Study of conformational equilibria of G-quadruplex and i-motif structures by means of multivariate analysis

Sanae Benabou[1], Anna Aviñó[2], Ramon Eritja[2], Anna de Juan[1], Raimundo Gargallo[1]

[1] *Department of Chemical Engineering and Analytical Chemistry, University of Barcelona, Barcelona, Spain (***ub.edu/gesq/dna***)*

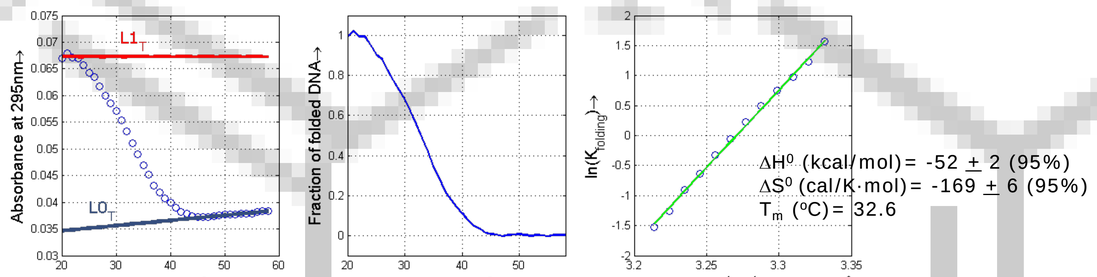
[2] *Institute for Advanced Chemistry of Catalonia (IQAC-CSIC), Networking Center on Bioengineering, Biomaterials and Nanomedicine, Barcelona, Spain*

1.Overview

- Apart from the well-known double-stranded helix firstly proposed by Watson and Crick, DNA may adopt other complex structures, such as *i*-motif or G-quadruplex [1-3].
- The **stability** of these structures depends on **factors** like pH, temperature, ionic strength, and the presence of ligands, such as proteins or drugs.
- The knowledge about the stability of these structures is obtained from biophysical experiments. Traditionally, these experiments have been monitored measuring an **instrumental signal** at one single channel. The **univariate approach** has several drawbacks like the difficulty of modeling systems where intermediate species are present.
- Here, several **examples of application** of multivariate approaches to the study of complex DNA structures are given.

2.The univariate approach

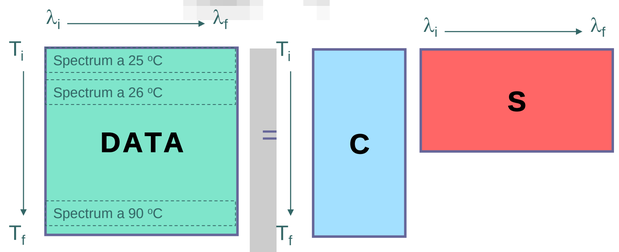
- As example, melting experiments are usually monitored at one single wavelength, such as 260 nm (for duplex structures) or 295 nm (for G-quadruplex or i-motif structures).
- The univariate approach is fast and very useful when a two-state process is involved. However, for processes involving intermediates and/or when baseline is difficult to be drawn, this univariate approach **may produce erroneous results**.



3.The simplest multivariate approach: analysis of a single data matrix

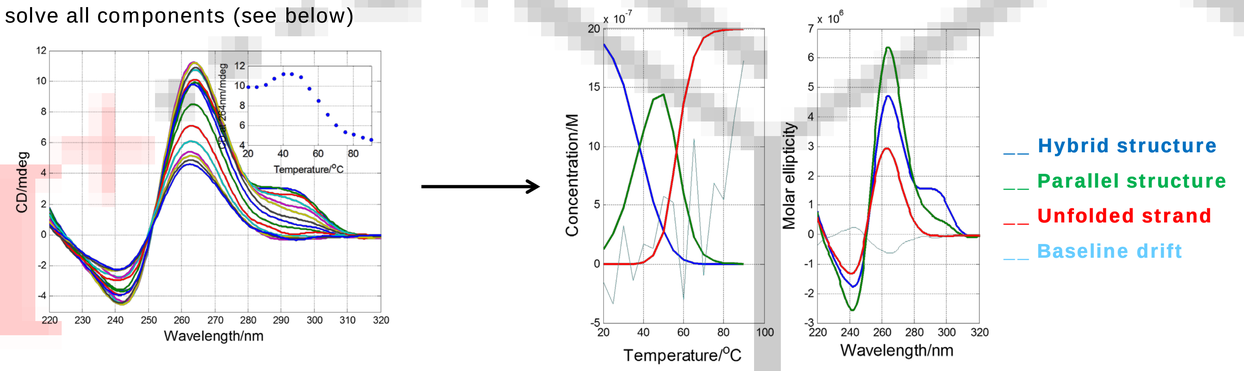
Using appropriate multivariate methods, it may be possible:

