Extracting useful information from evolving FTIR data sets using MCR-ALS and 2DCoS

Bernhard Lendl

Institute for Chemical Technologies and Analytics
Vienna University of Technology, Austria
Outlook

• Information Content in IR and Raman Spectra
• Chemometric Methods
  – MCR-ALS: Multivariate Curve Resolution – Alternating Least Square
  – 2DCoS: Two Dimensional Correlation Spectroscopy
  – Enzymatic hydrolysis of albumin by proteinase K
• Concept for time resolved FTIR spectroscopy of chemical reactions in solution
• Chemistry on the fly
• Summary
Information Contained in Mid-IR and Raman Spectra

- Functional groups, fingerprint
  - Identification of analytes
  - Simultaneous determination of several analytes

- Inter- and intramolecular Interactions
  - Determination of secondary structure of Proteins

- Information on latent variables
  - Octane number, wine varieties, cancer
Example: Elucidation of Secondary Structure of Proteins

Albumin
α-Helix

Concanavalin
40% sheet

Minispec, 2007
Example: Determination of Latent Variable Classification of Red Wines - Concept

Solid Phase Extraction
1. Conditioning
2. Sample loading
3. Washing
4. Elution

1 µL of wine extract

Diamond

ZnSe Focusing Element

From Interferometer
To Detector

Minispec, 2007
Classification of Red Wines - Results

Spectra of phenolic extracts

Hierarchical clustering

ZW: Zweigelt, SL: St. Laurent, BF: Blaufränkisch,
ME: Merlot, CS: Cabernet Sauvignon, PN: Pinot Noir

Minispec, 2007
Multivariate Curve Resolution (MCR-ALS)

- Concept: Modeling of experimental data matrix $D$
- Multi-component systems can be often be described with a simple model consisting of the composition-weighted sum of signals of their pure components
- Only condition: linear structure of data set

$$D = CS^T + E$$

$D(r \times c)$ is the original data matrix
$C(r \times n)$ and $S^T(n \times c)$ contain pure response profiles related to the data variation in the row $(r)$ and column $(c)$ direction of $D(r \times c)$
$E(r \times c)$ is the error matrix
Multivariate Curve Resolution: Graphical Representation

\[ \begin{align*}
\text{exp. data } D & = C + E \\
& = c_1 s_1^T + c_2 s_2^T + c_3 s_3^T + E
\end{align*} \]

c_i...concentration profiles with time    s_i...spectra of pure substances
Uncertainty in MCR Results due to Ambiguities

• Magnitude ambiguity:

\[ D = \sum_{i=1}^{n} \left( \frac{1}{k_i} c_i \right) (k_i S_i^T) \]

\[ D = C'S'^T \]

• Rotational ambiguity:

\[ D = CS^T \]
\[ D = C(TT^{-1})S^T \]
\[ D = (CT)(T^{-1}S^T) \]
\[ D = C'S'^T \]

Aim to reduce these ambiguities
Exploratory Data Analysis

• Important to find good (correct) initial estimates for starting the iteration process in of MCR
  – Initial estimates on concentration profiles and responses (spectra) need to be made

• Typical questions to be answered:
  – How many components shall be assumed to be present?
    • PCA, EFA, FSMW-EFA,…
  – Is there an special „purest“ variable in concentration or spectral axis available?
    • SIMPLISIMA, PCA
Constraints

• Property which are fulfilled for the whole or part of the system

• Examples:
  – Non-Negativität
  – Unimodality
  – Mass balance (closure)
  – Application of physico-chemical models
  – Known spectra / Concentration profile (are set invariant during the iteration process) => access to quantitative information
  – Local rank,…
Characteristics of MCR - ALS

• Clear criteria for stopping the iteration process

\[
\% \text{ lack of fit} = 100 \left[ \frac{(\sum r_{ij}^2)}{(\sum d_{ij}^2)} \right]^{1/2}
\]
Typical value: 0.1 %

• Quantitative information
  – Number of components
  – Spectral properties
  – Concentration profiles

• Simultaneous analysis of several data sets

• Limitation: Difficulties in definition of spectral properties of „components“
Overview: Two Dimensional Correlation Spectroscopy (2DCoS)

- Perturbation: Mechanical, electrical, Chemical, magnetic, Optical, thermal, etc.
- Spectroscopic Analysis (e.g., IR, UV, ..)
- System
- Dynamic Spectra

I. Noda, Lecture 2DCoS-3, 2005

Minispec, 2007
2DCoS

Idea: Looking for correlation in the experimental output $D$

Correlation analysis

\[
\Theta(\nu_1, \nu_2) + i\Psi(\nu_1, \nu_2) = \frac{1}{\pi T} \int_{0}^{\infty} \overline{Y_1}(\sigma) + i\overline{Y_2}(\sigma)
\]

Time resolved spectra

Time/pressure/
Temperature,..

