretratamientos son mínimas, aunque el modelo desarrollado únicamente con SNV como pretratamiento es el que presenta los mejores resultados en ambos casos. La utilización de 1ª Derivada más SNV en el espectro completo presenta valores casi iguales de RMSECV y de los parámetros de la recta que la utilización únicamente de SNV con un componente más, por lo que se ha escogido el pretratamiento de SNV en el intervalo B para desarrollar el modelo final.

Se ha estudiado si, para el modelo elegido, existía alguna muestra anómala en el conjunto de calibración que pudiera influenciar en los resultados obtenidos.

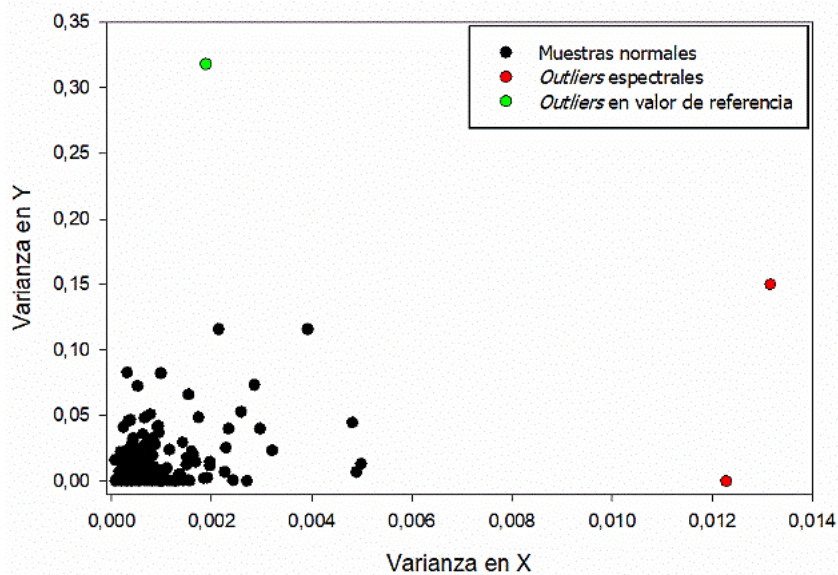


Figura V - 14. Gráfico de varianzas en los espectros (X) y en las concentraciones (Y) del primer factor del modelo. Las muestras en rojo son anómalas en el espectro, y la verde presenta un valor anómalo en el valor de referencia.

Tal y como se puede observar en la Figura V - 14, de las 142 muestras del conjunto de calibración, dos presentan un valor de la varianza en los datos de los espectros NIR (X) mucho mayores que el resto de muestras (muestras en rojo), por lo que se han considerado *outliers* en los espectros y se han eliminado del conjunto de muestras de calibración. En la misma figura, se observa que una muestra con una varianza en las X similar al resto presenta una varianza en las concentraciones (Y) mucho mayor a las demás (muestra en verde). Esto indica que esta muestra seguramente presente un error

en el valor obtenido en el método de referencia, por lo que se considera también como una *outlier*.

El número óptimo de factores, 3, se ha escogido como el necesario para que el error de la validación cruzada (RMSECV) y la varianza explicada de Y se estabilicen (Figura V - 15).

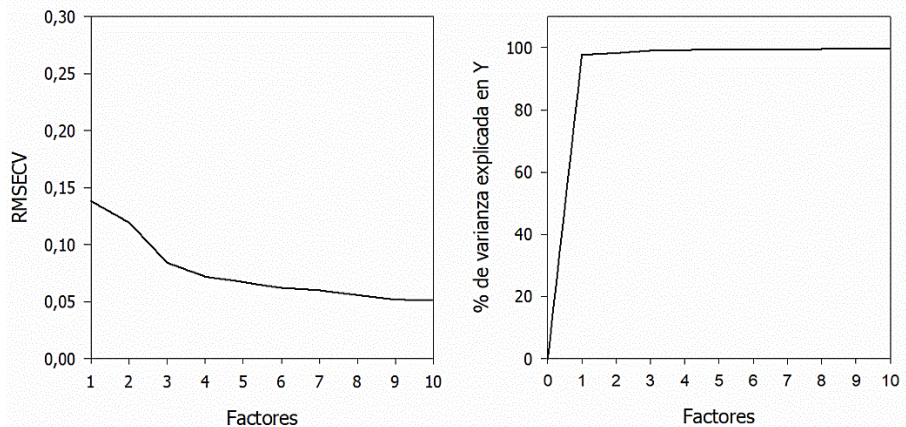


Figura V - 15. A la izquierda se puede observar la variación del valor de RMSECV en función del número de factores del modelo, mientras que a la derecha se observa el aumento del porcentaje de la varianza explicada de Y en función de los factores del modelo.

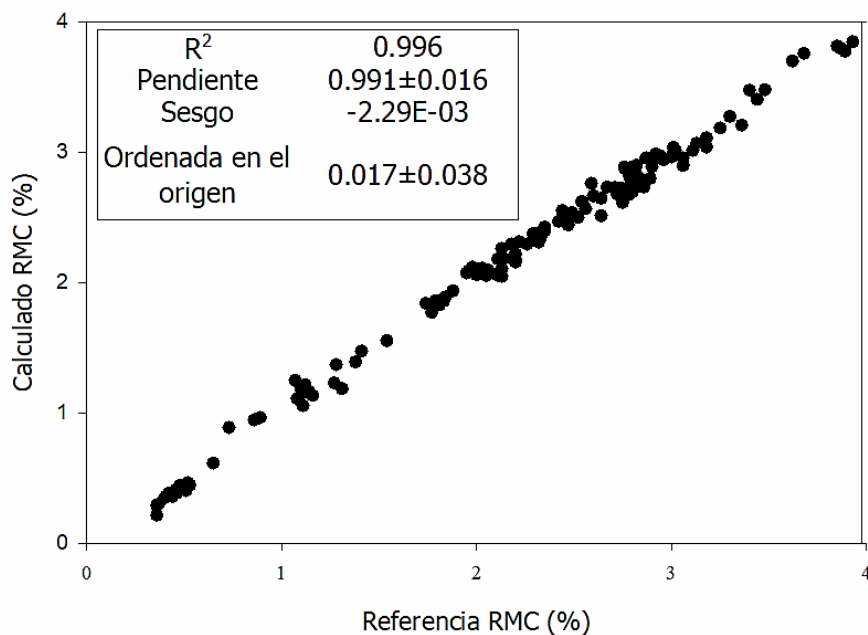


Figura V - 16. Recta de calibración del modelo NIR, con los valores de la pendiente, ordenada en el origen, sesgo y R^2 .

En la Figura V - 16 se muestra, para el conjunto de muestras de calibración, el valor de RMC calculado por validación cruzada con el modelo, frente a los valores de referencia. Tanto la pendiente como la ordenada en el origen incluyen, respectivamente, los valores de 1 y 0 en su intervalo de confianza. El sesgo no es estadísticamente significativo (el

valor de la probabilidad de la prueba t , a un nivel de significación del 95 %, es de 0,750), lo que permite concluir que no hay ningún error sistemático en el modelo. Además, el valor de R^2 es muy próximo a 1 (0,996), lo que indica que los datos de calibración tienen una buena habilidad predictiva. La linealidad de la recta se ha comprobado evaluando estadísticamente la falta de ajuste de la recta; éste ha dado un valor de probabilidad de 0,086 para $\alpha = 0,05$.

En la siguiente tabla se puede ver un resumen del modelo desarrollado y que se validará en la siguiente sección.

Tabla V - 15. Resumen de los principales parámetros del modelo NIR desarrollado.

Parámetro	Resultado	Parámetro	Resultado
Nº de muestras	139	RMSEC	0,077
Muestras industriales	9	RMSECV	0,084
Intervalo de RMC (%)	0,36 a 3,93	R^2	0,996
Intervalo de longitudes de onda	1350 nm a 2350 nm	Pendiente	0,991±0,016
Pretratamiento	SNV	Ordenada de origen	0,017±0,038
Nº de factores PLSR	3	Sesgo	-2,29E-03
% varianza explicada de Y	99,14	Test de linealidad	Lineal

3.4. Validación del modelo cuantitativo

Para validar el modelo cuantitativo NIR se ha utilizado un conjunto de 84 muestras, independientes del empleado durante el desarrollo del modelo. Estas muestras provienen de 4 lotes diferentes preparados en la planta piloto y de un lote industrial.

Tal y como se puede observar en la Figura V - 17, las 84 muestras se encuentran distribuidas en el intervalo de RMC comprendido entre 0,37 % y 3,78 %, cubriendo de esta manera todo el intervalo de concentraciones del modelo a validar (0,36 % a 3,93 %). En las siguientes secciones se comentan los parámetros evaluados durante la validación del modelo.

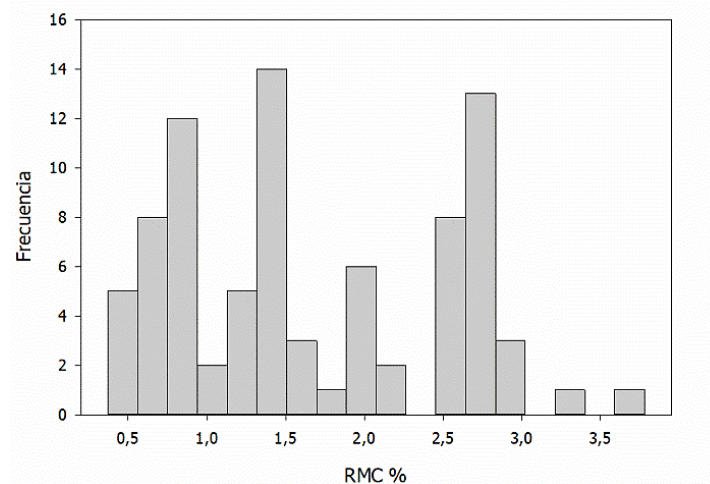


Figura V - 17. Histograma del RMC (%) de las 84 muestras utilizadas para validar el modelo cuantitativo

3.4.1. Especificidad

La especificidad del modelo para asegurar que el agua es la propiedad modelada se ha realizado evaluando los *loadings* y los *scores* del modelo.

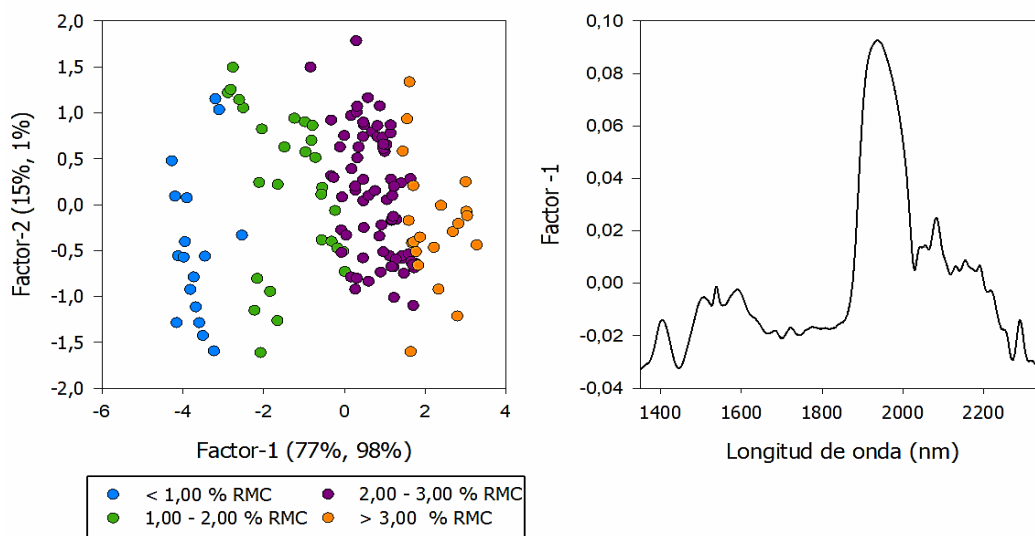


Figura V - 18. A la derecha se puede observar los *loadings* del primer factor del modelo PLSR, con un máximo en la banda de mayor intensidad del agua, y a la izquierda el gráfico de los *scores*, donde se puede observar que el factor 1 está relacionado con el contenido de RMC (%).

Tal y como se observa en Figura V - 18, los *loadings* con un mayor peso en el factor 1, que explica el 98 % de la variabilidad en Y, se observan en las longitudes de onda relacionadas con la banda de mayor intensidad del agua (1940 nm). A la izquierda en la misma figura, se observa como las muestras quedan distribuidas a lo largo del factor 1 según su concentración de RMC (%).

Durante la predicción de las muestras del conjunto de validación no se detectó ninguna muestra anómala y los espectros de las muestras quedaron proyectados en el modelo, formando un único grupo.

3.4.2. Linealidad

Para evaluar este parámetro se ha realizado un análisis de regresión de los valores de RMC calculados con el modelo NIR, frente a los valores obtenidos con el método de referencia, para las 84 muestras descritas anteriormente.

Tabla V - 16. Resumen de los parámetros evaluados para determinar la linealidad del modelo.

R ²	Pendiente	Ordenada en el origen	Sesgo	Probabilidad de la falta de ajuste
0,990	1,017±0,020	0,010±0,037	0,037	0,342

Como se puede observar en Tabla V - 16, el valor del coeficiente de determinación es próximo a 1 y, tanto la pendiente como la ordenada en el origen incluyen en sus intervalos de confianza el 1 y el 0, respectivamente. Referente a la falta de ajuste, el valor de probabilidad obtenido en la prueba *t* es superior al 0,05, por lo que no se considera estadísticamente significativo, demostrando así la linealidad del intervalo de concentraciones del modelo.

3.4.3. Intervalo de linealidad

El intervalo de linealidad en el cual se ha demostrado que el modelo es lineal (sección 3.4.2), exacto (sección 3.4.4) y preciso (sección 3.4.5) comprende entre 0,37 % y 3,93 %. Por lo que cualquier valor fuera del intervalo comentado deberá expresarse como <0,37 % o >3,93 % sin indicar el valor numérico.

3.4.4. Exactitud

Desde el valor mínimo de RMC al máximo, se ha tomado aleatoriamente una muestra a intervalos de 0,1 %; en total 25 muestras. Se ha realizado una prueba *t* de muestras apareadas, comparando la concentración calculada con el modelo NIR y la encontrada con el método de referencia, para un nivel de confianza del 95 %. El valor obtenido para la *t* experimental ha sido de 1,721, menor que el valor de la *t* crítica bilateral (2,064). Esto indica que no hay diferencias significativas entre ambos métodos.

El valor de RMSEP obtenido en el cálculo de la linealidad, con 84 muestras, es de 0,086, un valor similar a los obtenidos por el RMSEC (0,077) y RMSECV (0,084) obtenidos para

el conjunto de calibración (Tabla V - 15) y al valor del SEL del método de referencia (0,22 descrito en la sección 1.2.2). Además, este valor del RMSEP cumple con el criterio propuesto en la guía de la EMA sobre modelos NIR donde se comenta que el RMSEP debe ser menor o igual a 1,4 por el SEL.⁸

3.4.5. Precisión

Teniendo en cuenta que el objetivo del modelo no contempla su uso en otro laboratorio, se han evaluado solamente los niveles de precisión intermedia y repetibilidad.

- Precisión intermedia: para evaluar este parámetro se han utilizado 2 muestras una de un lote producido en la planta piloto y otra del lote industrial. Las muestras han sido analizadas 6 veces por dos analistas y en dos días distintos (Tabla V - 17).

Tabla V - 17. Precisión intermedia. Resultados del RMC (%) calculados por el modelo al reanalizar ambas muestras 6 veces en dos días distintos y por dos analistas distintos (A1 y A2).

Rep.	Muestra laboratorio				Muestra industrial			
	Día 1		Día 2		Día 1		Día 2	
	A1	A2	A1	A2	A1	A2	A1	A2
1	1,53	1,50	1,50	1,50	1,90	1,91	1,91	1,91
2	1,49	1,46	1,47	1,48	1,87	1,82	1,89	1,88
3	1,48	1,46	1,48	1,47	1,87	1,83	1,87	1,86
4	1,49	1,46	1,47	1,47	1,85	1,85	1,86	1,85
5	1,48	1,46	1,46	1,46	1,89	1,83	1,85	1,85
6	1,47	1,47	1,46	1,46	1,84	1,82	1,84	1,84
Promedio	1,49	1,47	1,47	1,47	1,87	1,87	1,84	1,87
Des. Est.	0,021	0,016	0,015	0,015	0,023	0,026	0,034	0,026
RSD (%)	1,41	1,09	1,02	1,02	1,22	1,39	1,87	1,39

Para saber si hay diferencias entre las variables días, analistas o la interacción entre ambos se realiza un ANOVA de dos factores. Los resultados, resumidos en la Tabla V - 18, muestran que el valor de la probabilidad es superior a 0,05 en todos los casos, al igual que la $F_{exp} < F_{crítica}$, por lo que se puede decir que no hay diferencias entre las variables estudiadas.