1. To determine the number of species of conformations present throughout the experiment,
2. To quantify their relative concentration (**distribution diagram, matrix C**),
3. To recover their **pure spectra (matrix S)**



- The analysis of **spectroscopic data** may be done:
  - Applying a **physicochemical model**, such as the mass action law (for chemical equilibria) or the van't Hoff equation (for melting experiments): **hard-modeling**
  - **Avoiding the application of a physicochemical model**: **soft-modeling**
  - A **combination** of both procedures: **hybrid- modeling**

- In this example, the melting of the C GGG C GGG CGCGA GGG AT GGG sequence at pH 7.0, 150mM KCl, 20mM phosphate buffer is shown [4].
- **Ellipticity data** were collected from 220 to 320 nm, and from 20 to 90ºC.
- A **hybrid-modeling** approach has been applied. Multivariate analysis allowed the calculation of the **distribution diagram** and of the **pure molar ellipticity spectra**.
- The sequence showed a complex unfolding process, and multivariate analysis based on hybrid modeling was needed to solve all components (see below)



T_m values provide qualitative information on the stability of DNA folded structures.

The hybrid approach not only provides thermodynamic parameters related to the folding/unfolding equilibria, but also allows modeling of baseline drifts.

Baseline drifts at high temperatures are usually due to evaporation phenomena.

4. A step forward: analysis of a process simultaneously monitored with two techniques

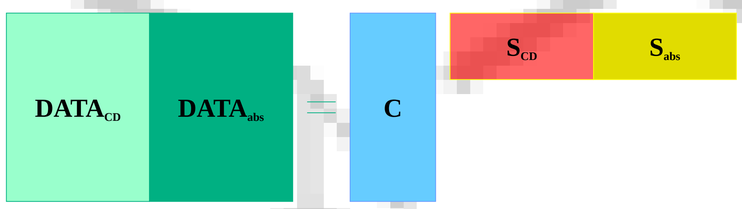
In this case, a single experiment is monitored with two different spectroscopic techniques [5].

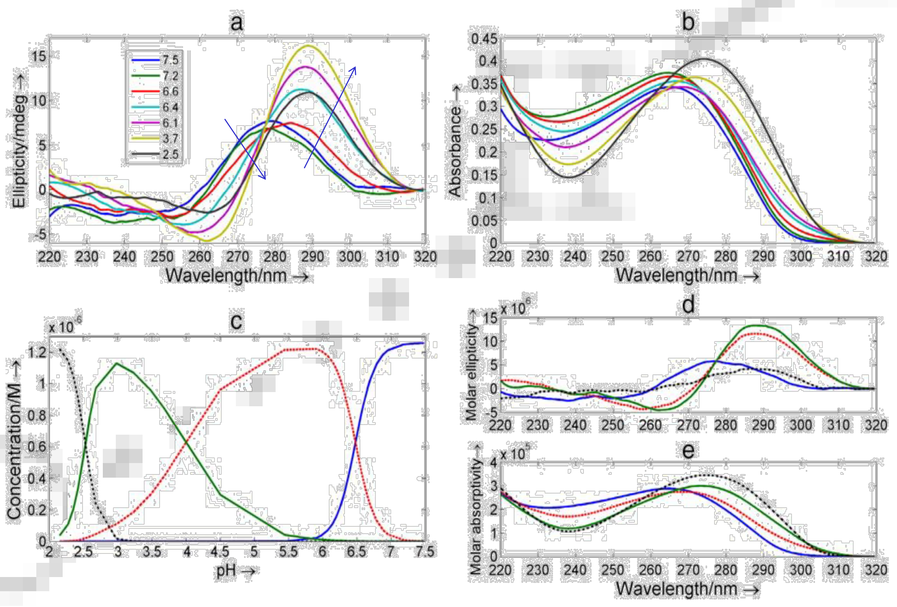
As example, the acid-base titration of the 5'-ACC CCC TGC ATC TGC ATG CCC CCT CCC ACC CCC T-3' sequence has been monitored by means of molecular absorption and circular dichroism spectroscopies.

The sample at pH 7.1 was titrated with successive additions of HCl up to pH 2, approximately at 25ºC and 150 mM KCl.

CD (panel a) and **absorbance (panel b)** spectra were organized in a single, augmented, data matrix.

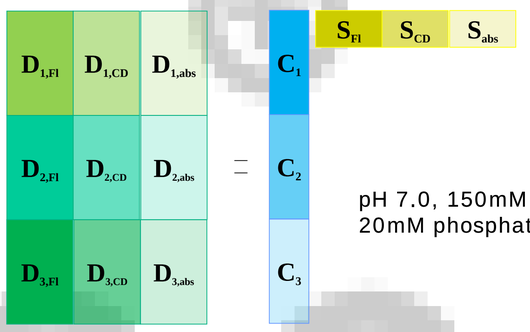
Hard-modeling-based analysis allowed the calculation of the **distribution diagram (panel c)**, and the pure spectra of **molar ellipticity (panel d)** and **absorptivity (panel e)**.