Spectral Variable

I. Noda Lecture 2DCoS-3, 2005
Synchronous Correlation Map

Autopeaks at diagonal positions represent the extent of perturbation-induced dynamic fluctuations of spectral signals.

Cross peaks represent simultaneous changes of spectral signals at two different wavenumbers, suggesting a coupled or related origin of intensity variations.

Same signs at cross peaks show simultaneous increase/decrease at corresponding wavenumbers.
Different signs: different behavior.
Asynchronous Correlation Map

Cross peaks appear only if changes at corresponding wavenumbers are occurring at a different rate.

The **signs** of synchronous and asynchronous cross peaks become the **same** if the intensity change at \( \nu_1 \) occurs **before** \( \nu_2 \).

The **signs** of synchronous and asynchronous cross peaks become **different** if the intensity change at \( \nu_1 \) occurs **after** \( \nu_2 \).
Characteristics of 2DCoS

- Through emphasizing spectral changes of the system under study in the two-dimensional correlation maps a kind of “resolution enhancement” is achieved.
- No model (components, spectra) => does not require estimates.
- No quantitative information.
- No objective criteria telling success of 2DCoS analysis.
- Frequently very useful to get a „feeling“ for the behavior of the system under study.
Enzymatic Hydrolysis of BSA with Proteinase K

Secundary structure:
67% α-helix,
10% β-turn
23% extended chain,
No β-sheet

Proteinase K
Family of subtilisin proteinases
No selectivity for cleaving at certain amino acids
High activity at 50 – 60°C

Heating of BSA:
Until 50°C reversible conformational changes
Unfolding of α-helix irreversible from 52 - 60°C
Continued Temperature Increase:
Start of β-aggregation
From 70°C on gel formation takes place

Minispec, 2007
Experimental Conditions – Recorded FTIR Spectra

30 mg/ml BSA
0.5 mg/ml Proteinase K
T = 60°C
Phosphate buffer pD 7.4
Reaction time: 320 min

50 µm CaF$_2$
Spectr. res.: 2 cm$^{-1}$
According to PCA 2 components explain 99.99% of spectral variance => Assumption of 2 components
2DCoS Analysis

Synchronous Map

Asynchronous Map

1654: α-Helix
1641: unordered structure.
1594: Carboxylate
1675: β-turn
1616: β-sheet

There must be more than 2 components!!
Following “Noda Rules“: 1654/1641; 1594/1675
Syn + / asyn - => \( \nu_1 \) after \( \nu_2 \)

Minispec, 2007
Residual Matrix $E$

$$E (r,c) = D - (CS^T)$$
Residual Matrix $E$ studied with 2DCoS

Strongest band at 1616 1/cm

=> At least one more component with band at 1616 1/cm
MCR-ALS Analyse II

- Assumption of 3 components, PCA explains now 99,999 %
- Analysis of a single experiment

Obviously there are 2 products: Intermediate and final product
This is unlikely from a chemical viewpoint
MCR – ALS of Experiment without Enzyme

- Only temperature induced effects are visible

Denaturation due to increasing temp.:
Increasing portion of disordered structures:
Amide I: 1654 -> 1651 1/cm
Also slight increase in content in β-turns:
Shoulders at 1616 (strong) and 1685 are being formed

Minispec, 2007
Matrix Augmentation

• Simultaneous Analysis of several experiments
Result of MCR-ALS Matrix Augmentation

- Native Albumin
- Defolded Albumin (random structures shift toward 1641 as well as $\beta$-sheet)
- Reaction product

Results of Bioligand Interaction Studies

Purpose: Screening of Potential Drugs
Principle: Comparison of two FTIR Spectra

Füchsle, PhD thesis 1999
Results of Bioligand Interaction Studies

Difference Spectra

- Benzamide
- Leupatin
- P-Leupatin
- Benzaldehyde
- Hydrocortisone

Active Molecules => Possible Drug

Inactive Molecules

Fuchsle, PhD thesis 1999
Difficulties of TR FTIR of Chemical Reactions in Solution