Tabla V - 18. Resultados de la ANOVA de dos factores de las dos muestras estudiadas.

Parámetro estadístico	Muestra laboratorio			Muestra industrial		
	Probabilidad	F _{exp}	F _{crítica}	Probabilidad	F _{exp}	F _{crítica}
Día	0,409	0,710		0,349	0,921	
Analista	0,133	2,449	4,351	0,176	1,968	4,351
Interacción	0,133	2,449		0,349	0,921	

- Repetibilidad: para evaluar este parámetro se ha utilizado el mejor y el peor resultado de 6 repeticiones obtenidas durante la precisión intermedia en un mismo día y analista (Tabla V - 17). Los valores de RSD (%) mínimos y máximos para la repetibilidad del modelo son de 1,02 % y 1,87 % respectivamente.

3.4.6. Límite de detección y cuantificación

Teniendo en cuenta que la aplicación final de este modelo es en el ámbito farmacéutico y siguiendo las normativas que rigen sobre este sector, se ha considerado como límite de detección y cuantificación el valor mínimo obtenido durante el desarrollo y la validación del modelo. Por lo que se puede considerar como LOD y LOQ 0,36 %.¹⁶

Sin embargo, considerando el estudio realizado por G. Clua-Palau *et al.*, estos resultados, aunque son válidos desde un punto de vista regulatorio, no son realistas desde la perspectiva académica.¹⁷ Dado el doble carácter industrial-académico de este trabajo, se han calculado tanto el LOD como el LOQ empleando la aproximación multivariante presentada por Allegrini y Olivieri (Tabla V - 19). Esta aproximación se basa en la estimación de un límite de detección multivariable expresado como un intervalo de LOD (LOD_{min} y LOD_{max}), que depende de la variabilidad de las muestras de calibración en el hiperplano de n-dimensiones. Esta aproximación también considera los errores de medida en las variables independientes en modelos de regresión PLS. De este modo, el intervalo LOD se convierte en un parámetro que caracteriza el modelo de calibración PLS general, y no cada muestra en particular.

Tabla V - 19. Valores del LOQ y LOD calculados según la aproximación multivariante presentada por Allegrini y Olivieri.

LOQ		LOD	
LOQ _{mín.}	LOQ _{máx.}	LOD _{mín.}	LOD _{máx.}
0,21 %	0,34 %	0,07 %	0,11 %

3.4.7. Robustez

Aunque los parámetros que podían afectar a la señal NIR ya han sido evaluados y mitigados durante el estudio de viabilidad, el estudio de la robustez se ha planteado llevar a cabo evaluando los resultados obtenidos de los lotes en rutina cuando estén disponibles.

3.4.8. Conclusiones de la validación del modelo

Teniendo en cuenta los resultados de las secciones anteriores, se puede concluir que el modelo descrito en la Tabla V - 15 es válido para cuantificar el contenido de agua residual en el producto N. La validación del modelo se ha realizado siguiendo las directrices de las distintas agencias reguladoras del sector farmacéutico, dando conforme en todos los parámetros evaluados, por lo que puede ser utilizado para aplicaciones dentro de Reig Jofre. En la siguiente sección se menciona una de estas.^{7,8,16}

3.5. Aplicación del modelo

Tal y como se ha comentado en secciones anteriores, la principal aplicación del modelo NIR desarrollado consiste en analizar las muestras de los lotes de PPQ del producto N, con el propósito de estudiar la homogeneidad del RMC, tanto dentro de un mismo lote, como entre distintos lotes y liofilizadores.

Con el fin de poder evaluar este CQA, el número de muestras escogido para analizar se ha basado en la ISO-3951-1:2012, con un nivel de inspección II, para una inspección normal para un tamaño del lote de 46875 viales.¹⁸ Para poder cumplir con este valor, se han estudiado 5 posiciones de cada piso del liofilizador, adquiriendo 6 muestras por cada posición (la descripción de las posiciones se puede ver en la Figura V - 19).

Durante la predicción no se detectó ninguna muestra anómala y los espectros de las muestras de los tres lotes de PPQ quedaron proyectados en el centro del modelo, formando un único grupo, indicando que la variabilidad industrial está incluida en el modelo.

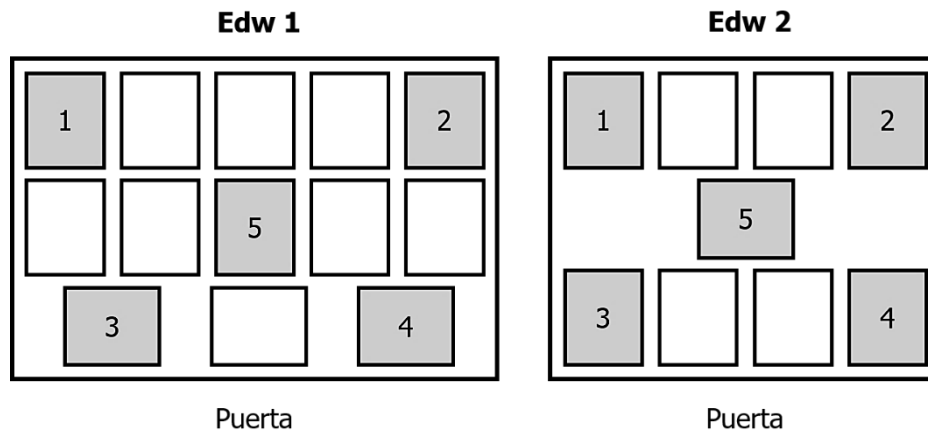


Figura V - 19. Esquema de las posiciones estudiadas para los dos liofilizadores utilizados para la validación del proceso del producto N.

En la Tabla V - 20 se puede ver un resumen estadístico de los resultados calculados con el modelo de estos tres lotes.

Tabla V - 20. Estadística descriptiva de los valores del mapeo obtenidos con el modelo y valores del método de referencia del lote.

Lote	Liofilizador	Nº de muestras	Media	Des. estándar	Mínimo	Máximo	Valor de KF	Des. estándar
Lote 1	Edw 2	420	0,80	0,14	0,53	1,54	0,70	0,03
Lote 2	Edw 1	300	0,93	0,16	0,59	1,45	0,80	0,15
Lote 3	Edw 2	420	0,84	0,14	0,62	1,54	0,70	0,09

Como se puede observar, todos los valores obtenidos son inferiores a la especificación máxima del producto (5 %) y los valores obtenidos con el método de referencia se encuentran en todos los casos, dentro del intervalo obtenido por el método NIR.

Para estudiar la homogeneidad de los liofilizadores, se realizó un ANOVA entre los distintos lotes, obteniendo un valor de probabilidad $<0,001$ inferior al 0,05. Esto indica que hay diferencias estadísticas significativas entre los lotes. Como se puede observar en la Figura V - 20, los lotes 1 y 3 presentan un valor medio de humedad menor que el del lote 2. Esto se puede deber a que estos dos lotes fueron liofilizados en el liofilizador Edw 2, el cual es más nuevo y como indican los resultados, tiene una mayor capacidad de secado que el liofilizador Edw 1.

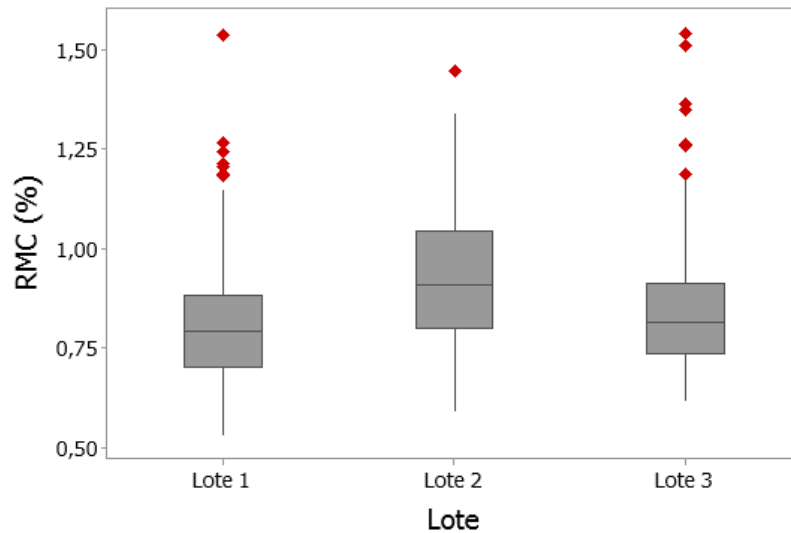


Figura V - 20. Gráfico de cajas de los tres lotes, en los que se puede observar la media (línea central de las cajas), el 1º cuartil representado con la parte inferior de las cajas y el 3º cuartil representado por la parte superior de estas. En rojo se pueden observar las muestras anómalas.

Puesto que los lotes no son estadísticamente iguales entre ellos, la evaluación entre pisos y posiciones se ha realizado de forma individual. El estudio de la distribución de la humedad en cada uno de los lotes, por la prueba de Kolmogorov-Smirnov, indica que siguen una distribución no normal, por lo que se realizó una transformación de Box-Cox para su normalización y un ANOVA de dos vías en cada lote. Por la falta de homogeneidad de varianzas, se realizó la prueba de Kruskal-Wallis para el factor "piso" y el factor "posición", que equivale a un análisis de varianza no paramétrico.

En los tres lotes, se halló que el valor de la humedad media en alguno de los pisos era distinto al del resto, mientras que únicamente en el lote 1 se encontraron diferencias en el valor de la humedad en alguna de las posiciones del mismo piso.

Teniendo en cuenta que todos los valores cumplen con la especificación, se ha decidido realizar una representación 3D de los liofilizadores para evaluar cualitativamente si existe algún patrón sistemático en la posición de las muestras relacionada con el contenido de humedad que se muestra en los gráficos de cajas.

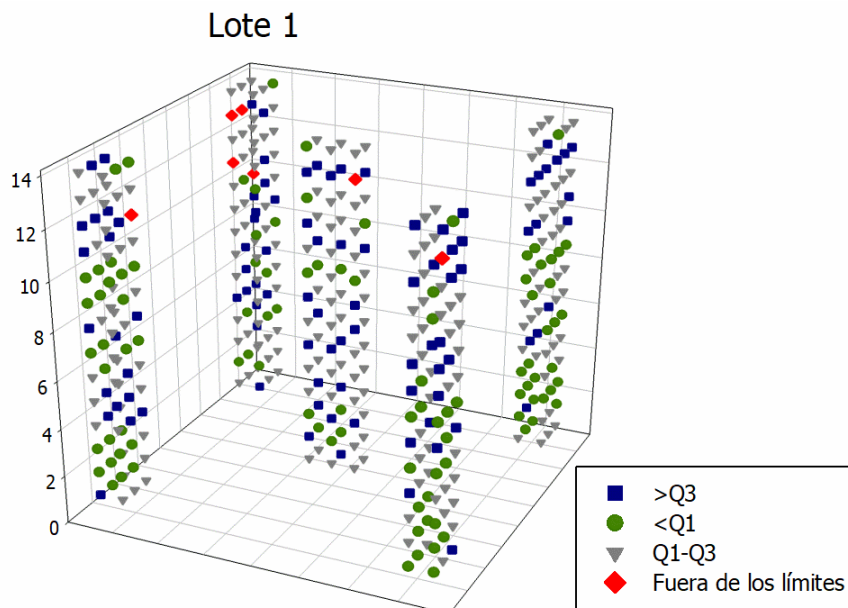


Figura V - 21. Distribución 3D de las muestras del mapeo dentro del liofilizador del lote 1, clasificadas por colores dependiendo de en qué cuartil se encuentren. Q1 y Q3 indican los cuartiles 1 y 3, respectivamente.

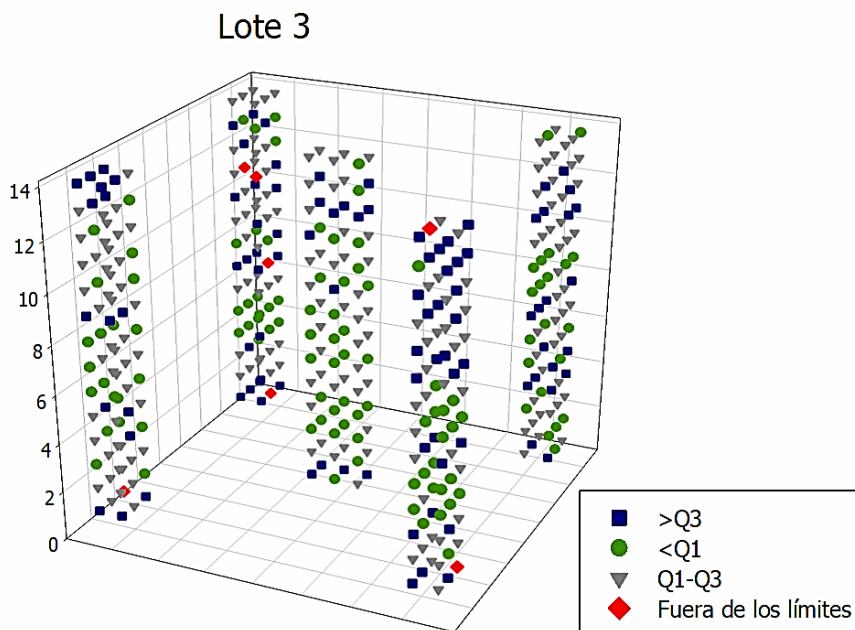


Figura V - 22. Distribución 3D de las muestras del mapeo dentro del liofilizador del lote 3, clasificadas por colores dependiendo de en qué cuartil se encuentren. Q1 y Q3 indican los cuartiles 1 y 3, respectivamente.

Como se puede observar en la Figura V - 21 y Figura V - 22, los dos lotes liofilizados en el Edw 2 muestran una distribución aleatoria de las muestras en los distintos cuartiles y no se observa que las muestras que se encuentran fuera de los límites estén localizadas en unos pisos o posiciones concretas del liofilizador. Esto indica que el proceso de liofilización es suficientemente homogéneo dentro de la cámara.

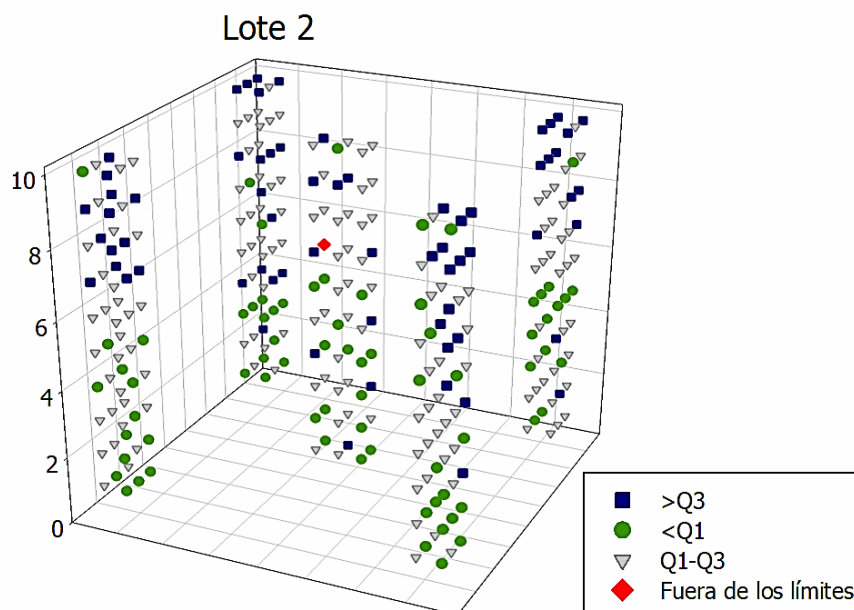


Figura V - 23. Distribución 3D de las muestras del mapeo dentro del liofilizador del lote 2, clasificadas por colores dependiendo de en qué cuartil se encuentren. Q1 y Q3 indican los cuartiles 1 y 3, respectivamente.

Como se puede observar en la Figura V - 23, el único lote liofilizado en el Edw 1 muestra una distribución de las muestras clasificadas por cuartiles con una leve tendencia de agrupación de las muestras más secas (<Q1) en los pisos 1-6 y de las muestras más húmedas (>Q3) en los pisos 7-10. Sin embargo, dado que sólo se dispone de un lote y que en la mayoría de los pisos hay muestras de los distintos cuartiles, el proceso de liofilización es también suficientemente homogéneo.

4. CONCLUSIONES

En este trabajo, se ha desarrollado y validado de forma satisfactoria un modelo cuantitativo NIR que permite calcular la humedad residual del producto N a liberación, utilizando PLSR.

Previamente, se ha evaluado el riesgo en el desarrollo del modelo empleando un análisis modal de fallos y efectos. De este análisis se han detectado dos riesgos potenciales proponiendo dos acciones de mitigación para solventarlos y evitar que tengan un impacto en el proceso.

Los espectros de los replicados de cada muestra se han promediado y tratado con un pretratamiento de SNV en el intervalo de longitudes de onda de 1350 nm a 2350 nm. Se han utilizado 139 muestras en el conjunto de calibración y se ha utilizado el método de validación cruzada para optimizar y evaluar el modelo con 3 factores PLSR.

