RESOLUTION METHODS APPLIED TO THE ANALYSIS OF SPECTROSCOPICAL DATA FROM BIOANALYTICAL PROCESSES

UNIVERSITAT DE BARCELONA

Joaquim Jaumot and Raimundo Gargallo

Grup de Quimiometria, Dept. de Química Analítica, Universitat de Barcelona. Av. Diagonal 647, E-08208, Barcelona, Spain

INTRODUCTION

The study of conformational transitions of biomolecules (proteins and/or nucleic acids) and the interactions between them and other species (metal complexes, organic compounds, …) has traditionally been carried out by monitoring the spectral changes at a single wavelength.

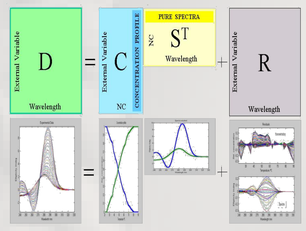
This method has certain drawbacks such as the difficulty in the estimation of the number of species when no selective wavelengths for every species of conformations is present. These difficulties can be solved by applying multivariate data analysis methods.

Among all the multivariate analysis methods, **Multivariate Curve Resolution by Alternating Least Squares (MCR-ALS)**[1-3] has proved to be a powerful tool for the resolution of complex systems without previous knowledge. Moreover, this method allows the simultaneous analysis of one experiment monitored with several techniques and several experiments monitored with one or more techniques.

In this work, examples of the usefulness of MCR are shown. These examples showed some of the different biomolecular interactions and spectroscopical techniques that might be analyzed by MCR.

THEORY OF MULTIVARIATE CURVE RESOLUTION

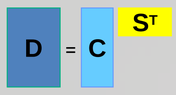
Resolution methods allow the decomposition of the initial data matrix **D** into the product of two data matrices **C** and $\mathbf{S}^T$, each of them including the pure response profiles associated with the row and column direction of the initial data matrix.



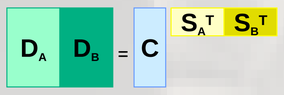
MCR is a procedure based on Factor Analysis methods that solve the equation ($\mathbf{D} = \mathbf{C}\,\mathbf{S}^T + \mathbf{R}$) iteratively by an Alternating Least Squares (ALS) algorithm. This optimization requires initial estimations either of **C** or $\mathbf{S}^T$ and allows the application of several constraints in the spectra and or concentration profiles to give more chemical meaning to the optimal mathematical solution.

Simultaneous MCR analysis of multiple experiments under different conditions allows the resolution of complex experimental data structures. There are 4 different data matrices structures:

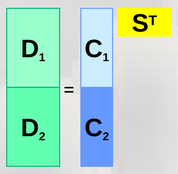
A) Individual Data Matrix.



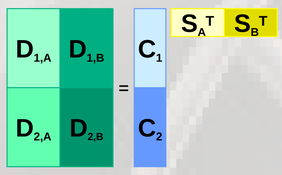
2) Data Matrix built up by one experiment monitored by several spectroscopies.



3) Data Matrix built up by several experiments monitored by one spectroscopy.

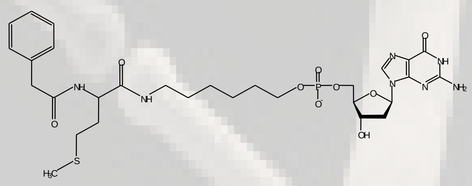


4) Data Matrix built up by several experiments monitored by several spectroscopies.



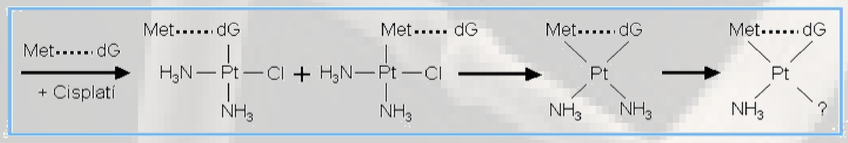
INTERACTIONS BETWEEN CISPLATIN AND A OLIGOPEPTIDE BY 2D-NMR

The wide use and effectiveness of cisplatin as anticancer drug makes the study of the reaction of this drug with Phac-Met-linker-p5'dG of interest. The therapeutic effect of cisplatin is believed to result from the formation of a chelate with two adjacent guanines in a DNA chain.