• FTIR spectroscopy
  – Small spectral changes expected =>
  – Signal averaging
  – Repeated initiation of the event under study (chemical reaction) for high time resolution

• Liquid handling
  – Fast, reproducible mixing of two liquids
  – Low sample consumption

Minispec, 2007
Short Optical Path in FTIR Imposes Difficulties for Rapid Mixing

• Flow rates for fast turbulent mixing require high pressure
  – Large reagent consumption
    (repeated experiments for good s/n-ratio)
  – Problems with constant optical path

• Low flow rates (max. 0.5 ml/min)
  – No problem with constant optical path, reduced reagent consumption, but
  – Reynolds number: \( \sim 10 \) =>
    \textbf{Strongly laminar flow:...slow mixing in conventional systems}

Minispec, 2007
TR - FTIR Spectroscopy in Aqueous Solution

Problem:
• Optical pathlength in the µm range
• Strongly laminar flow
• No turbulence

Solution: Miniaturization
\[ t = \frac{L^2}{D} \]
\[ t: \text{ time} \]
\[ L: \text{ diffusion distance} \]
\[ D: \text{ Diffusion coefficient} \]

Mixing is based on diffusion only!!

• Reduction of the diffusion length
• Generation of short inter-stream distances

Minispec, 2007
“Lasgane“ Micromixer

Reduction of diffusion lengths by superposition of streamlines

Cross section through the flow-cell

IR - Beam

Reagent B
Reagent A

20 µm

~ 1 mm

\[ t = \frac{L^2}{D} \]

Minispec, 2007
Details of the Fabricated Micro-mixer

„Lasagne mixer“

Schematic view

SEM pictures

Minispec, 2007
Operation of the Flow System

- Flow rate: 100 µl/min
- Linear flow rate: 11 cm/s
- Residence time: 22 ms
- Reynolds number: 1
- Optical path: 12 µm
Principles of Rapid Scan FTIR spectroscopy

FTIR data acquisition

Operation of the Flow System

Stopped flow: Diffusion based mixing => reaction
Flow “on”: Spaghetti stream lines => no reaction

Averaged Interferograms
Time resolved FTIR Spectra

Mirror movement [cm]

Single interferograms

Time [s]

Wavenumber [1/cm]

Δt

Minispec, 2007
Experimental Set-up
Details on the Mixing Process

With boundary layer, $X_A = 0.01$ in top channel, flow rate: 50 µl/min each cross sections every 500µm, cell size 2µm

$D = 10^{-9}$
Details on the Mixing Process

With boundary layer, $X_A = 0.01$ in top channel, flow rate: 50 µl/min each cross sections every 500µm, cell size 2µm

$D = 10^{-9}$
Details on the Mixing Process

With boundary layer, $X_A = 0.01$ in top channel, flow rate: 50 µl/min each cross sections every 500µm, cell size 2µm

$D = 10^{-9}$
Test of Experimental Set-up

Performance test with Acetic Acid - NaOH system
Interaction of Vancomycin and Ac-L-Lys(Ac)-D-Ala-D-Ala

Vancomycin:
- Glycopeptide antibiotic
- binds to –Lys-D-ala-D-ala end of cell-wall precursors of gram positive bacteria
- binding proceeds via
  - hydrogen bonding, and
  - hydrophobic interactions

Minispec, 2007
Spectral Differences (static) due to Binding

Vancomycin (V)
Tripeptide (TP)

Sum of V+TP
Formed Complex of V-TP

Minispec, 2007
Time Resolved (ms) FTIR Data

Original Spectra

\[ \Delta t = 65 \, \text{ms} \]

Difference Spectra

Minispec, 2007
Multivariate Curve Resolution: Graphical Representation

c_i...concentration profiles with time
s_i...spectra of pure substances
Result of MCR-ALS Analysis of the Data

Concentration profiles

2-step binding process

before reaction
intermediate endproduct (complex)

FTIR spectra

Minispec, 2007
TR-FTIR Provides Complementary Information to Crystallography

X-Ray Crystal Structure of Cocrystallized Vancomycin - $N_{\alpha},N_{\omega}$-Diacetyl-L-Lys-D-Ala-D-Ala

In crystal, complex is associated to hexamer
Motivation

• Develop a versatile platform for on-line monitoring of arbitrary chemical reactions in nL volumes