El modelo se ha validado con 84 muestras independientes de las de calibración y distribuidas en el intervalo de concentraciones de 0,37 % a 3,78 %. Los resultados de la validación son conformes en todos los parámetros, con un valor de RMSEC de 0,077 y un RMSEP de 0,086.

Este modelo se ha aplicado en el ámbito industrial para evaluar la homogeneidad del proceso de liofilización. Para ello se ha calculado el valor de RMC de 1140 muestras y se han evaluado las diferencias entre lotes y dentro de un mismo lote.

Aunque todos los resultados analizados se encontraron dentro de la especificación (<5 %), al realizar un ANOVA entre lotes se observaron diferencias significativas entre las muestras provenientes de ambos liofilizadores industriales (Edw 1 y Edw 2), con una menor humedad residual media en el liofilizador adquirido más recientemente. En el estudio de cada lote para evaluar las diferencias entre pisos y posiciones, se han observado diferencias significativas en los resultados de RMC de algún piso respecto al resto, aunque únicamente en uno de los lotes había diferencias en el secado en una de las posiciones respecto a las otras. A pesar de ello, no se observa una sistemática que relacione pisos y posiciones concretas con el secado en el liofilizador.

5. **BIBLIOGRAFÍA**

- (1) McCoy, T.; Affleck, R.; Khamar, D. Considering Residual Moisture Measurement in Lyophilized Drug Product. *Am. Pharm. Rev.* 2019.
- (2) Cook, I. A.; Ward, K. R. Headspace Moisture Mapping and the Information That Can Be Gained about Freeze-Dried Materials and Processes. *PDA J. Pharm. Sci. Technol.* 2011, 65 (5), 457–467.
- (3) Clavaud, M.; Lema-Martinez, C.; Roggo, Y.; Bigalke, M.; Guillemain, A.; Hubert, P.; Ziemons, E.; Allmendinger, A. Near-Infrared Spectroscopy to Determine Residual Moisture in Freeze-Dried Products: Model Generation by Statistical Design of Experiments. *J. Pharm. Sci.* 2020, 109 (1), 719–729.
- (4) Flórez-Borges, P. Desarrollo de Un Liofilizado Oral Antihistamínico de Segunda Generación, Universidad de Barcelona, 2016.
- (5) Baheti, A.; Kumar, L.; Bansal, A. K. Excipients Used in Lyophilization of Small Molecules. *J. Excipients Food Chem.* 2010, 1 (1), 41–54.
- (6) European Pharmacopoeia. 2.5.32. Water Micro Determination. 2019, No. 1, 20532.
- (7) The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. *ICH Q2 (R1) Validation of Analytical Procedures: Text and Methodology*, 2011; Vol. 20.
- (8) European Medicines Agency. Guideline on the Use of near Infrared Spectroscopy

-
- by the Pharmaceutical Industry and the Data Requirements for New Submissions and Variations. 2014, p 28.
- (9) European Medicines Agency (EMA). ICH Guideline Q9 on Quality Risk Management. 2014, pp 1–20.
- (10) Siesler, H. W.; Ozaki, Y.; Kawata, H.; Heise, H. . *Near-Infrared Spectroscopy Principles, Instruments, Applications*; 2006.
- (11) European Pharmacopoeia. 2.2.40. *Near-Infrared Spectroscopy*; 2019.
- (12) USP. <1119> *NEAR-INFRARED SPECTROSCOPY*; 2010.
- (13) Nunes, C.; Suryanarayanan, R.; Botez, C. E.; Stephens, P. W. Characterization and Crystal Structure of D-Mannitol Hemihydrate. *J. Pharm. Sci.* 2004, *93* (11), 2800–2809.
- (14) Cao, W.; Xie, Y.; Krishnan, S.; Lin, H.; Ricci, M. Influence of Process Conditions on the Crystallization and Transition of Metastable Mannitol Forms in Protein Formulations During Lyophilization. *Pharm. Res.* 2013, *30* (1), 131–139.
- (15) Dierks, T. M.; Korter, T. M. Origins of the Relative Stabilities of Anhydrous and Hydrated α -Mannitol Crystals. *J. Phys. Chem. A* 2016, *120* (33), 6629–6636.
- (16) European Pharmacopoeia. 5.21. Chemometric Methods. 2016, 5641–5658.
- (17) Clua-Palau, G.; Jo, E.; Nikolic, S.; Coello, J.; Maspoch, S. Finding a Reliable Limit of Detection in the NIR Determination of Residual Moisture in a Freeze-Dried Drug Product. *J. Pharm. Biomed. Anal.* 2020, *183*, 113163.
- (18) Asociación Española de Normalización y Certificación. *Procedimientos de Muestreo Par Ainspección Por Variables*; España, 2012; p 111.

VI. DETERMINACIÓN DE OXÍGENO EN EL
HEADSPACE DE UN PRODUCTO
LIOFILIZADO

1. INTRODUCCIÓN

Un elevado contenido de humedad o de oxígeno son las principales causas que pueden afectar a la estabilidad de los productos farmacéuticos. En capítulos anteriores, se ha comentado la influencia de la humedad en los productos liofilizados y se han desarrollado métodos para su determinación; en este capítulo, se estudia el impacto del oxígeno.

Las formulaciones de productos farmacéuticos liofilizados que contienen biomoléculas o moléculas orgánicas de gran tamaño acostumbran a presentar una alta sensibilidad al oxígeno. Esto puede tener algunos efectos negativos en la calidad del producto final, por ejemplo, una disminución de la eficacia, la formación de subproductos de incluso alta toxicidad o el cambio de coloración del producto.^{1,2} Para proteger estas formulaciones, la concentración de gas debe ser disminuida antes de concluir el ciclo de liofilización y de cerrar los viales. Por esta razón, en el caso de productos parenterales liofilizados, se cierran los viales directamente con una baja presión en su interior (correspondiente a la presión del secado secundario) o se introduce en la cámara de liofilización una atmósfera inerte (ausencia de vapor de agua y oxígeno) para realizar una rotura parcial del vacío. De esta manera, se controla la composición del *headspace* de los viales de todo el lote, reduciendo al mínimo la presencia de oxígeno.^{3,4}

Para mantener estas condiciones de atmósfera inerte a lo largo del tiempo es importante escoger un buen acondicionado (vial, tapón y cápsula), haciendo el sistema impermeable a la entrada de aire del exterior. A pesar de esto, existe la posibilidad de que la atmósfera inerte del producto se vea afectada, como cuando se producen fugas debidas a un defecto en el material o a cualquier problema durante el cierre del vial. El primer caso implica un flujo constante de gases hacia el interior del producto, mientras que el segundo se acostumbra a solventar con el cierre mediante el capsulado del vial, manteniendo su hermeticidad y esterilidad.⁵⁻⁷

Una forma no destructiva de cuantificar la concentración de oxígeno presente en el *headspace* de los productos, y saber si se ha mantenido o no el ambiente de atmósfera inerte, es utilizar la espectroscopia láser TDLAS-FMS. Durante la validación del proceso de fabricación, se puede comprobar el contenido de oxígeno mediante un muestreo extensivo que puede llegar a ser de cientos de viales. Aunque es una medida sencilla, requiere un esfuerzo de tiempo y personal.

Aunque cumpla con todas las especificaciones, el producto liofilizado que se estudia en este capítulo puede mostrar una coloración rosada debida a la presencia de oxígeno. A lo largo del tiempo, esta coloración puede aumentar, pudiendo aparecer en una misma caja viales con distinta coloración del producto. La aparición del color en el sólido sugiere que se puede intentar relacionar la concentración de oxígeno en el vial con medidas espectroscópicas de la coloración del producto liofilizado. En este trabajo, se estudia la posibilidad de aprovechar la información contenida en el espectro visible-NIR, que se registra para la determinación el contenido de humedad residual del sólido liofilizado, para relacionarla con el contenido de oxígeno en el *headspace*. De esta manera, con un mismo espectro se puede obtener el doble de información, con la consiguiente reducción del tiempo de análisis.^{8,9}

El objetivo final del estudio es el desarrollo de un modelo que, a partir del espectro visible-NIR, permita saber qué productos pueden desarrollar una mayor coloración con el tiempo, con el fin de evitar tener muestras con distinto tono de color y homogenizar el aspecto del lote.

2. MATERIALES Y MÉTODOS EXPERIMENTALES

2.1. Muestras

Las muestras estudiadas consisten en un producto parenteral liofilizado, producto J, que contiene 1 g de API y no presenta excipientes en su formulación. Una vez liofilizado, el producto tiene un aspecto de polvo blanco o ligeramente rosa, envasado en un vial incoloro de tubo de 20 ml, con un tapón de 20 mm de diámetro de clorobutilo y una cápsula de aluminio con tapa.

Para el desarrollo de este estudio se han utilizado lotes de validación de proceso, realizados entre 2018 y 2019 para el mapeo de la humedad, y lotes de producción liofilizados en los liofilizadores Edw 1 y Edw 2.

2.2. Determinación de la absorbancia del producto disuelto

La determinación del grado de coloración del producto J, se basa en la espectroscopia de absorción visible y, para ello, se ha utilizado el espectrofotómetro UVVis LAMBDA 25. Primero se realiza el blanco del equipo, utilizando dos cubetas de cuarzo de 1 cm de camino óptico con agua. A continuación, se disuelve un vial con 10 ml de agua,

obteniendo una concentración de 100 mg/ml, y se realiza la medida de absorbancia a las longitudes de onda de 370 nm y 450 nm. La especificación para cada longitud de onda es que la absorbancia debe ser <0,65 y <0,1 respectivamente.

2.3. Método de referencia

El método analítico empleado para la cuantificación del oxígeno en el *headspace* de las muestras del producto J, se basa en la técnica de espectroscopia de absorción láser de diodo sintonizable (TDLAS-FMS).

El procedimiento consiste en realizar el centrado de la señal del láser, utilizando los patrones de concentración de oxígeno del mismo formato que el vial del producto (viales incoloros de tubo de 20 ml) y, seguidamente, establecer la recta de calibración. A continuación, se introduce la muestra en el compartimento y se realiza la medida. El valor obtenido consiste en la concentración de oxígeno (en %), promedio de las medidas realizadas al rotar la muestra 360° en el eje vertical.

2.4. Adquisición de los espectros

Los espectros de las muestras utilizadas se han registrado empleando el espectrofotómetro NIR Labspec500, descrito anteriormente. El procedimiento se basa en el estudio de viabilidad realizado por la compañía durante el desarrollo del modelo NIR para la determinación de RMC del producto J. Consiste en colocar la parte inferior del vial encima de la fuente de luz del accesorio ASD Muglight, y registrar por triplicado su espectro, realizando una rotación de 120° entre cada uno y retirando la muestra de la fuente de luz después de cada registro.⁸

Previamente al desarrollo o aplicación del modelo, los tres espectros de cada muestra, registrados en unidades de reflectancia, se transforman a unidades de absorbancia y se promedian, con el fin de tener un único espectro por muestra.

3. RESULTADOS Y DISCUSIÓN

3.1. Estudios previos

Durante los estudios de estabilidad del producto J, se observa que algunos viales presentan una coloración más rosácea que otras muestras del mismo lote. En estudios iniciales realizados en la compañía, se demostró que ambos grupos de productos son

estancos (no tienen fugas) y tienen la misma humedad residual. Sin embargo, al utilizar el equipo de TDLAS-FMS se observó que la concentración de oxígeno en los viales rosáceos era mayor que en los viales blancos.

Se seleccionaron 10 viales con distinta concentración de oxígeno y, como se puede observar en la Figura VI - 1, a medida que la concentración de oxígeno aumenta (viales de la derecha), se observa una mayor coloración en el sólido. Como se puede observar, la diferencia de coloración entre los viales de los extremos es evidente, sin embargo, entre viales de concentración parecida la diferenciación a simple vista es muy subjetiva.

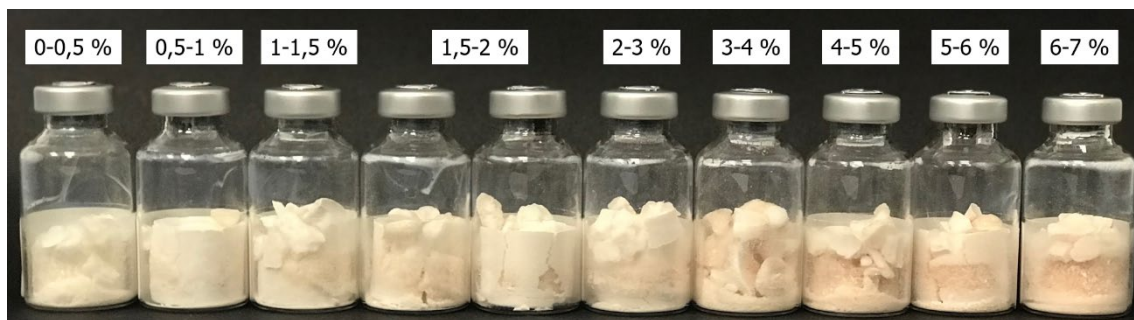


Figura VI - 1. Imagen de las muestras con distinta concentración de oxígeno en el *headspace* (indicada en la parte superior del producto) donde se puede observar como la intensidad del color aumenta con la concentración de oxígeno.

A la vista de estas diferencias, se decidió estudiar la relación que había entre la concentración de oxígeno en el *headspace* y la absorbancia de la solución del producto liofilizado. Antes de analizar la absorbancia de la solución siguiendo el procedimiento descrito en la sección 2.2, se demostró que todos los viales eran estancos comprobando que no había cambios en el contenido de oxígeno justo después de la fabricación y en el momento de la medida.

Como se puede observar en la Figura VI - 2, a las longitudes de onda de la especificación (370 nm y 450 nm), a medida que aumenta la concentración de oxígeno en el *headspace* aumenta la absorbancia. A pesar de este incremento, el lote se considera válido, dado que todos los valores de absorbancia se encuentran por debajo del límite de especificación del producto.

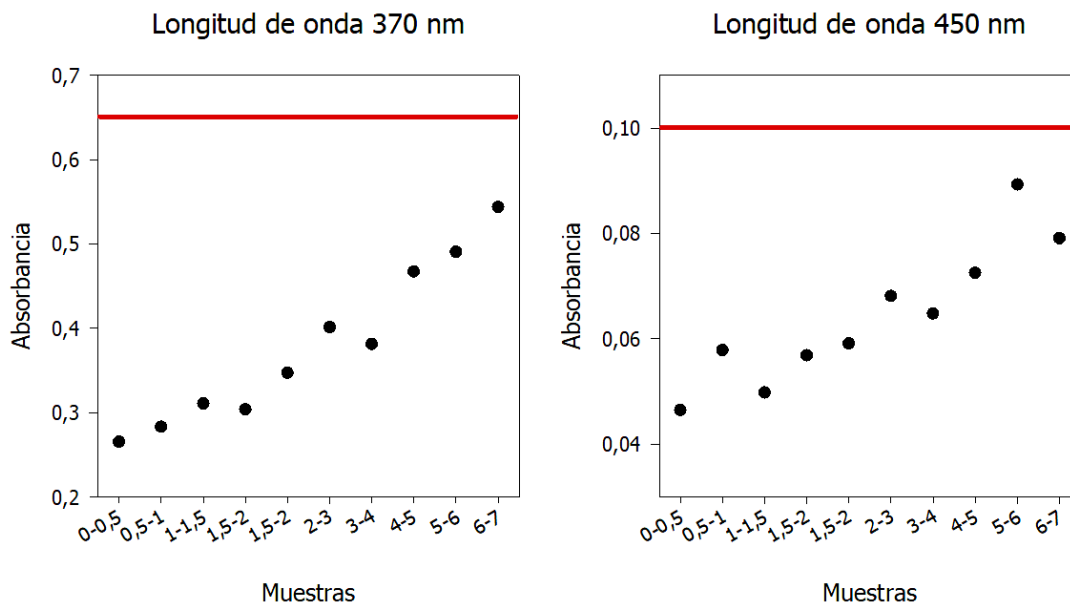


Figura VI - 2. Valores de absorbancia de la disolución de las muestras, a las dos longitudes de onda indicadas en el método.

Este método es capaz de identificar las muestras dependiendo de la coloración del producto, sin embargo, dado su carácter laborioso y destructivo no es viable su uso industrial. Por este motivo es necesario el desarrollo de un método rápido y no destructivo, como la espectroscopia NIR.

3.2. Desarrollo de un modelo cuantitativo

El número de lotes disponibles para desarrollar un modelo multivariable en un proceso industrial puede ser limitado, debido a que la fabricación depende de la demanda del mercado. Asimismo, la variabilidad de las muestras es limitada, debido a que la producción está optimizada para conseguir un intervalo muy estrecho. Para la construcción del modelo, se ha utilizado como conjunto de calibración las muestras de tres de los cuatro lotes de la validación del proceso de fabricación (PPQ) que se describen en la Tabla VI - 1 (lote 2, 3 y 4), dado que contienen un mayor número de muestras con contenido alto de oxígeno. El lote restante, junto con otros de producción, se ha utilizado para el conjunto de validación.