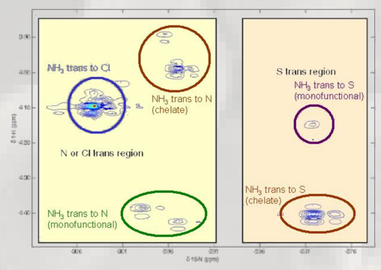
From previous 1D NMR and HPLC studies, a mechanism for the platinaction reaction was proposed.:

At the start of the reaction two products are formed: Monofunctional Pt-S adduct (major product) by the coordination of Pt with the S atom of the methionine and Monofunctional Pt-N7 adduct (minor product) by the coordination of platinum with the N7 atom of the guanine.

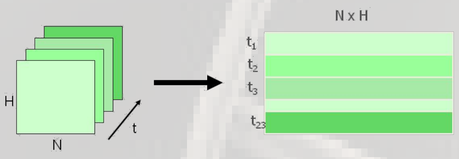


These two Monofunctional adducts evolve to the Pt-{S,N7} chelate. This chelate loses the amine in trans to the S atom to yield the final product

The reaction was monitored by [^{1}H,^{15}N]-HSQC-NMR because this is a technique especially useful to monitor complexation reactions in a very simplified way. Only correlation signals corresponding to the ^{15}N-labelled amine bonded to the platinum atom are observed in three different regions of the ^{15}N chemical shifts.

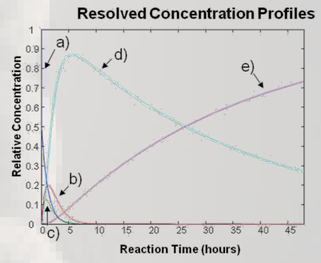


Each 2D NMR correspond to a complete data matrix (size H x N). When the whole reaction time was considered, a three-way array was built up (size H x N x t).

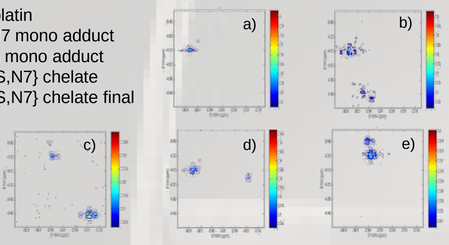


MCR-ALS cannot work directly with three-way data arrays and an unfolding of the three-way array to a two-way data matrix was required previous to the analysis.

MCR resolved concentration profiles and spectra for each one of the species [4] are:



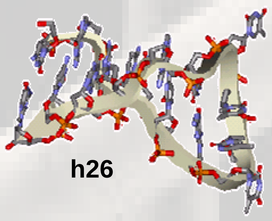
a) Cisplatin
b) Pt-N7 mono adduct
c) Pt-S mono adduct
d) Pt-{S,N7} chelate
e) Pt-{S,N7} chelate final



TRIPLE HELIX FORMING OLIGONUCLEOTIDES

There is an increasing interest in the biophysical processes involving triplex structures due to their potential applications in antigene therapy. A detailed knowledge of the conditions for triplex formation and stability is required. Although parallel triplexes are easily detected, antiparallel triplexes seem to be more interesting from the biomedical point of view. The study of antiparallel triplexes formation by UV spectroscopy is frequently hindered by the fact that only a small hypochromism is observed.

The triplex forming equilibria between the the single stranded S11CT (5'- CTTCCTCCTCT -3') and S11AG (5'- AGAGGAGGAAG-3') and the target hairpin structure h26 (5'- GAAGGAGGAGA - TTTTT-TCTCCTCCTC -3') were studied[5].

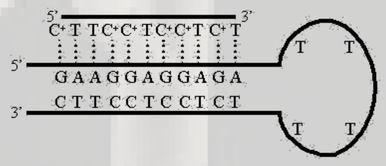


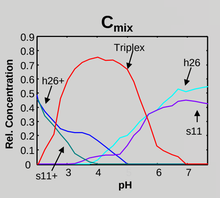
In each case, three pH titrations were performed:

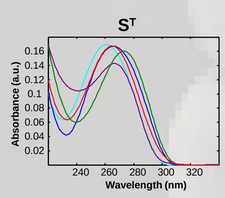
- Target hairpin h26
- S11CT or S11AG
- Mixture of S11CT and h26, or S11AG and h26

Rank analysis of the spectral data obtained in the titration of the mixture indicated that 3 species are present indicating a rank deficiency problem. Only the simultaneous analysis of the three experiments allowed the complete resolution of the experiment.

Parallel triplexes are defined by triads d(T-A-T) and d(G-C-C+). Protonation of the N3 of the cytosines is required to form the bond.