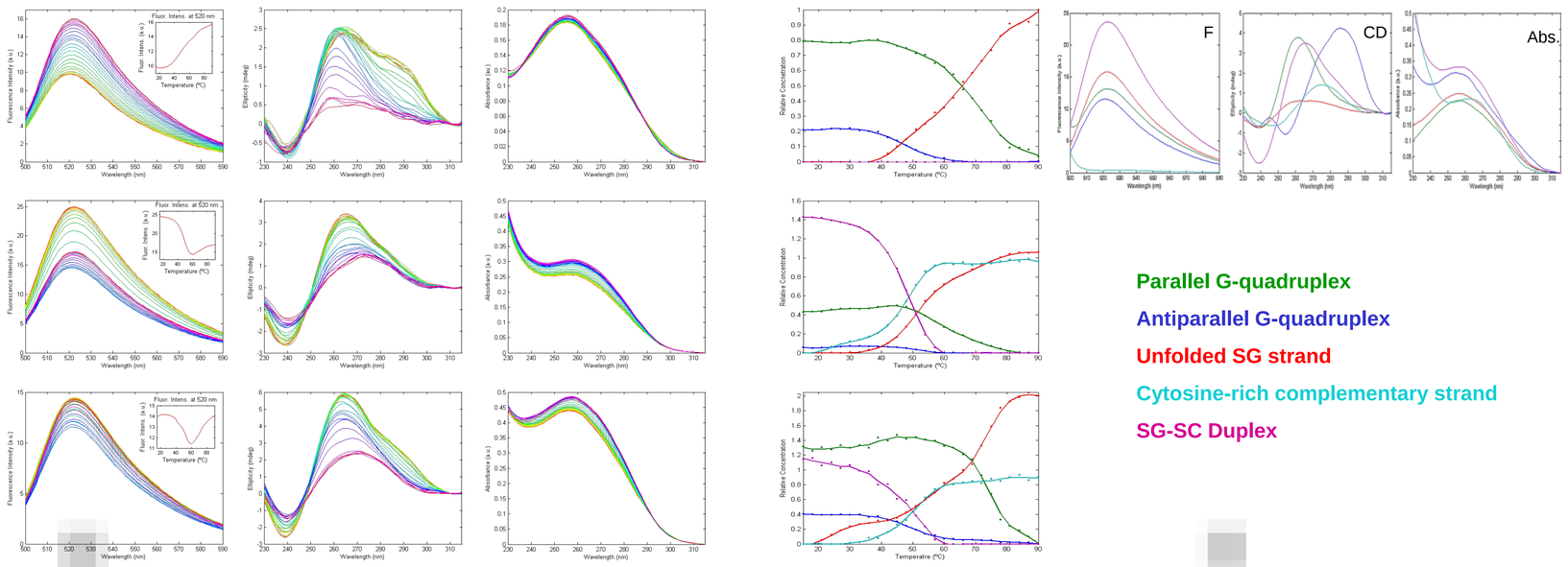


5. Multiple processes monitored with several techniques

- In this case, three melting experiments were monitored with three different spectroscopic techniques (molecular fluorescence, circular dichroism and molecular absorption) [6]. **Soft-modeling**-based analysis was applied.
  1. **Melting of 5'-Fluorescein-TAGGGTTAGGGT-Dabsyl-3' (SG)**
  2. **Melting of the 1:1 mixture of SG and its complementary sequence**
  3. **Melting of the 2:1 mixture of SG and its complementary sequence**



pH 7.0, 150mM KCl,
20mM phosphate buffer



References

[1] Benabou, S.; Aviñó, A.; Eritja, R.; González, C.; Gargallo, R. *RSC Advances* **2014**, 4, 26956-26980
[2] Neidle, S. *The FEBS Journal,* **2010**, 277, 1118-1125
[3] Biffi, G.; Antonio, M.Di; Tannahill, D.; Balasubramanian. S. *Nat. Chem.* **2014**, 6, 75
[4] Gargallo, R. *Anal. Biochem.* **2014**, 466, 4-15
[5] Benabou, S.; Ferreira, R.; Aviñó, A.; González, C.; Lyonnais, S.; Solà, M.; Eritja, R.; Jaumot, J.; Gargallo, R. *Biochim. Biophys. Acta* **2014**, 1840, 41-52
[6] Jaumot, J.; Eritja, R.; Tauler, R.; Gargallo, R. *Nucleic Acids Res.* **2006**, 34, 206-216

Acknowledgments

Funding from the Spanish government and recognition from the Catalan autonomous government are acknowledged.

The poster background shows the CH⁺·C base pair, the building block of the i-motif DNA.