Tabla VI - 1. Distribución de las muestras de los lotes de validación del proceso de fabricación del producto J, indicando el porcentaje de las muestras según su concentración de oxígeno.

Lote	Nº de muestras	Concentración de oxígeno					
		0-1 %	1-2 %	2-3 %	3-4 %	4-5 %	>5 %
1	353	95 %	2 %	1 %	0 %	0 %	1 %

Lote	Nº de muestras	Concentración de oxígeno					
		0-1 %	1-2 %	2-3 %	3-4 %	4-5 %	>5 %
2	290	93 %	2 %	2 %	1 %	1 %	1 %
3	283	71 %	5 %	5 %	2 %	3 %	13 %
4	408	72 %	5 %	3 %	3 %	5 %	12 %

Dada la gran cantidad de muestras en las concentraciones inferiores a 2 % y con el fin de evitar añadir al modelo una sobrecarga de información redundante, se ha reducido el número de muestras en estas concentraciones bajas, eliminándolas de forma aleatoria.

Teniendo en cuenta la relación observada en la Figura VI - 2, se ha realizado un modelo cuantitativo PLSR para intentar hallar la concentración de oxígeno en el *headspace* del vial a partir de medidas de reflectancia del sólido liofilizado, utilizando los datos centrados y la validación cruzada como método de evaluación del modelo, con segmentos de 20 muestras cada uno, tomados de forma aleatoria.

La selección de los pretratamientos e intervalos de longitudes de onda se ha ido realizando de manera iterativa, seleccionando aquellos modelos que presentaban unos mejores resultados.

En primer lugar, utilizando el intervalo de longitudes de onda desde 350 nm a 900 nm se han evaluado diferentes pretratamientos y combinaciones de éstos (SNV, 1º y 2º derivada Savitzky-Golay, las combinaciones de SNV con ambas derivadas y OSC) para realizar una primera selección de aquellos que presentan un modelo más simple con unos mejores resultados. Dado que los espectros presentaban mucho ruido en el extremo inicial del espectro, todos los pretratamientos utilizados se han aplicado después de realizar un suavizado al espectro con el método de Savitzky-Golay, con un polinomio de grado 2 y 11 puntos de ventana móvil. Durante la construcción de los modelos, se han ido eliminando las muestras anómalas detectadas.

El modelo construido con OSC presenta unos mejores resultados en todos los parámetros evaluados por lo que se procedió a optimizar el intervalo de longitudes de onda utilizados usando este pretratamiento.

En este caso, como se puede observar en la Figura VI - 3, la zona inicial del espectro muestra un elevado ruido, mientras que la zona final no muestra ninguna banda de absorción. La selección de los intervalos de longitudes de onda evaluados se ha basado

en la reducción de ruido y en la eliminación de zonas del espectro con ausencia de la señal.

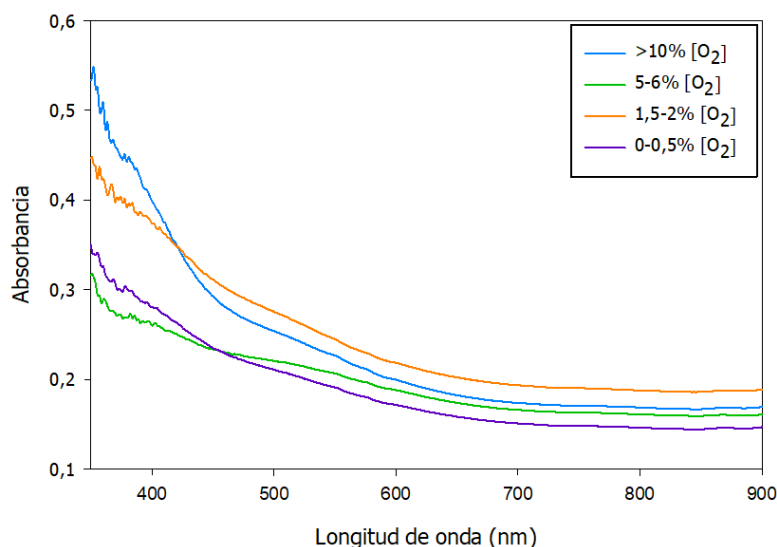


Figura VI - 3. Espectros, sin suavizar, de 4 muestras con distinta concentración de oxígeno en el *headspace*.

En la siguiente tabla, se muestran los resultados obtenidos al evaluar los modelos construidos con el pretratamiento de OSC y utilizando distintos intervalos de longitudes de onda.

Tabla VI - 2. Resultados de la evaluación de los modelos construidos con OSC en diferentes intervalos de longitudes de onda.

Intervalo longitudes de onda	Nº de factores PLSR	RMSEC	RMSECV	Varianza explicada en Y (%)
350 nm a 900 nm	4	0,962	1,066	79,6
400 nm a 900 nm	6	0,755	0,848	86,9
400 nm a 700 nm	3	0,878	0,911	84,4
450 nm a 700 nm	2	0,902	0,918	84,5
450 nm a 650 nm	3	0,893	0,924	84,2

Como se puede observar en la Tabla VI - 2, los modelos presentan valores de RMSECV que mejoran al eliminar la zona del espectro de 350 nm a 400 nm. Los mejores resultados se obtienen en el intervalo de longitudes de onda de 400 nm a 900 nm, sin embargo, es el modelo que presenta una mayor complejidad debido a que requiere 6 factores PLS. Los otros tres modelos presentan resultados parecidos, pero el construido entre 450 nm y 700 nm requiere de un factor menos y explica algo más de varianza en

Y, por lo que se escoge como modelo final. En la Tabla VI - 3 se muestra un resumen del modelo escogido.

Tabla VI - 3. Resumen de los principales parámetros del modelo PLSR. Los valores de R^2 , pendiente y ordenada en el origen corresponden a la recta de los valores calculados por el modelo en *validación cruzada* frente a los valores obtenidos por el método de referencia.

Parámetro	Resultado	Parámetro	Resultado
Nº de muestras	261	Contenido de oxígeno (%)	0 % a 9,5 %
Intervalo de longitudes de onda	450 nm a 700 nm	Pretratamiento	Suavizado de 11 puntos + OSC
Nº de factores OSC	3	Nº de factores PLSR	2
% varianza explicada de Y	84,5	RMSEC	0,902
RMSECV	0,918	R^2	0,917
Pendiente	0,845	Ordenada de origen	0,478

El modelo propuesto presenta los mejores resultados, pero la recta de los valores calculados frente a los valores de referencia presenta una pendiente baja, por lo que no es adecuado para realizar una cuantificación correcta del contenido de oxígeno en el vial, por lo que se decide cambiar el enfoque.

3.3. Desarrollo de un modelo cualitativo

Para poder controlar que el lote no tenga viales con un alto contenido de oxígeno, lo que puede provocar que las muestras estén ligeramente coloreadas, no es necesario conocer el contenido exacto en cada vial, solamente es necesario distinguir las muestras con alto contenido de oxígeno del resto. Teniendo en cuenta este razonamiento, se ha decidido construir un modelo cualitativo que, a partir de un valor de especificación, pueda discriminar las muestras dependiendo de su concentración de oxígeno en el *headspace*.

La elección del valor de la especificación debe hacerse realizando un estudio de estabilidad completo (hasta caducidad) con muestras a distinta concentración de oxígeno. Sin embargo, para el desarrollo de este estudio se ha elegido como especificación provisional, el valor presentado en la patente de la referencia.⁹ En esta patente se comenta que, independientemente de que el producto siga cumpliendo con las especificaciones, para evitar la coloración del producto y prolongar su estabilidad, la concentración de oxígeno debe ser inferior al 5 %; recomendando contenidos inferiores

al 2 % que se puede utilizar como límite de aviso. Este último valor ha sido el elegido como umbral del método de clasificación según el contenido de oxígeno.

La aplicación de este límite concuerda con la observación visual de las muestras sólidas. Como se puede ver en la Figura IV - 1, la coloración en las muestras empieza a ser notable a partir de una concentración de oxígeno del 4 %, mientras que, por debajo de 2-3 % la diferencia de coloración entre las muestras es inapreciable.

Teniendo en cuenta el límite definido, se procede a desarrollar un modelo cualitativo para realizar un *screening* de las muestras, que permita clasificarlas, según si su concentración de oxígeno se encuentra por encima o por debajo de este límite. En la siguiente figura se muestra el % de muestras que se dispone para cada clase en el conjunto de entrenamiento.

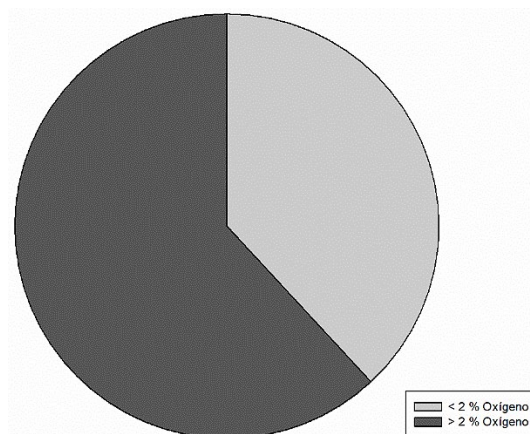


Figura VI - 4. Porcentaje de muestras de cada una de las clases en el conjunto de calibración, con 38,1 % de muestras con una concentración de oxígeno inferior al 2 %, y un 61,8 % de muestras con una concentración de oxígeno superior al 2 %.

En el caso de un método cualitativo, los parámetros que definen la calidad del método y proporcionan información sobre su desempeño están basados en respuestas de tipo binario (cumplir/no cumplir) sobre la especificación. Los resultados obtenidos se reducen a cuatro posibles situaciones, ver Figura VI - 5.

- Positivos verdaderos. Los positivos verdaderos (TP, siglas del inglés *True Positives*) son aquellas muestras que son clasificadas como positivas (en este caso por encima del límite de 2 % de oxígeno) y realmente son positivas.
- Negativos verdaderos. Los negativos verdaderos (TN, siglas del inglés *True Negative*) son las muestras clasificadas como negativas (por debajo del límite de concentración de oxígeno del 2 %) y que realmente son negativas.
- Falsos positivos. Los falsos positivos (FP, siglas del inglés *False Positive*) son aquellas muestras que realmente son negativas (la concentración de oxígeno es

inferior al límite), pero que el modelo las clasifica como positivas (muestras con concentración de oxígeno superior al límite).

- **Falsos negativos.** Los falsos negativos (FN, siglas del inglés *False Negative*) son aquellas muestras realmente negativas (la concentración de oxígeno es superior al límite), pero que el modelo las clasifica negativas (muestras con concentración de oxígeno inferior al límite).

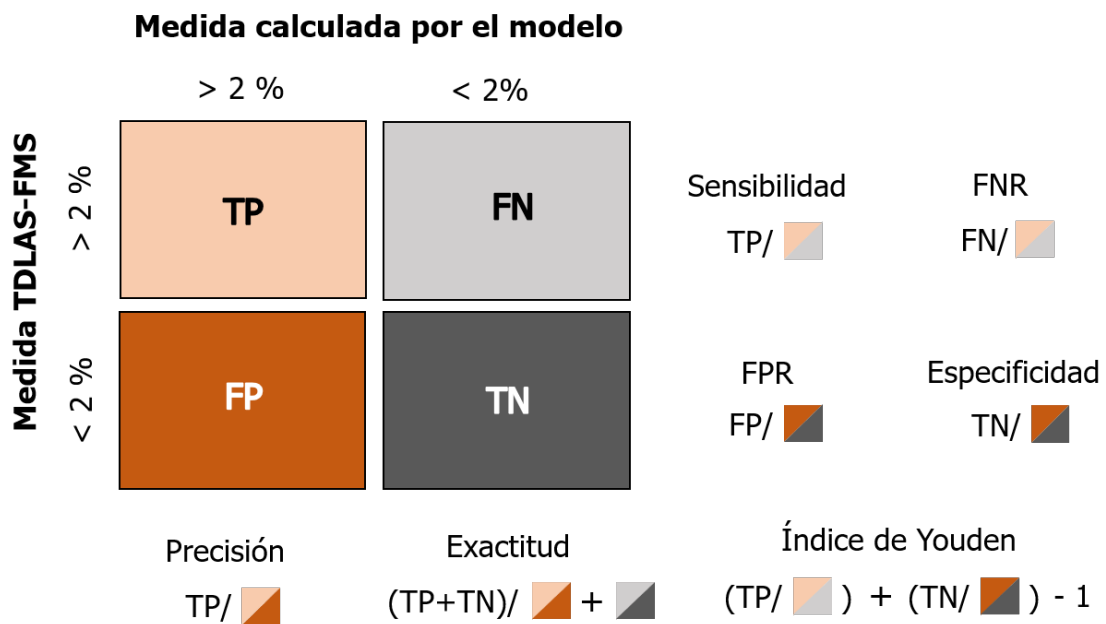


Figura VI - 5. Tabla de contingencia donde se indica cómo se clasifican las muestras respecto al límite de concentración de oxígeno y los diferentes parámetros calculados.

A partir de estos valores se pueden calcular las propiedades analíticas de un método cualitativo, las cuales servirían para evaluar los modelos. Estas propiedades de los métodos cualitativos pueden tener el mismo nombre que las de los cuantitativos, pero su significado es distinto:¹⁰⁻¹²

- **Exactitud.** En este caso la exactitud se define como la relación entre el valor total de muestras clasificadas correctamente, ya sean positivas o negativas, y el número total de muestras. Este parámetro se calcula con la Ecuación VI - 1.

$$Exactitud = \frac{TN + TP}{TN + TP + FN + FP} \quad \text{Ecuación VI - 1}$$

- **Precisión.** En los modelos cualitativos este parámetro se define como la relación entre los resultados verdaderamente positivos y los resultados positivos designados por el modelo. Se calcula con la Ecuación VI - 2.

$$\text{Precisión} = \frac{TP}{TP + FP} \quad \text{Ecuación VI - 2}$$

- Sensibilidad. En un método cualitativo esta propiedad se refiere a la habilidad del modelo para reconocer que una muestra es realmente positiva (y no clasificarla como negativa, generando falsos negativos). La sensibilidad sería la tasa de positivos verdaderos y se expresa como una relación o porcentaje y se calcula de acuerdo con la Ecuación VI - 3.

$$\text{Sensibilidad} = \frac{TP}{TP + FN} \quad \text{Ecuación VI - 3}$$

- Especificidad. En un método cualitativo este concepto se refiere a la habilidad del modelo para reconocer que una muestra es realmente negativa (y no como positiva, generando falsos positivos). La especificidad sería la tasa de negativos verdaderos. Al igual que la sensibilidad, se expresa como una relación o porcentaje y se calcula de acuerdo con la Ecuación VI - 4.

$$\text{Especificidad} = \frac{TN}{TN + FP} \quad \text{Ecuación VI - 4}$$

A partir de los valores de sensibilidad y especificidad se pueden hallar distintas propiedades:

- Tasa de falsos positivos. La tasa de falsos positivos (FPR, siglas del inglés *False positive rate*) es la proporción de casos que se detectan erróneamente como positivos respecto a todos los casos que realmente son negativos. Es el valor complementario de la especificidad ($1 - \text{Especificidad}$), y se calcula directamente de acuerdo con la Ecuación VI - 5.

$$FPR = \frac{FP}{TN + FP} \quad \text{Ecuación VI - 5}$$

- Tasa de falsos negativos. La tasa de falsos negativos (FNR, siglas del inglés *False negative rate*) es la relación entre los casos que se detectan erróneamente como negativos respecto al total de resultados positivos obtenidos en la prueba. Es el valor complementario de la sensibilidad ($1 - \text{Sensibilidad}$), y se calcula directamente de acuerdo con la Ecuación VI - 6.

$$FNR = \frac{FN}{TP + FN} \quad \text{Ecuación VI - 6}$$

- Índice de Younde (J). El índice de Youden toma valores comprendidos entre 0 y 1 y se calcula siguiendo la Ecuación VI - 7. Este parámetro resume el rendimiento del método analítico, por lo que un valor de 1 indica que la prueba es perfecta y 0 que carece de importancia analítica.

$$\text{Índice de Youden (J)} = \text{Sensibilidad} + \text{Especificidad} - 1 \quad \text{Ecuación VI - 7}$$

Para clasificar las muestras por encima o por debajo de la especificación dependiendo de la concentración de oxígeno en el *headspace* los modelos cualitativos se han construido utilizando un PLS-DA. Los valores de la Y de las muestras se han codificado como 1 si el contenido de oxígeno era mayor que el 2 %, y como 0 si eran menor. Aquí también se ha centrado los datos y se ha realizado una validación cruzada como método de evaluación del modelo, con segmentos de 20 muestras cada uno, tomados de forma aleatoria.

Utilizando el rango de longitudes de onda del modelo PLSR seleccionado en la sección anterior (desde 450 nm a 700 nm), se han evaluado los mismos pretratamientos que se evaluaron durante el desarrollo del modelo cuantitativo.