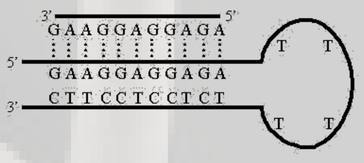


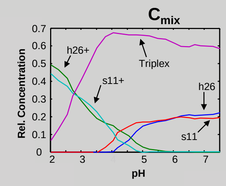


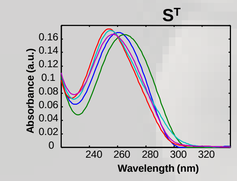


Concentration profiles showed clearly the pH dependence of the formation of the parallel triplex

Antiparallel triplexes are based on d(C-G-G), d(T-A-A) and d(T-A-T). No protonation is required and pH-independent binding is observed







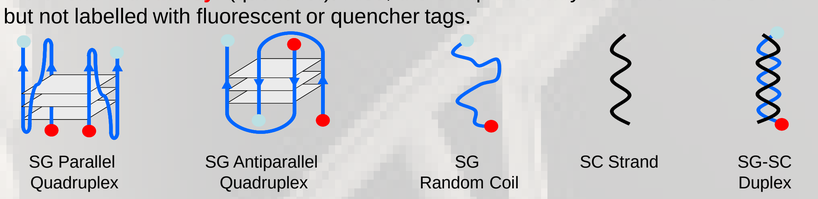
Despite of the small changes of the spectra, a high concentration of antiparallel triplex could be determined at neutral pH

G-QUADRUPLEX FORMING OLIGONUCLEOTIDES

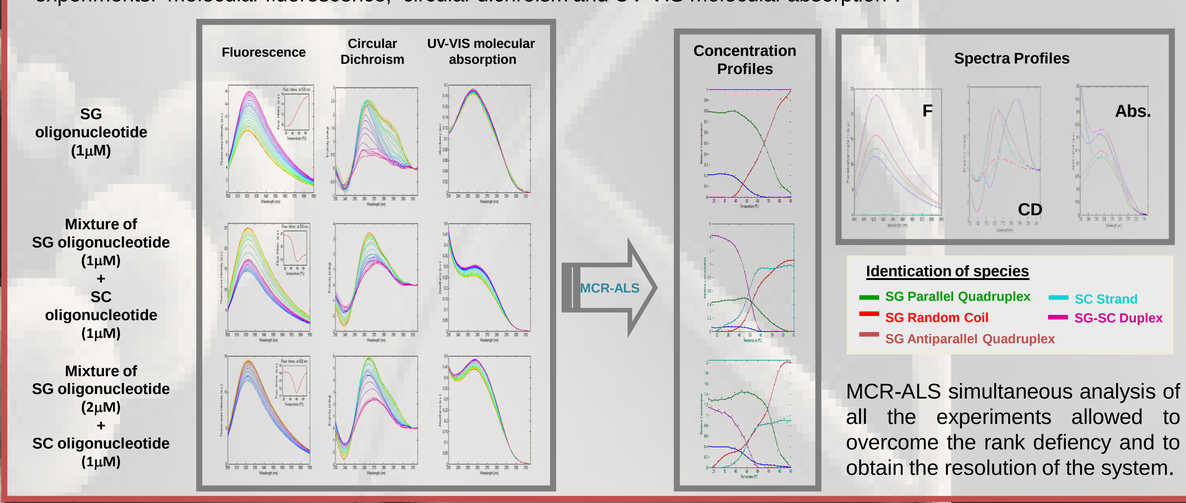
DNA G-quadruplexes have generated considerable attention because they could be molecules of interest for drug design. Oligonucleotides containing one, two or four G-stretches can form tetrameric, dimeric or monomeric G-quadruplexes, respectively. In dimeric G-quadruplexes, parallel and antiparallel structures have been characterized and preferred structures depend on the nature of monovalent cations, such as potassium or sodium.

The oligonucleotide 5'> F - TAG GGT TAG GGT – Q <3' (SG) was labeled at the 5' end with **fluorecein** (fluorophore, λ_{exc}=492 nm., λ_{em}=520 nm.) and at the 3' end with **dabsyl** (quencher). Also, the complementary strand 5'> ACC CTA ACC CTA <3' (SC) was synthesized but not labelled with fluorescent or quencher tags.

Melting experiments of SG and mixtures of SG and SC at different concentration ratios were performed. 5 different species can be expected according to the literature:



3 experiments have been carried out and 3 different spectroscopic techniques have been used to monitor these experiments: molecular fluorescence, circular dichroism and UV-VIS molecular absorption[6].