• Application in all fields of chemistry
  – Analytical (Surface Enhanced Raman Spectroscopy)
  – Synthesis (combinatorial chemistry)
  – Bio-medical and bio-chemical
FT Raman Spectroscopy of Levitated Drops

Minispec, 2007
FT Raman Spectrum of ~300 nL Ethanol
Surface Enhanced Raman Scattering

1. Rough nobel metal surfaces (Ag, Au,....) Colloids or rough solid state substrates

2. Match between absorption of SERS substrate and excitation wavelength

High electrical fields in close proximity to surface => enhanced Raman scattering
Single molecule detection possible
Single Molecule Detection Using Surface-Enhanced Raman Scattering (SERS)

Raman microscope - Silver colloid

Results of Single Molecule Detection by SERS

- 1174 cm\(^{-1}\) Raman line
- Average of 0.6 crystal violet molecules
Desired Characteristics of SERS Substrate Preparation

- Simple and reproducible generation of SERS active surfaces
- High reproducibility of enhancement factors
- No memory effects
- Fast
Frequently Encountered Problems Using SERS Spectroscopy

• Colloidal SERS substrates
  – Tedious preparation procedure
    (Lee-Meisl method: refluxing for reduction of AgNO₃ with citrate)
  – Ageing of SERS colloids, stability
  – Batch to batch reproducibility
  – Adsorption to walls, tubings
    => memory effects
Solution: I) Hydroxylamine Reduced Silver Sol

\[
\text{NH}_2\text{OH} + \text{Ag}^+ \rightarrow \text{Ag} + \text{N}_2,.. \ (\text{N}_2\text{O})
\]
Solution: II) SERS Synthesis and Application in Levitated Droplets

Raman spectrometer

Ultrasonic trap

Micro-dispenser

Automated flow system

collection optic

Reflector

Sample

Sound pressure

Ultrasonic vibrator

Waste

Hydroxylamine in NaOH

Analyte NaCl

AgNO₃

Water

Minispec, 2007
Flow-through Micro-dispenser

Operation in flow-through mode

- Droplets: 50 pL – 300 nL
- Single shot – 500 Hz
- Initial velocity: 1 m/s

Co-operation with T. Laurell, Lund Sweden

Minispec, 2007
50 Picoliter Droplets at 100 Hz

Dispensing_640.avi
Detail of Experimental Set-up

- Ultrasonic Levitator
- Sample
- Micro-dispenser

Minispec, 2007
Detail of Experimental Set-up
On-line Synthesis of Ag-sol and SERS Spectra of 9-Aminoacridine in a Drop

Injection of
- 432 nL $10^{-3}$ M AgNO$_3$
- 48 nL Hydroxylamine $10^{-2}$ M
- 6 nL 9-Aminoacridine $10^{-5}$

Minispec, 2007
Sensitivity in the Low Femtogram Region

Analyte: 9-Aminoacridine

Minispec, 2007
Reproducibility in SERS Spectroscopy

Advantage of the micro-dispenser / levitation - concept

• Off-line preparation of Ag–colloids:
  r.s.d. ~ 30 %

• Preparation of Ag-colloids using automated flow systems:
  r.s.d ~ 7 %

• In-situ synthesis and use of Ag-colloids in a levitated droplet:
  r.s.d. ~ 2 %
Summary

• Chemometrics
  – MCR-ALS delivers easily understandable results
    • Limitation in case it gets difficult to determine the number of components present
  – 2DCoS delivers results difficult to explain at first glance, difficult to extract time sequence of events
    • But very useful for selecting initial estimates required in MCR-ALS
      – Hint on number of components
      – Initial estimates
  – Combination of both techniques advantageous

• Time resolved FTIR spectroscopy of chemical reactions in solution

• Chemistry-on-the-fly: New approach in chemical reaction monitoring
Acknowledgement

• Co-workers:
  Josef Diewok, Peter Hinsmann (TU Vienna)
  Mercedes Lopez-Pastor (Universidad de Córdoba, Spain)
  María Jose Ayora Cañada (Universidad de Jaen, Spain)
  Ana Dominguez Vidal (TU Wien)

• Financial support:
  Spanish Ministry of Education and Science
  Austrian Science Fund, FWF
  European Union, Marie Curie Training Site, ADVIS

www.iac.tuwien.ac.at/cavs