Tabla VI - 4. Comparativa de los modelos PLS-DA construidos con distintos pretratamientos.

Pretratamiento	Nº de factores PLS-DA	Especificidad	Sensibilidad
SNV	4	0,87	0,91
1Der	5	0,90	0,91
2Der	5	0,76	0,82
SNV + 1Der	3	0,86	0,91
SNV + 2Der	5	0,68	0,82
OSC	2	0,98	0,93

Como se puede observar en la Tabla VI - 4, los modelos construidos con segunda derivada presentan una menor especificidad y sensibilidad que el resto de los modelos, por lo que se descarta este pretratamiento. Respecto a los demás modelos, todos muestran una especificidad y sensibilidad próxima al 90 %. Sin embargo, destacan los valores obtenidos con el modelo construido con OSC los cuales son próximos al 100 % en el caso de la especificidad y presentan un valor de sensibilidad superior a los demás. Estos resultados se obtienen sin utilizar un elevado número de factores PLSR, por lo que se escoge este pretratamiento para construir el modelo final.

Tabla VI - 5. Resumen de las características del modelo PLS-DA utilizado.

Parámetro	Resultado	Parámetro	Resultado
Nº de muestras	244	Contenido de oxígeno (%)	0 % a 9,5 %
Intervalo de longitudes de onda	450 nm a 700 nm	Pretratamiento	Suavizado de 11 puntos + OSC
Nº de factores OSC	3	Nº de factores PLS-DA	2

En la Figura VI - 6 se puede observar los *scores* del modelo donde se observa la separación entre ambos grupos de muestras.

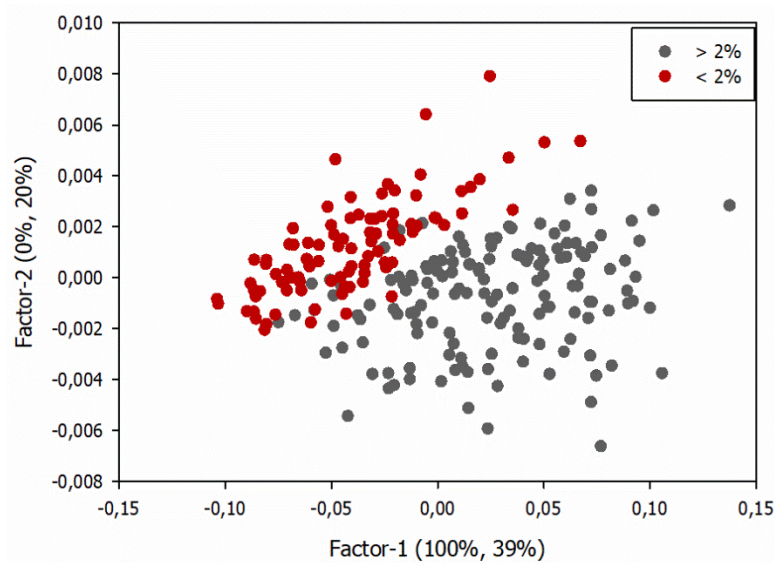


Figura VI - 6. Representación de los *scores* del modelo PLS-DA utilizado.

En la Tabla VI - 6 se muestra la tabla de contingencia del modelo seleccionado más los demás parámetros evaluados.

Tabla VI - 6. Parámetros de la evaluación del modelo cualitativo.

Positivos verdaderos	140	Tasa de falsos positivos	0,02	Tasa de falsos negativos	0,07
Negativos verdaderos	91	Sensibilidad	0,93	Especificidad	0,98
Falsos positivos	2	Precisión	0,99	Índice de Youden	0,91
Falsos negativos	11	Exactitud	0,95	Muestras totales	244

Además, de los parámetros evaluados, es necesario conocer el límite de detección del método teniendo en cuenta, a la vez, tanto los falsos positivos (se concluye que la

concentración de oxígeno es mayor al umbral, cuando en realidad no lo es) y los falsos negativos (se concluye que la concentración de oxígeno está por debajo del umbral, aunque no lo esté). En la referencia adjuntada, se sugiere el uso de dos magnitudes diferentes, el límite de decisión, $CC\alpha$, y la capacidad de detección, $CC\beta$, marcando cada una de ellas un límite de concentración relacionado con cada tipo de error.¹³ El error α o error de tipo I, definido como la "probabilidad de que la muestra analizada sea realmente conforme, aunque se haya obtenido una medición no conforme («decisión de falso no conforme»)". Y el error abreviado como β o error de tipo II, definido como la "probabilidad de que la muestra analizada sea realmente no conforme, aunque se haya obtenido una medición conforme («decisión de falso conforme»)".¹⁴

- Límite de decisión, $CC\alpha$, se define como el "límite en el cual y a partir del cual se puede concluir con una probabilidad de error α que una muestra no es conforme".¹⁴ Habitualmente, α se define a un nivel de significación de 0,05. Es decir, es la concentración límite por debajo de la cual, la probabilidad de obtener falsos positivos es menor que el 5 %.
- La capacidad de detección, $CC\beta$, se define como el "contenido mínimo de la sustancia que puede ser detectado, identificado o cuantificado en una muestra, con una certeza estadística de $1 - \beta$ ".¹⁴ Habitualmente, β se define a un nivel de significación de 0,95. Representa la concentración límite por encima de la cual la probabilidad de obtener falsos negativos es inferior al 5 %.

El intervalo de concentraciones entre ambos límites se define como la zona de falta de fiabilidad (*unreliability region*) del modelo cualitativo. Los valores de $CC\alpha$ y $CC\beta$, así como la zona de falta de fiabilidad se pueden calcular a partir de la curva característica de rendimiento (PCC, siglas del inglés *Performance Characteristic Curve*).¹⁰ En la PCC se representa la probabilidad de obtener resultados positivos frente a la concentración de oxígeno, $P(O_2)$. Los datos experimentales se ajustan a una función sigmoide mediante un ajuste no lineal y se trazan dos líneas paralelas al eje de abscisas desde la probabilidad $P(O_2) = 95 \%$ y $P(O_2) = 5 \%$. Los puntos de corte con la sigmoide ajustada definirán $CC\beta$ y $CC\alpha$, respectivamente.

Las concentraciones de oxígeno experimentales se han agrupado en intervalos, cada 0,5 % desde 0 a 4 %, y cada 1 % para valores superiores. En cada intervalo se ha calculado la probabilidad de obtener una muestra positiva y los valores se han representado en la Figura VI - 7. Los valores experimentales se han ajustado a la

función: $P(O_2) = 1 - e^{-aO_2^b}$ mediante una regresión no lineal por mínimos cuadrados, utilizando el algoritmo de Gauss-Newton.

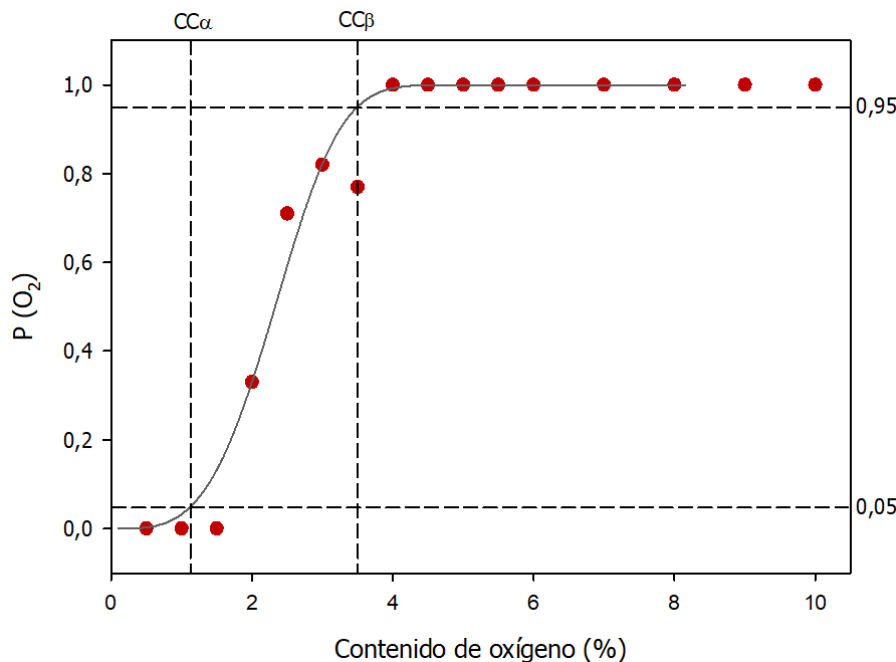


Figura VI - 7. Curva característica de rendimiento del modelo cualitativo. Los puntos son los valores experimentales y la línea la función ajustada. Se indican los valores de $CC\alpha$ y $CC\beta$ obtenidos.

A partir del valor de corte para las probabilidades de 0,05 y 0,95, se ha calculado el valor de $CC\alpha$, que es de 1,13 %, y el valor de $CC\beta$, que es de 3,5 %. Para concentraciones de oxígeno en el vial entre 0 y $CC\alpha$, los resultados del método indican no contaminación de oxígeno. Para concentraciones $\geq 3,5$ %, el método indica contaminación de oxígeno. La zona de falta de fiabilidad, entre los dos límites, es aceptable, debido a que el principal interés del método es que las muestras con un contenido de oxígeno ≥ 5 % siempre se clasifiquen como positivas respecto a la interferencia.

El umbral marcado de un 2 % de oxígeno en el *headspace* de las muestras, con los resultados obtenidos, parece ser demasiado restrictivo, debido a que los falsos negativos se han encontrado con muestras con un contenido de oxígeno inferior a 3,2 %. Esto queda confirmado con la zona de falta de fiabilidad, la cual se encuentra comprendida en el intervalo de concentraciones de 1,13 % a 3,5 %. Sería posible la elección de un umbral mayor, pero su valor seguiría siendo subjetivo sin un estudio de estabilidad de las muestras hasta su caducidad. Con el umbral actual se asegura que ninguna muestra con un contenido superior al 5 %, con una alta probabilidad de generar color con el tiempo, se identifique erróneamente como una muestra negativa. Por lo que cualquier

muestra que se clasifique incorrectamente como negativa, tendrá una coloración rosa aceptable.

Teniendo en cuenta estos resultados, se decide validar el modelo PLS-DA construido con OSC después de realizar un suavizado espectral de 11 puntos de ventana móvil en el intervalo de longitudes de onda desde 450 nm hasta 700 nm.

3.4. Validación del modelo

Para validar la capacidad de clasificación del modelo se estudiaron los dos parámetros que se recomiendan en las guías de las agencias reguladoras la especificidad y la robustez.¹⁵⁻¹⁷

Con el fin de evaluar ambos, se han utilizado muestras independientes de las utilizadas en el conjunto de calibración para obtener el conjunto de validación. Este conjunto lo constituyen 4 lotes, uno de ellos el restante del PPQ y los otros tres adquiridos en rutina. En la siguiente tabla se muestra un resumen de la información de estos lotes.

Tabla VI - 7. Distribución de las muestras de los lotes del conjunto de validación según su concentración de oxígeno.

Lote	Nº de muestras	Concentración de oxígeno					
		0-1 %	1-2 %	2-3 %	3-4 %	4-5 %	>5 %
1 ^{er} PPQ	61	45	7	5	1	1	2
1 rutina	20	19	-	-	-	-	1
2 rutina	20	16	2	1	-	-	1
3 rutina	20	19	-	-	1	-	-

Como se puede observar, hay muy pocas muestras por encima del umbral definido en el método, lo que es un indicativo de la calidad del proceso de fabricación.

Antes de realizar la clasificación de las muestras, se proyectan en los *scores* del modelo para ver si pertenecen a la misma población.

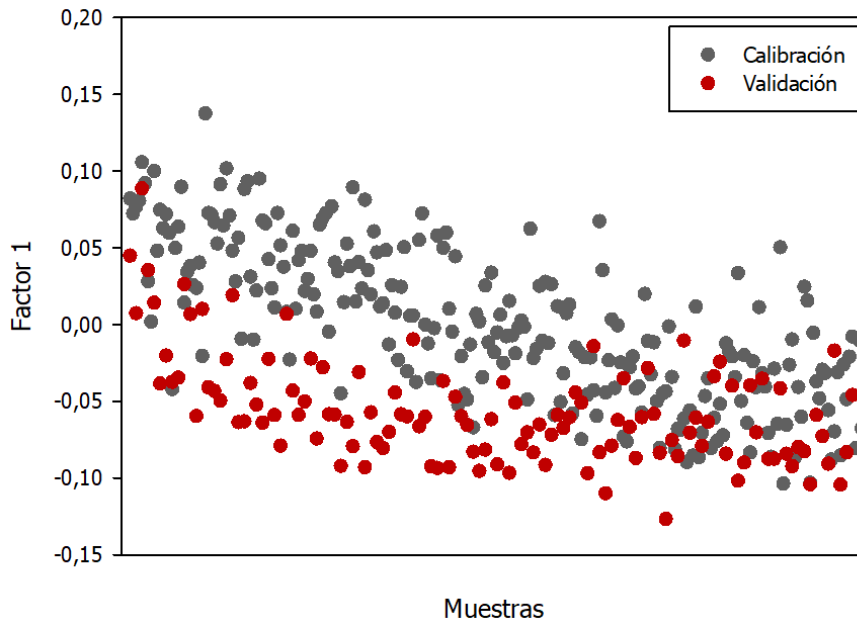


Figura VI - 8. Representación de la proyección de las muestras del conjunto de validación en los *scores* del factor 1 del modelo.

Como se puede observar en la Figura VI - 8, las muestras de los lotes quedan proyectadas dentro del grupo de muestras del conjunto de calibración. Las muestras de concentración alta corresponden a un score positivo, y las de concentración baja a un score negativo, tanto en las muestras de calibración (en gris) como en las muestras de validación (en rojo).

Para la validación, se han calculado los mismos parámetros que en la evaluación del modelo. Los resultados se muestran en la siguiente tabla:

Tabla VI - 8. Resultados de la validación del modelo construido.

Positivos verdaderos	11	Tasa de falsos positivos	0,02	Tasa de falsos negativos	0,07
Negativos verdaderos	106	Sensibilidad	0,85	Especificidad	0,98
Falsos positivos	2	Precisión	0,85	Índice de Youden	0,83
Falsos negativos	2	Exactitud	0,97	Muestras totales	121

Como se puede observar en la tabla, la tasa de falsos positivos y negativos es inferior al 10 %, y la sensibilidad y especificidad son próximas a 1. Las muestras clasificadas como falsos positivos y negativos tienen una concentración próxima al umbral del 2 % y confirman la zona de fiabilidad del modelo. Teniendo en cuenta estos resultados se

puede decir que el modelo evaluado es específico y robusto para el propósito por el cual se ha diseñado.

4. CONCLUSIONES

En este capítulo se ha desarrollado y validado de forma satisfactoria un modelo cualitativo de espectros visible-NIR para clasificar las muestras dependiendo de la concentración de oxígeno que presentan en su *headspace*.

El modelo construido consiste en un PLS-DA de dos factores construido con 244 muestras y evaluado con validación cruzada. El pretratamiento que mejores resultados se ha construido con un pretratamiento de suavizado Savitzky-Golay seguido de un OSC en el intervalo de longitudes de onda de 450 nm a 700 nm. Los resultados de la evaluación del modelo muestran una buena especificidad (98 %) y sensibilidad (93 %), además de un índice de Youden del 0,91 próximo a 1. La zona de falta de fiabilidad se encuentra comprendida en el intervalo de concentraciones de 1,13 % a 3,5 %. Aunque el límite superior de este intervalo se encuentre por encima del umbral escogido del 2 %, está por debajo del límite superior del 5 % por lo que cualquier muestra que se clasifique incorrectamente como negativa, tendría una coloración rosa aceptable. Queda abierta la posibilidad de modificar el umbral a partir del estudio de estabilidad completo (hasta caducidad) con muestras con distinta concentración de oxígeno en el *headspace*.

El modelo se ha validado con 4 lotes independientes al conjunto de calibración, uno de PPQ y los tres restantes de rutina. La robustez del modelo ha sido demostrada con la clasificación de las muestras de 3 lotes en rutina. Con las muestras del conjunto de validación se ha demostrado que el modelo cualitativo presenta una sensibilidad y precisión del 85 %, una especificidad del 98 % y una exactitud del 97 %. Los resultados de la validación son conformes para los dos parámetros evaluados (especificidad y robustez).

Con todo esto se puede concluir que ha sido posible desarrollar un método utilizando el mismo espectro empleado para la cuantificación de la humedad residual del producto liofilizado J, que permita saber qué productos pueden desarrollar una mayor coloración con el tiempo.