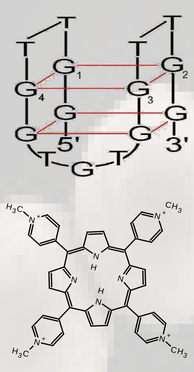


MCR-ALS simultaneous analysis of all the experiments allowed to overcome the rank defiency and to obtain the resolution of the system.

INTERACTION BETWEEN THE TBA G-QUADRUPLEX AND THE TMPYP4 PORPHYRIN

The G-quadruplex structure has been found in telomers, entity related with the cycle of life in the cell, and in some oncogens and aptamers. These parts of the genome have been suggested as potentials targets for anti-cancer therapies. For this reason, nowadays, there is an increasing interest in the development of drugs which could stabilize this structure. Several interaction mechanisms between TmPyP4 and TBA have been proposed. One of them shows the porphyrin inserted between the two planes defined by the G-tetrads. Another mechanism shows the drug interacting through the loop.

The interaction between the TBA G-quadruplex and the TmPyP4 porphyrin has been studied using RP-HPLC-DAD and two different spectrophotometric techniques (UV-VIS molecular absorption and circular dichroism)[7].

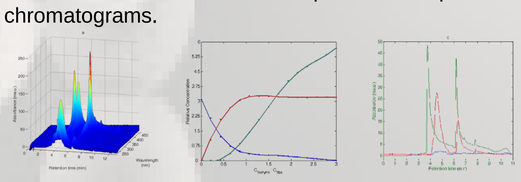


Mole- ratio experiments by HPLC and CD-UV spectroscopies

HPLC-DAD

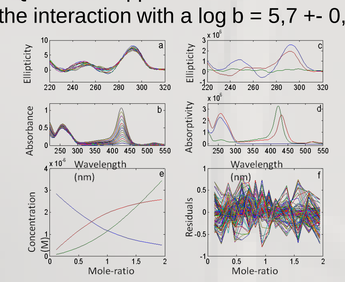
A new species related to the complex appeared when the ligand concentration was increased.

MCR analysis of the HPLC-DAD data matrix allows to recover mole-ratio concentration profiles and species chromatograms.



CD-UV SPECTROSCOPIES

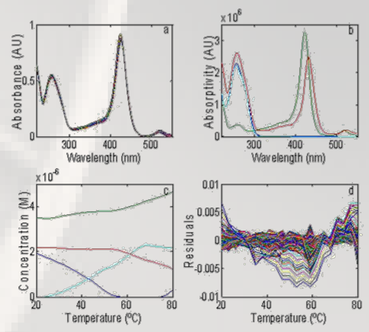
EQUISPEC approach allows to model the interaction with a log b = 5,7 +- 0,1



Thermal stability of the complex

UV melting experiments were carried out to determine the stability of the complex.

Melting temperature of the TBA quadruplex is approximately 60 ºC and the complex 45ºC



CONCLUSIONS

MCR-ALS has been shown to be particularly useful in the cases when no prior information is available and there is no possibility to postulate a known chemical model.

MCR success in the resolution of complex data sets can be explained by:

- ✓ **Flexibility** in the application of constraints during the optimization,
- ✓ **Adaptability** to analyze spectroscopic data from different spectroscopic techniques,
- ✓ **Simplicity** to analyze several experiments at the same time (changing experimental conditions and/or spectroscopic techniques)

REFERENCES

1. Tauler, R.; Kowalski, B. *Journal of Chemometrics* (**1995**), 9(1), 31-58.
2. Tauler, R. *Chemometrics and Intelligent Laboratory Systems* (**1995**), 30(1), 133-146.
3. Jaumot, J; Gargallo, R; de Juan, A; Tauler, R. *Chemometrics and Intelligent Laboratory Systems* (2005), 76(1), 101-110.
4. Jaumot, J.; Marchan, V.; Gargallo, R.; Grandas, A.; Tauler, R. *Analytical Chemistry* (**2004**), 76(23), 7094-7101.
5. Jaumot, J; Aviñó, A; Eritja, R; Tauler, R; Gargallo, R. *Journal of Biomolecular Structure & Dynamics* (**2003**), 21(2), 267-278.
6. Jaumot, J; Eritja, R; Tauler, R; Gargallo, R. *Nucleic Acids Research* (**2006**), 34(19), 206-216.
7. Del Toro, M.; Gargallo, R.; Eritja, R.; Jaumot, J. *Analytical Biochemistry* (**2008**), 379(1), 8-15.

ACKNOWLEDGEMENTS

This research was supported by MEC (CTQ2006-15052-C02-01).