5. BIBLIOGRAFÍA

- (1) Baertschi, S. W.; Alsante, K. M.; Reed, R. A. *Pharmaceutical Stress Testing*, 2^o.; Baertschi, S. W., Alsante, K. M., Reed, R. A., Eds.; CRC Press, 2005.
- (2) Lighthouse Instruments. *Application Note: In-Process Monitoring of Headspace Oxygen Levels in Parenteral Containers*.
- (3) Nema, S.; Ludwig, J. D. *Pharmaceutical Dosage Forms - Parenteral Medications*, 3rd ed.; Nema, S., Ludwig, J. D., Eds.; CRC Press, 2016; Vol. 1.
- (4) Franks, F. *Freeze-Drying of Pharmaceuticals and Biopharmaceuticals*; Royal Society of Chemistry: Cambridge, 2007.
- (5) Brown, H.; Mahler, H.-C.; Mellman, J.; Nieto, A.; Wagner, D.; Schaar, M.; Mathaes, R.; Kossinna, J.; Schmitting, F.; Dreher, S.; et al. Container Closure Integrity Testing—Practical Aspects and Approaches in the Pharmaceutical Industry. *PDA J. Pharm. Sci. Technol.* 2017, 71 (2), 147–162.
- (6) de Castro, M. D. L.; García, J. L. L. *Acceleration and Automation of Solid Sample Treatment*, 1^o.; Elsevier Science (USA), 2002; Vol. 24.
- (7) Hede, J. O.; Fosbøl, P. L.; Berg, S. W.; Dahl, S. Container Closure Integrity Testing—Method Development for Freeze-Dried Products Using Laser-Based Headspace Oxygen Analysis. *PDA J. Pharm. Sci. Technol.* 2019, 73 (2), 170–180.
- (8) Clua Palau, G. Application of Quality by Design and near Infrared Spectroscopy in Manufacturing and Control of Freeze-Dried Drug Products, Universitat Autònoma de Barcelona, 2018.
- (9) Ardia, R.; Magliano, B. EP Patent 2520282A2, 2012.
- (10) López, M. I.; Callao, M. P.; Ruisánchez, I. A Tutorial on the Validation of Qualitative Methods: From the Univariate to the Multivariate Approach. *Anal. Chim. Acta* 2015, 891, 62–72.
- (11) Kotu, V.; Deshpande, B. Model Evaluation. In *Data Science*; Elsevier, 2019; pp 263–279.
- (12) Nisbet, R.; Miner, G.; Yale, K. Model Evaluation and Enhancement. In *Handbook of Statistical Analysis and Data Mining Applications*; Elsevier, 2018; pp 215–233.
- (13) Currie, L. A. Limits for Qualitative Detection and Quantitative Determination. Application to Radiochemistry. *Anal. Chem.* 1968, 40 (3), 586–593.
- (14) La comisión de las comunidades Europeas. *2002D0657-ES-10.01.204*; 2003; pp 1–36.
- (15) The Internacional Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. *ICH Q2 (R1) Validation of Analytical Procedures: Text and Methodology*; 2011; Vol. 20.
- (16) European Medicines Agency. Guideline on the Use of near Infrared Spectroscopy by the Pharmaceutical Industry and the Data Requirements for New Submissions and Variations. 2014, p 28.
- (17) European Pharmacopoeia. 5.21. Chemometric Methods. 2016, 5641–5658.

VII. CONSTRUCCIÓN Y VALIDACIÓN
DE UNA BIBLIOTECA DE ESPECTROS
NIR

1. INTRODUCCIÓN

Como se ha comentado en la sección 3.4.2 del capítulo I, la identificación de los materiales, tanto de recepción como de liberación, es uno de los análisis de rutina realizados en los laboratorios de control de calidad de cualquier empresa farmacéutica. Para poder realizar una buena identificación del producto hace falta la compilación de distintos resultados procedentes de diversas técnicas analíticas, lo que se traduce en un trabajo tedioso que requiere un gasto de tiempo y recursos considerable. Una de las alternativas es utilizar alguna técnica rápida para realizar este tipo de análisis de identificación. La espectroscopia NIR es adecuada para ello, ya que es rápida y no destructiva; sin embargo, teniendo en cuenta que las bandas de los espectros NIR no están bien resueltas, sino que presentan un elevado solapamiento, se requiere el uso de la quimiometría.^{1,2}

Los espectros de los distintos productos que se deben identificar se recogen en lo que se denomina una biblioteca. A ella se le aplican técnicas de reconocimiento de pautas (PRM) para estudiar la manera en que se pueden distinguir los productos entre sí. Si el espectro de una nueva muestra coincide con alguno de los productos que forman parte de la biblioteca, se dará como identificado. Por el contrario, si no coincide con ninguno, querrá decir que no pertenece a ningún producto de la biblioteca. Esta biblioteca general, o de nivel 1, puede que contenga productos muy similares que se confundan entre sí, con lo que sería imposible identificar sin ambigüedad a cuál de los productos confundidos pertenece una nueva muestra.

En estos casos para poder resolver ambigüedades la estrategia consiste en crear una nueva biblioteca con los espectros de los productos confundidos, que se denominará sub-biblioteca, o biblioteca de nivel 2. En esta sub-biblioteca, se necesitará aplicar una combinación de intervalo de longitudes de onda, pretratamiento y/o PRM distintos al de la biblioteca general, con el fin de conseguir una mayor selectividad para la discriminación entre productos. Si aún quedaran productos confundidos, se continuaría el proceso creando una nueva sub-biblioteca con los espectros de los productos no separados, y se repetiría el proceso hasta conseguir una identificación unívoca de los productos. Cuando en la identificación se utiliza más de un nivel, distintas sub-bibliotecas, se dice que se trabaja con bibliotecas de identificación en cascada. Una muestra se compara con las bibliotecas desde el nivel superior hasta que se consigue su identificación inequívoca. Si no es posible, la muestra a identificar se cataloga como

diferente de todas las otras existentes. En la construcción de estas bibliotecas, es crítico que los espectros contengan la variabilidad habitual esperada en los productos; para a partir de ella se pueden definir los criterios de aceptación que han de cumplir las nuevas muestras para decidir que se identifican como un producto determinado.

En la industria farmacéutica, el uso de este tipo de procedimiento es común para la identificación de los materiales de partida, en cambio en el producto terminado no está tan ampliamente implementado. Las principales diferencias se basan en la regulación en ambos y en que el producto terminado debe identificarse en el mismo envase.^{3,4}

El objetivo de este apartado consiste en realizar una primera prueba piloto para construir una biblioteca de espectros con la finalidad de poder identificar los productos liofilizados terminados antes de su liberación y en el mismo envase. Con los resultados de este estudio se decidirá si es factible su implementación en los Laboratorios Reig Jofre.

2. MATERIALES Y MÉTODOS EXPERIMENTALES

2.1. Muestras

En el desarrollo de este apartado, se han utilizado muestras de producción de 13 productos terminados los cuales tienen distintos API. En la Tabla VII - 1 se puede observar una pequeña descripción de los productos analizados, indicando el nombre que se utilizará durante este estudio para mantener la confidencialidad de la empresa.

Tabla VII - 1. Codificación de los productos estudiados con la descripción de su principal función, los excipientes que contienen y la función de éstos.⁵

Producto	Descripción	Excipientes	Finalidad del excipiente
A	Antiviral	α	Ajuste del pH.
B	Antibiótico	α y β	Ajuste de pH y agente regulador de pH.
C	Antídoto	-	-
D	Antiviral	-	-
E	Tinte para contrastes	-	-
F	Antifúngico	α , β y γ	Ajuste de pH, agente regulador de pH y azúcar que actúa como agente de carga (aumentar el contenido en sólidos del producto).
G	Antiepiléptico	-	-

Producto	Descripción	Excipientes	Finalidad del excipiente
H	Analgésico	δ y ε	Aminoácido que actúa como agente de carga y ajuste de pH.
I	Antibiótico	-	-
J	Antibiótico	ζ	Agente tonificante (propiedades isotónicas).
K	Cardiovascular	η y α	Azúcar que actúa como agente de carga y ajuste de pH.
L	Antiarrítmico	δ y α	Aminoácido que actúa como agente de carga y ajuste de pH.
M	Anticoagulante	η y α	Azúcar que actúa como agente de carga y ajuste de pH.

Las muestras utilizadas en este estudio se han ido recopilando durante finales del 2017 hasta finales del 2019, una vez el lote había sido identificado como conforme por el departamento de control de calidad.

2.2. Adquisición de los espectros

Los espectros de las muestras se han registrado empleando el espectrofotómetro NIR Labspec500 descrito anteriormente.

El procedimiento para adquirir los espectros, como se ha podido ver en capítulos anteriores, puede verse afectado por distintos factores como el tiempo de exposición al haz de luz NIR de la muestra, posición de la muestra, etc. En este caso, teniendo en cuenta que el objetivo de este modelo es cualitativo, sólo se han considerado aquellos factores relacionados con las etapas del registro del espectro.

Tal y como se describe en la Tabla V – 7 y Tabla V - 8, solamente los parámetros de duración del análisis y posición de la muestra tienen un riesgo inicial por encima del umbral aceptado de RPN, 20. Sin embargo, al modificar el objetivo del modelo, se deben modificar las acciones de mitigación y recalcular el riesgo residual. Estas acciones se pueden ver en la Tabla VII - 2.

Tabla VII - 2. Análisis de riesgos asociados al registro del espectro, para la construcción de la biblioteca de espectros NIR.

Registro espectro											
Parámetros	Modo de fallo potencial	Efecto del potencial fallo	S	P	D	Riesgo inicial	Acciones de mitigación	S	P	D	Riesgo residual
Duración del análisis	La exposición prolongada de la muestra sobre el haz de luz calienta la muestra.	La superficie de la muestra escaneada podría variar observando diferencias en el espectro.	10	7	1	70	Teniendo en cuenta que este parámetro tiene un efecto relevante en alguno de los productos estudiados, como acción de mitigación se realiza el procedimiento del producto más restrictivo. ⁶ En este caso, después de cada registro se retirará la muestra de la fuente de luz NIR.	10	1	1	10
Posición de la muestra	La colocación de las muestras sobre el accesorio no es repetitiva y no se cubre la máxima superficie del haz de luz.	Los niveles de error del modelo aumentarían debido a las variaciones de la línea base y/o aumentaría el ruido del espectro.	10	4	4	160	Como se ha demostrado en apartados anteriores, pequeñas variaciones en la posición de la muestra durante el registro (siempre y cuando ésta se encuentre dentro de la ventana de la luz) no presentan ningún efecto apreciable en el espectro.	10	1	1	10

Teniendo en cuenta las mitigaciones comentadas, el procedimiento para el registro de los espectros de las muestras de la biblioteca consiste en colocar la parte inferior del vial encima de la fuente de luz del accesorio ASD Muglight. Después de cada registro, el vial de la muestra será retirado de la fuente de luz y se registrará el espectro tres veces realizando una rotación de 120° cada vez.

Para construir y validar la biblioteca de espectros NIR se ha utilizado el software OPUS. Dado que este programa solamente puede importar espectros con su mismo formato, una vez registrados los espectros, se ha utilizado una macro para realizar la conversión del formato de los espectros ASD a un formato aceptado en OPUS.

3. RESULTADOS Y DISCUSIÓN

3.1. Descripción de las muestras

Previamente al inicio de la construcción de la biblioteca, se ha realizado un primer contacto con los productos a clasificar, con el fin de saber con cuáles es posible que haya problemas de confusión o mayor dificultad para diferenciarlos.

La Figura VII - 1 muestra los espectros promedio de los 13 productos estudiados. A la vez, para poder cuantificar de forma sencilla cuánto se parecen los productos entre sí, se ha calculado la matriz de correlación (Tabla VII - 3).

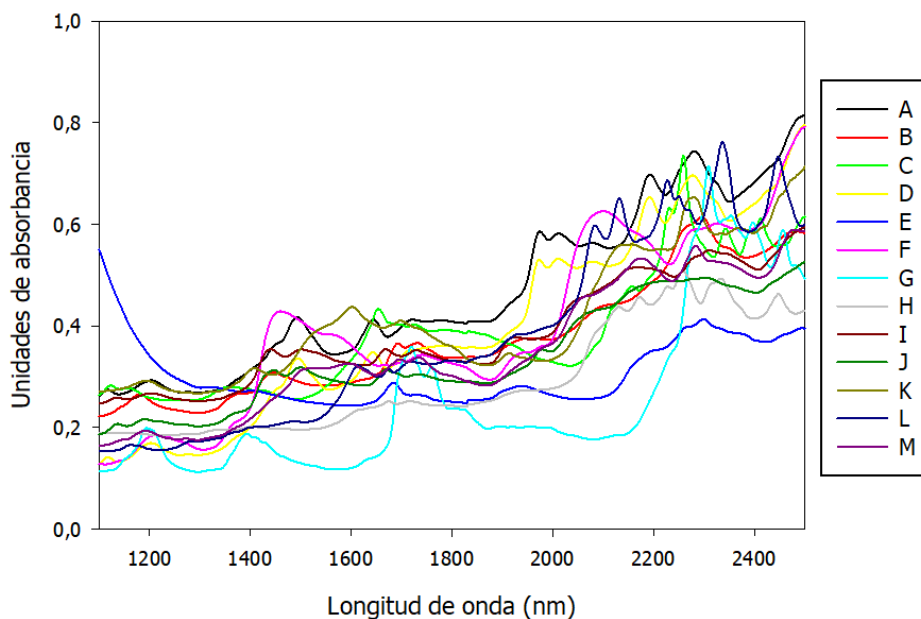


Figura VII - 1. Espectros NIR promedio de cada producto en el intervalo de longitudes de onda de 1100 nm a 2500 nm.

Como se puede observar en la Figura VII - 1, los espectros NIR de los productos E y G presentan un perfil más distinto al de los demás, especialmente el del producto E. Esto aparece reflejado en la matriz de correlación (Tabla VII - 3), donde se puede ver que E es el producto que presenta menor coeficiente de correlación con los demás productos. El producto G, también muestra la misma tendencia y es el segundo producto con menor coeficiente de correlación con el resto. Esto indica que, *a priori*, serán los productos más fáciles de diferenciar del resto.

En la matriz de correlación (Tabla VII - 3) también se observa que los productos pertenecientes a la misma familia terapéutica (ver Tabla VII - 1) presentan un mayor coeficiente de correlación entre sí por lo que *a priori* serán más difíciles de separar y puede hacer falta el uso de sub-bibliotecas. Por ejemplo, los antiviricos A y D tienen entre sí una correlación de 0,99, mayor que con el resto de los productos. Esto es debido a que ambos presentan la misma estructura molecular básica, cambiando únicamente algún grupo funcional. Lo mismo sucede con los antibióticos I y J, también con un coeficiente de correlación entre ellos del 0,99. En este caso, además de tener una estructura muy similar ambos productos se producen por fermentación bacteriana. El producto M es un anticoagulante y también tiene un coeficiente de correlación elevado con los antibióticos I y J. Esto se podría explicar porque los tres productos son macromoléculas orgánicas, las cuales presentan en su estructura azúcares y péptidos similares.

Tabla VII - 3. Matriz de correlación entre los espectros promedio de cada producto.

	A	B	C	D	E	F	G	H	I	J	K	L	M
A	1,00	0,97	0,88	0,99	0,38	0,92	0,78	0,94	0,97	0,97	0,90	0,95	0,97
B		1,00	0,93	0,96	0,45	0,90	0,87	0,96	0,97	0,96	0,93	0,96	0,95
C			1,00	0,87	0,48	0,77	0,84	0,92	0,86	0,87	0,91	0,90	0,86
D				1,00	0,31	0,93	0,76	0,93	0,97	0,97	0,90	0,95	0,97
E					1,00	0,21	0,56	0,47	0,37	0,33	0,43	0,36	0,30
F						1,00	0,67	0,87	0,97	0,96	0,91	0,90	0,97
G							1,00	0,77	0,77	0,74	0,79	0,76	0,73
H								1,00	0,95	0,96	0,93	0,98	0,94
I									1,00	0,99	0,94	0,96	0,99
J										1,00	0,93	0,96	0,99
K											1,00	0,91	0,95
L												1,00	0,96
M													1,00

Como ya se ha descrito en la Tabla VII - 1, los productos no son moléculas puras, sino una mezcla de API y distintos excipientes. Los espectros NIR de los excipientes pueden tener también una importancia en la similitud de los espectros de los productos. Si se considera los productos con los mismos excipientes, independientemente de la familia a la que pertenezcan los API, se observa que los productos que contienen los mismos excipientes de ajuste de pH o tónicos no se parecen entre sí. Esto se debe a que se añaden en cantidades muy pequeñas en comparación con el API, por lo que su señal tiene poco impacto en el espectro NIR del producto. Sin embargo, por lo que respecta a los excipientes con propiedades de carga (que se añaden en mayor cantidad), se observa que los productos H y L, los cuales contienen el excipiente δ , tienen un coeficiente de correlación de 0,98 aunque no pertenezcan a la misma familia y el API tenga distinta estructura.

3.2. Construcción de la biblioteca de espectros

Para poder construir una biblioteca representativa de los espectros de producto terminado, se han registrado espectros de lotes aleatorios de producción, analizando un mínimo de 4 muestras por lote.

Como se puede observar en la Tabla VII - 4, el número de lotes no es el mismo para cada producto, debido a que la producción depende de la demanda del mercado y ésta no es homogénea. A pesar de ello, se ha utilizado el máximo número de lotes posible para cada producto, para poder incluir la mayor variabilidad a la biblioteca (lotes de API y excipientes distintos, diferente época del año, etc.).⁷

Tabla VII - 4 Número de lotes analizados para cada producto, indicando también el número de lotes destinado a cada conjunto de muestras.

Producto	Nº lotes total	Conjunto entrenamiento		Conjunto evaluación		Conjunto validación	
		Nº lotes	Espectros	Nº lotes	Espectros	Nº lotes	Espectros
A	28	15	272	7	114	6	105
B	10	6	102	2	33	2	33
C	13	7	117	3	53	3	54
D	3	2	36	1	18	-	-
E	7	3	51	2	32	2	36
F	4	3	48	1	12	-	-

Producto	Nº lotes total	Conjunto entrenamiento		Conjunto evaluación		Conjunto validación	
		Nº lotes	Espectros	Nº lotes	Espectros	Nº lotes	Espectros
G	9	5	84	2	30	2	36
H	55	27	463	15	249	13	221
I	22	12	201	6	114	4	66
J	34	17	269	9	141	8	129
K	1	-		-		1	18
L	2	-		-		2	33
M	2	-		-		2	39

Cuando ya se disponía de los espectros de las muestras el primer paso fue la evaluación de estos espectros de cada producto con el fin de detectar los espectros anómalos.

Una vez eliminados los espectros anómalos, se realizó un PCA de cada producto. La representación gráfica de los *scores* se ha utilizado para distribuir los lotes en los tres conjuntos de muestras: entrenamiento, evaluación y validación. El conjunto de entrenamiento se utiliza para definir los modelos que, posteriormente, permitirán identificar las nuevas muestras. Para conseguir este objetivo, es importante que este conjunto de muestras incluya aquellas que presenten una mayor variabilidad del conjunto de espectros registrados. Por lo tanto, para definir el conjunto de entrenamiento, se seleccionaron los lotes que presentaban las muestras más alejadas en el gráfico de *scores* del PCA.^{VII} Los lotes restantes se distribuyeron aleatoriamente entre los conjuntos de evaluación y validación. En la Tabla VII - 4 se muestra la distribución de lotes para cada producto. Como se puede observar en la misma tabla, los productos de los que se disponía menos de 3 lotes se han utilizado sólo en el conjunto de validación, para evaluar la robustez del modelo frente a productos no incluidos en el mismo. Cuando algún producto ha tenido desde 3 hasta menos de 5 lotes, se han distribuido sus lotes entre los conjuntos de entrenamiento y evaluación.^{7,8}

Una vez eliminados los espectros anómalos y formados los distintos grupos de muestras, se ha pasado a la construcción de la biblioteca. En esta etapa de construcción, se deben escoger los parámetros que mejor permitan identificar los productos a partir de las

^{VII} Las muestras de un mismo lote siempre se han mantenido en el mismo grupo (entrenamiento, evaluación o validación).

muestras del conjunto de entrenamiento, y se ha de estudiar su comportamiento con el conjunto de muestras de evaluación. Los parámetros estudiados son los siguientes:

- Tipo de medida utilizada para evaluar la similitud entre productos. Se ha utilizado la distancia Euclídea entre espectros y la misma distancia aplicada a los *scores* de los espectros después de realizar un PCA. En este caso no se ha evaluado el uso del coeficiente de correlación debido a que, como se ha demostrado en la Tabla VII - 3, hay varios productos que presentan un coeficiente de correlación muy elevado entre ellos.
- Pretratamiento de los espectros. Los pretratamientos evaluados han sido los disponibles en el software: SNV, primera y segunda derivada (de ahora en adelante 1Der o 2Der) Savitzky-Golay con un polinomio de segundo orden y 9 puntos de ventana móvil y la combinación de ambos.
- Intervalo de longitudes de onda. Intervalo del espectro utilizado para realizar la identificación de las muestras.
- Umbral de aceptación (U). Es el valor máximo que permite asignar una muestra a un grupo de manera inequívoca. Este parámetro se define utilizando las muestras de entrenamiento.⁹
- Selectividad (S) de la biblioteca. Este parámetro se evalúa utilizando las muestras del conjunto de evaluación, y se calcula siguiendo la ecuación:

$$S = \frac{D}{U_1 + U_2} \quad \text{Ecuación VII - 1}$$

Donde D es la distancia entre los espectros promedio de los dos productos comparados y $U_1 + U_2$ la suma de los umbrales definidos por esos dos mismos productos. Dependiendo del valor de la selectividad, se pueden encontrar tres situaciones, tal como se esquematiza en la Figura VII - 2.¹⁰ Para obtener una biblioteca que permita identificar los productos con un bajo índice de falsos positivos, la selectividad debe ser siempre mayor a 1.

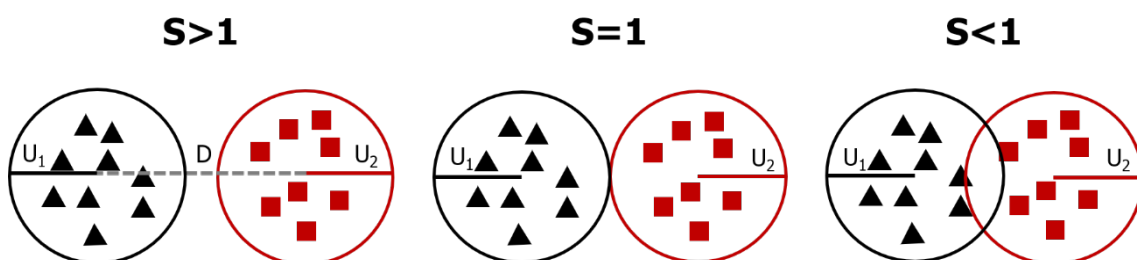


Figura VII - 2. Esquema de las posibles situaciones entre dos clases según el valor de la selectividad.

El procedimiento seguido para construir la biblioteca de espectros consistió en realizar un primer cribado con el objetivo de descartar aquellas combinaciones de parámetros que no alcanzaban a separar bien los grupos estudiados. Para poder comparar qué combinación presentaba unos mejores resultados, se ha mantenido constante el intervalo de longitudes de onda y se ha evaluado cuántos productos se podían identificar sin confusión y el % de selectividades con un valor inferior a 1, entre 1 y 2, y mayores a 2.

Tabla VII - 5. Resultados de la construcción de la biblioteca utilizando la distancia Euclídea entre los espectros y entre los *scores* en el intervalo de longitudes de onda de todo el espectro.

Medida	Pretratamiento	Nº de productos sin confusión	Selectividad		
			S<1	1≤S≤2	S>2
Distancia Euclídea entre los espectros	SNV	2	24 %	29 %	47 %
	1Der (9p)	0	38 %	60 %	2 %
	2Der (9p)	0	76 %	24 %	0 %
	1Der (9p) + SNV	3	9 %	64 %	27 %
	2Der (9p) + SNV	0	65 %	7 %	28 %
Distancia Euclídea entre los <i>scores</i>	SNV	5	4 %	31 %	65 %
	1Der (9p)	0	47 %	53 %	0 %
	2Der (9p)	0	47 %	51 %	2 %
	1Der (9p) + SNV	1	2 %	16 %	82 %
	2Der (9p) + SNV	1	7 %	11 %	82 %

Con los resultados de la Tabla VII - 5, se puede ver que en ninguno de los casos estudiados se consigue separar los 10 productos; siempre hay productos que se confunden, por lo que en todos los casos es necesaria la construcción de sub-bibliotecas. Respecto al tipo de distancia utilizada, los resultados muestran que el uso de la distancia Euclídea entre los *scores* mejora la identificación frente a la distancia Euclídea entre los espectros. En cuanto a los pretratamientos, se observa que aquellos modelos construidos con SNV (individual o combinado) muestran mejores resultados, exceptuando la combinación SNV y 2Der al utilizar la distancia Euclídea entre los espectros. Los modelos que presentan un menor porcentaje de clases con S<1, en el caso de utilizar la distancia Euclídea entre espectros, son los que combinan los pretratamientos de 1Der con SNV;

mientras que, utilizando la distancia entre los *scores*, cualquier pretratamiento que utilice SNV, tiene un bajo porcentaje de clases con $S < 1$.

A partir de esta selección de modelos, se ha acotado el intervalo de longitudes de onda, eliminando las partes del espectro que no presentan ninguna banda de absorción o que presentan una banda común en todos los productos. En este caso, se ha estudiado el rango desde 2490 nm hasta 1860 nm, desde 1775 nm hasta 1400 nm y desde 1250 nm hasta 1110 nm.

Tabla VII - 6. Resultados de la construcción de la biblioteca utilizando la distancia Euclídea entre los espectros y entre los *scores* en el intervalo de longitudes de onda desde 2490 nm hasta 1860 nm, desde 1775 nm hasta 1400 nm y desde 1250 nm hasta 1110 nm.

Medida	Pretratamiento	Nº de productos sin confusión	Selectividad		
			$S < 1$	$1 \leq S \leq 2$	$S > 2$
Distancia Euclídea entre los espectros	1Der (9p) + SNV	4	8 %	66 %	27 %
	SNV	4	7 %	24 %	69 %
Distancia Euclídea entre los <i>scores</i>	1Der (9p) + SNV	5	9 %	7 %	84 %
	2Der (9p) + SNV	2	9 %	64 %	27 %

Como se puede observar en la Tabla VII - 6, en todos los casos se sigue requiriendo el uso de sub-bibliotecas, aunque en casi todos los casos el número de productos que se separan directamente ha mejorado en comparación con los modelos construidos con todo el intervalo de longitudes de onda, exceptuando cuando se utiliza la distancia entre *scores* de los espectros tratados con SNV. También se aprecia que el modelo de la distancia Euclídea entre *scores* de los espectros tratados con SNV presenta un porcentaje de $S < 1$ menor que los demás modelos. Sin embargo, el modelo que consigue un mayor porcentaje de valores de selectividad mayores que 2, es el modelo que utiliza distancias entre *scores* después de un pretratamiento con 1Der, con una ventana móvil de 9 puntos, y SNV, por lo que se escoge como la combinación de pretratamiento y distancia para construir la biblioteca final.

Utilizando esta combinación y realizando una acotación del intervalo de longitudes de onda, se ha construido la biblioteca final, la cual separa todos los productos, excepto los dos antiviricos, A y D, para los cuales se ha creado una biblioteca de nivel 2.

Estos dos antiviricos son los que ya en la Tabla VII - 3 presentaban un elevado coeficiente de correlación. Ambas moléculas presentan la misma estructura y sólo

difieren en un grupo funcional. Representando los espectros con los distintos pretratamientos, se observó que, utilizando un pretratamiento de 1Der, con una ventana móvil de 9 puntos, y SNV, en el intervalo de longitudes de onda desde 1778 nm hasta 1707 nm los productos se diferenciaban.

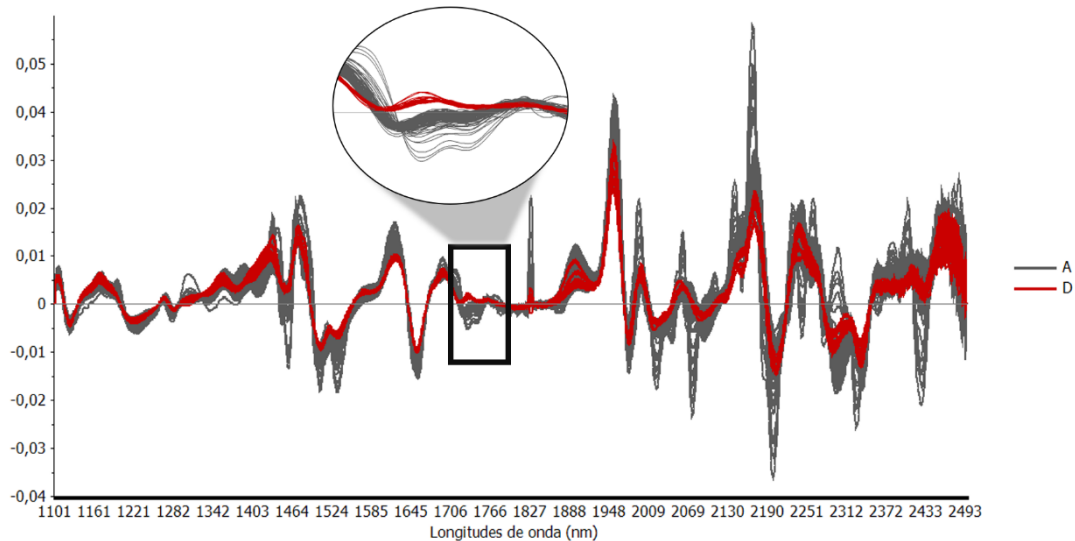


Figura VII - 3. Representación de los espectros de las muestras del conjunto de entrenamiento de los productos A y D con un pretratamiento de 1Der con 9 puntos de ventana móvil y SNV. En la misma figura se muestra un aumento del único intervalo de longitudes de onda (1707 nm a 1778 nm) donde se diferencian ambos productos.

Tanto el tipo de distancias como el pretratamiento utilizados en esta biblioteca de nivel 2, son los mismos que en la biblioteca de nivel 1, la única diferencia entre ellas es el intervalo de longitudes de onda utilizado.

En la siguiente tabla se muestran los parámetros finales de la biblioteca escogida, mostrando los distintos niveles y umbrales de aceptación.

Tabla VII - 7. Parámetros de la biblioteca final, indicando el umbral de cada grupo.

Nivel de la biblioteca	Intervalo de longitudes de onda	Nº de factores	Pretratamiento	Producto	Umbral
1	Desde 2491 hasta 1957, desde 1739 hasta 1425 y desde 1243 hasta 1111	3 (factores 3, 4 y 5)	1Der (9p) + SNV	A	0,070
				B	0,110
				C	0,072
				D	0,058
				E	0,040
				F	0,031

Nivel de la biblioteca	Intervalo de longitudes de onda	Nº de factores	Pretratamiento	Producto	Umbral
				G	0,064
				H	0,135
				I	0,054
				J	0,066
2	Desde 1778 hasta 1707	1 (factor 2)	1Der (9p) + SNV	A	0,485
				D	0,611

En la Figura VII - 4 se puede ver la representación 2D de la biblioteca de ambos niveles con todos los factores.

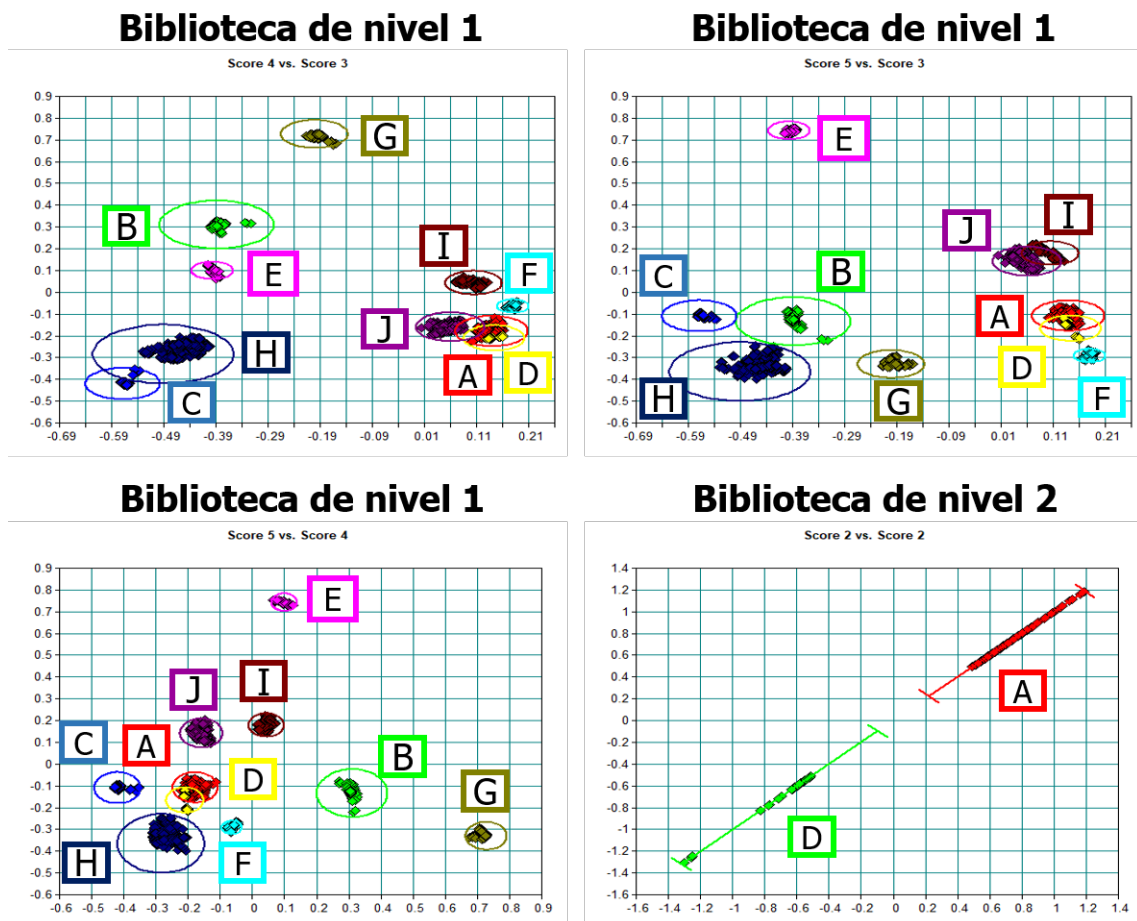


Figura VII - 4. Representación de los *scores* de las bibliotecas de nivel 1 y 2, indicando los grupos de cada producto.

Como se puede observar en la Figura VII - 4, ninguno de los valores de selectividad se encuentra por debajo de 1, lo que indica que no hay ningún grupo que se solape con otro.

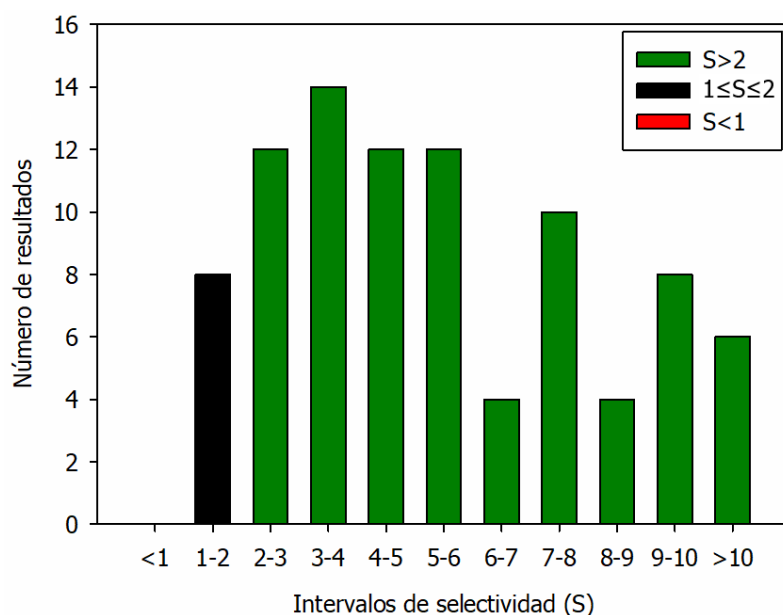


Figura VII - 5. Histograma de los valores de selectividad obtenidos.

3.3. Validación de la biblioteca

Una vez se ha construido la biblioteca de espectros NIR, se procede a su validación. Para ello se estudiaron los dos parámetros que se recomiendan en las guías de las agencias reguladoras: la especificidad y la robustez.^{8,11,12}

Para evaluar ambos, se han utilizado los lotes del conjunto de validación que se indican en la Tabla VII - 4, los cuales no han sido utilizados durante la construcción y evaluación de la biblioteca en cascada. Los espectros de las muestras de estos lotes se han identificado utilizando la biblioteca descrita en la Tabla VII - 7. Los resultados de la identificación se muestran en la Tabla VII - 8.

Tabla VII - 8. Resultados de la identificación de los 13 productos, clasificando los resultados como muestras identificadas correctamente, muestras no identificadas y muestras identificadas incorrectamente.

Producto	Muestras identificadas correctamente	Muestras no identificadas	Muestras identificadas incorrectamente
A	105	0	0
B	33	0	0
C	54	0	0
D	No utilizado en la validación.		
E	36	0	0
F	No utilizado en la validación.		

Producto	Muestras identificadas correctamente	Muestras no identificadas	Muestras identificadas incorrectamente
G	36	0	0
H	220	1	0
I	66	0	0
J	129	0	0
K	0	18	0
L	0	33	0
M	0	39	0

Como se puede observar, los distintos lotes de rutina evaluados se han identificado correctamente, lo que demuestra la robustez de la biblioteca.

Para evaluar la especificidad se ha calculado los valores de los parámetros descritos en la sección 3.3 del capítulo VI.

Tabla VII - 9. Evaluación de la especificidad de la biblioteca construida.

Positivos verdaderos	679	Tasa de falsos positivos	0
Negativos verdaderos	90	Sensibilidad	0,999
Falsos positivos	0	Precisión	1
Falsos negativos	1	Exactitud	0,999
Tasa de falsos negativos	0,002	Especificidad	1

Como se puede observar en la tabla, tanto la precisión como la especificidad de la biblioteca tienen un valor de 1. Únicamente una muestra no se ha identificado correctamente como perteneciente a su clase, aunque tampoco se ha identificado como perteneciente a otra clase de la biblioteca, por lo tanto, se ha obtenido un falso negativo. Este valor hace que la sensibilidad de la biblioteca sea de un 99,9 % al igual que la exactitud.

Además, se ha retado a la biblioteca a identificar los espectros de tres productos que no forman parte de ninguna clase, para ver si eran clasificados como muestras desconocidas.⁹ Como se puede observar en la Tabla VII - 8, las 90 muestras de los productos K, L y M, han sido identificadas como desconocidas. Esto demuestra que la

biblioteca es capaz de distinguir productos externos a ella, aunque tengan altos valores de coeficientes de correlación con los productos que conforman la biblioteca, como sería el caso del producto H y el producto L; o de los productos I y J con el producto M. Con estos resultados se puede concluir que la biblioteca de espectros NIR construida es específica.

4. CONCLUSIONES

En este trabajo, se ha desarrollado y validado de forma satisfactoria una biblioteca de espectros NIR que permite identificar 10 productos terminados de la compañía a liberación.

La biblioteca de espectros NIR consiste en un esquema de biblioteca en cascada de dos niveles, construida utilizando un conjunto de muestras de entrenamiento y evaluando su capacidad de identificación con el conjunto de muestras de evaluación. La biblioteca general se ha construido aplicando un pretratamiento a los espectros de primera derivada con 9 puntos de ventana móvil y SNV, en los intervalos de longitudes de onda desde 2491 nm hasta 1957 nm, desde 1739 nm hasta 1425 nm y desde 1243 nm hasta 1111 nm, y midiendo la distancia Euclídea entre los *scores* del PCA. En esta primera identificación se consiguen separar 8 de los 10 productos. Con la finalidad de poder identificar los dos productos que presentan confusión entre sí, se ha construido una biblioteca de nivel 2 aplicando el mismo pretratamiento y distancia, variando únicamente el intervalo de longitudes de onda (desde 1778 nm hasta 1707 nm).

Para validar la biblioteca de espectros NIR se ha utilizado un conjunto de muestras independiente (conjunto de validación). Con estas muestras se ha demostrado que la biblioteca construida presenta una sensibilidad y una exactitud del 99,9 %, una especificidad y precisión del 100 %, una tasa de falsos negativos del 0,2 % y de falsos positivos del 0 %. La robustez del sistema ha sido demostrada intentando identificar en la biblioteca los espectros de tres productos que no forman parte de ella.

Con todo esto se puede concluir que es factible construir una biblioteca de espectros para identificar los productos liofilizados terminados antes de su liberación en el mismo envase, cumpliendo con los requisitos regulatorios de la industria farmacéutica y que sería posible la implementación de este método de identificación en la compañía.

5. **BIBLIOGRAFÍA**

- (1) Blanco, M.; Coello, J.; Iturriaga, H.; Maspoch, S.; de la Pezuela, C. Near-Infrared Spectroscopy in the Pharmaceutical Industry. *Analyst* 1998, *123* (8), 135R-150R.
- (2) Luypaert, J.; Massart, D. L.; Vander Heyden, Y. Near-Infrared Spectroscopy Applications in Pharmaceutical Analysis. *Talanta* 2007, *72* (3), 865–883.
- (3) Szymańska, E.; Saccenti, E.; Smilde, A. K.; Westerhuis, J. A. Double-Check: Validation of Diagnostic Statistics for PLS-DA Models in Metabolomics Studies. *Metabolomics* 2012, *8* (S1), 3–16.
- (4) MacDonald, B. F.; Prebble, K. A. Some Applications of Near-Infrared Reflectance Analysis in the Pharmaceutical Industry. *J. Pharm. Biomed. Anal.* 1993, *11* (11–12), 1077–1085.
- (5) Baheti, A.; Kumar, L.; Bansal, A. K. Excipients Used in Lyophilization of Small Molecules. *J. Excipients Food Chem.* 2010, *1* (1), 41–54.
- (6) Clua Palau, G. Application of Quality by Design and near Infrared Spectroscopy in Manufacturing and Control of Freeze-Dried Drug Products, Universitat Autònoma de Barcelona, 2018.
- (7) Miró Vera, A. Y.; Alcalà Bernàrdez, M. Near-Infrared Spectroscopy in Identification of Pharmaceutical Raw Materials. In *Encyclopedia of Analytical Chemistry*; John Wiley & Sons, Ltd: Chichester, UK, 2017; pp 1–19.
- (8) European Pharmacopoeia. 5.21. Chemometric Methods. 2016, 5641–5658.
- (9) Blanco, M.; Romero, M. A. Near-Infrared Libraries in the Pharmaceutical Industry: A Solution for Identity Confirmation. *Analyst* 2001, *126* (12), 2212–2217.
- (10) Bruker. *User Manual IDENT OPUS Spectroscopy Software*, Versión 7.; Bruker, 2011.
- (11) Peinado, A.; Hammond, J.; Scott, A. Development, Validation and Transfer of a Near Infrared Method to Determine in-Line the End Point of a Fluidised Drying Process for Commercial Production Batches of an Approved Oral Solid Dose Pharmaceutical Product. *J. Pharm. Biomed. Anal.* 2011, *54* (1), 13–20.
- (12) The Internacional Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. *ICH Q2 (R1) Validation of Analytical Procedures: Text and Methodology*; 2011; Vol. 20.

VIII. CONCLUSIONES

1. **CONCLUSIONES**

Los trabajos presentados en esta tesis han ofrecido soluciones analíticas, capaces de mejorar la calidad del proceso de fabricación y resolver los inconvenientes concretos aparecidos durante la vida útil de distintos productos farmacéuticos liofilizados.

Centrándonos en los objetivos concretos definidos se puede concluir:

- Se ha desarrollado y validado un método basado en la valoración coulombimétrica Karl Fischer acoplada a un horno, para determinar el contenido de agua de los tapones de clorobutilo utilizados en el envase primario. En la validación del método se han seguido las recomendaciones de la ICH y se ha estudiado la especificidad (a la temperatura de trabajo únicamente se determina agua), exactitud (recuperaciones según los criterios de la Ph.Eur. a tres niveles de concentración), linealidad (pendiente y ordenada en el origen sin diferencias significativas a 1 y 0, respectivamente), intervalo de linealidad (350 μg - 7500 μg), precisión intermedia (sin diferencias significativas entre analistas para concentraciones del límite superior e inferior), repetibilidad (RSD de 9,68 %) y límite de cuantificación (103,3 μg de agua).

Este método ha permitido evaluar el comportamiento de la humedad de los tapones durante el ciclo de vida en condiciones de estabilidad acelerada de un producto parenteral liofilizado, denominado H. Se ha comprobado que, al igual que ocurre con el producto, el proceso de liofilización seca los tapones. Independientemente de la humedad inicial de los tapones, su contenido en agua aumenta con el tiempo del estudio hasta unos 2 mg. Se ha confirmado que la impureza 1 del producto estudiado está relacionada con la humedad procedente del tapón, mientras que la impureza 2 se produce a partir de la humedad residual de la muestra.

- Se ha desarrollado y validado según las recomendaciones de la ICH un método para la determinación de la humedad residual de un producto liofilizado, denominado N, basado en el uso de Karl Fischer coulombimétrico acoplado a un horno. En la validación se ha estudiado la especificidad (a partir del análisis termogravimétrico no hay degradación del producto a 80°C), exactitud (recuperaciones según los criterios de la Ph.Eur. en cinco concentraciones distintas), linealidad (pendiente y ordenada en el origen próximas a 1 y 0,

respectivamente), intervalo de linealidad (240 µg - 5260 µg), repetibilidad (RSD < 15%), precisión intermedia (sin diferencias significativas entre analistas y días para niveles altos y bajos de humedad) y límite de cuantificación (99,6 µg de agua).

Como parte integral del desarrollo del método analítico se ha realizado un análisis de riesgos basado en el análisis modal de fallos y efectos que ha permitido mitigar 8 parámetros con alto riesgo de fallo.

El método Karl Fischer, validado anteriormente para el producto N, se ha utilizado como método de referencia para el desarrollo y validación de un método rápido y no destructivo para la determinación de la humedad residual, a liberación del mismo producto. Es un método basado en la espectroscopia NIR, utilizando un modelo PLSR. El modelo se ha construido a partir de espectros NIR registrados entre 1350 t 2350 nm, con SNV como pretratamiento y 3 factores PLSR.

El método NIR ha sido validado según las recomendaciones de la EMA y de la ICH para métodos cuantitativos. Se ha estudiado la especificidad (tanto los *scores* como los *loadings* del modelo están relacionados con el contenido de agua), exactitud (a partir de una prueba t de muestras apareadas para 25 muestras en todo el intervalo de concentración), linealidad (se ha estudiado la recta de valores calculados por el modelo frente a los de referencia, y la pendiente y ordenada en el origen no presentan diferencias significativas de 1 y 0, respectivamente; el sesgo es estadísticamente igual a cero), intervalo de linealidad (para la misma recta, 0,37 % - 3,93 %), precisión intermedia (sin diferencias significativas entre analistas y días tanto para muestras de laboratorio como industriales), repetibilidad (valores máximos y mínimos de RSD=1,87 % y 1,02 %) y límite de cuantificación (desde el punto de vista normativo, tanto LOD como LOQ son el valor mínimo obtenido en el desarrollo y validación del método, 0,36 % de agua) El método ha sido implementado en el ámbito industrial con el fin de evaluar la homogeneidad del proceso de liofilización. Con los datos recopilados se ha podido concluir que, aunque todas las muestras se encuentran dentro de especificaciones, hay diferencias significativas dependiendo del liofilizador utilizado; con una menor humedad residual media en el liofilizador adquirido más recientemente. En el estudio de cada lote para evaluar las diferencias entre pisos y posiciones, se han observado diferencias de algún piso respecto al resto, aunque únicamente en uno de los lotes había diferencias en las posiciones. A

pesar de ello, no se ha apreciado una sistemática que relacione pisos y posiciones concretas con el secado en el liofilizador.

- En esta tesis se ha conseguido desarrollar y validar, de forma satisfactoria, un método cualitativo para la clasificación de las muestras del producto liofilizado etiquetado como J, según la concentración de oxígeno del *headspace* del vial. Se han utilizado medidas de reflectancia en el visible de los mismos espectros que se registran para hallar el RMC del producto en la zona NIR. Se ha utilizado un modelo PLS-DA para relacionar el contenido rosado del sólido con el contenido de oxígeno y separar las muestras con alto y bajo contenido a partir de un valor umbral del 2 % de oxígeno.

El estudio del comportamiento del método y su validación se han realizado teniendo en cuenta que se trata de un método cualitativo, utilizando los resultados de verdaderos y falsos negativos, así como los de verdaderos y falsos positivos, para el cálculo de los parámetros. El método propuesto tiene una tasa de falsos positivos del 2 % y de falsos negativos del 7 %, con una sensibilidad del 93 % y una especificidad del 98 %; la precisión es de un 99 % y la exactitud del 95 %. Se ha utilizado la curva característica de rendimiento para el cálculo de zona de falta de fiabilidad del método, comprendida entre un 1,13 % y un 3,5 % de oxígeno. Un futuro análisis completo a caducidad con muestras a distinta concentración de oxígeno puede llevar a una revisión del umbral.

- Se ha desarrollado y validado un método para la identificación, en el propio envase, de productos terminados antes de su liberación. Se ha utilizado la espectroscopia NIR para crear una biblioteca representativa de los espectros NIR de 10 productos y se han utilizado técnicas quimiométricas de PRM para su evaluación y validación. En la validación se han utilizado 8 productos distintos presentes en la biblioteca, con un total de 679 muestras, y 3 productos no presentes en ella, con un total de 90 muestras. Ninguna de las muestras ha sido identificada incorrectamente, todas las muestras de productos ausentes han sido identificadas como no pertenecientes a la biblioteca y únicamente una muestra de un producto, con 221 muestras, ha sido no identificada, aunque no confundida con otro producto. A partir de los resultados hallados, se puede concluir que es viable la implementación de un método de identificación multivariante para identificar los principales productos de la compañía, cumpliendo con los requisitos regulatorios de la industria farmacéutica.

CONCLUSIONES

Además de los objetivos comentados, con el desarrollo de esta tesis se ha podido realizar una transferencia de conocimiento de la universidad a la industria, cumpliendo así con el objetivo principal del proyecto de doctorados industriales.