Electronic Supplementary Information

Understanding the effect of nucleobases at the loops on the stability of the i-motif structure

Sanae Benabou[1], Miguel Garavís[2], Sébastien Lyonnais[3], Ramon Eritja[4], Carlos González[2], Raimundo Gargallo[1*]

1. Solution Equilibria and Chemometrics Group, Department of Analytical Chemistry, University of Barcelona, Diagonal 645, E-08028 Barcelona, Spain

2. Institute of Physical Chemistry "Rocasolano", CSIC, Serrano 119, E-28006 Madrid, Spain

3. Institute of Molecular Biology of Barcelona, IBMB-CSIC, Baldiri i Reixac 3-4, 08028 Barcelona, Spain

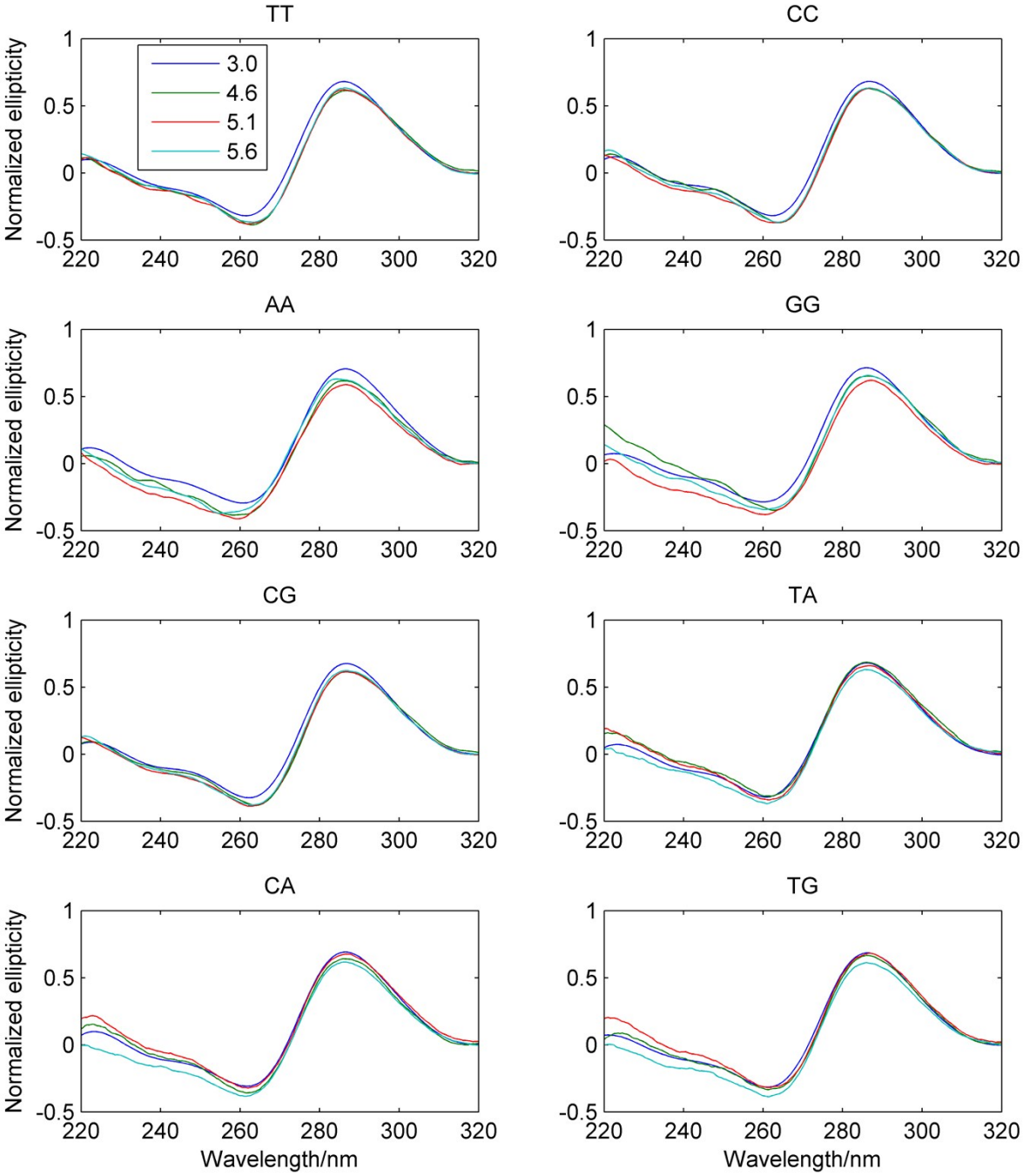
4. Institute for Advanced Chemistry of Catalonia (IQAC-CSIC), CIBER-BBN Networking Centre on Bioengineering, Biomaterials and Nanomedicine, Jordi Girona 18-26, E-08034 Barcelona, Spain

Contents

S1. CD spectra

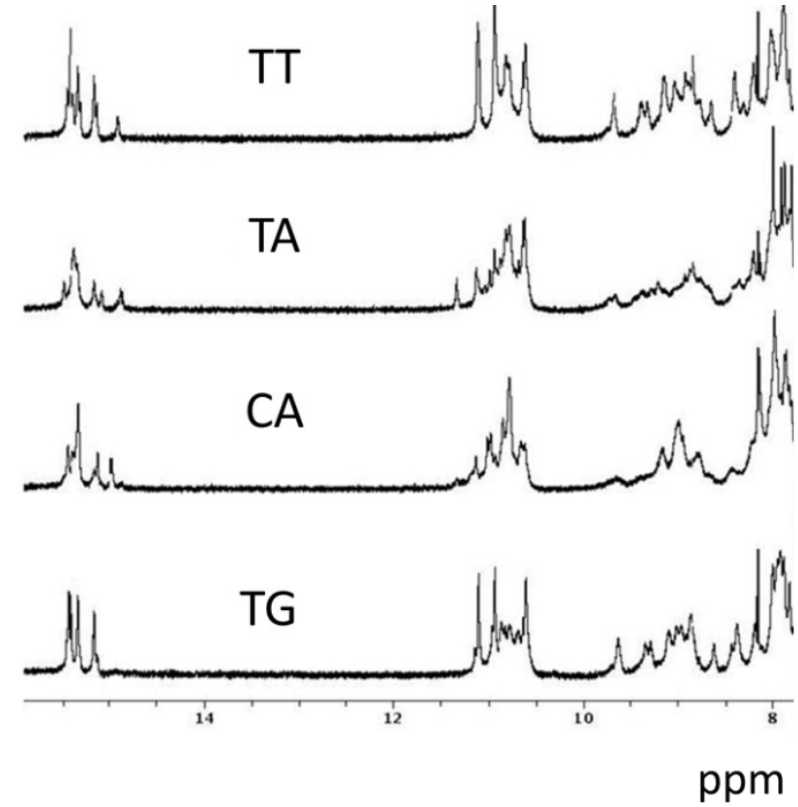
Normalized CD spectra measured at pH values 3.0, 4.6, 5.1, and 5.6 (20mM sodium acetate or phosphate buffer), 150mM KCl and 25°C. The normalization has been carried out to remove artifacts due to the potential differences in DNA concentration among different measurements. The normalization has been done as follows:

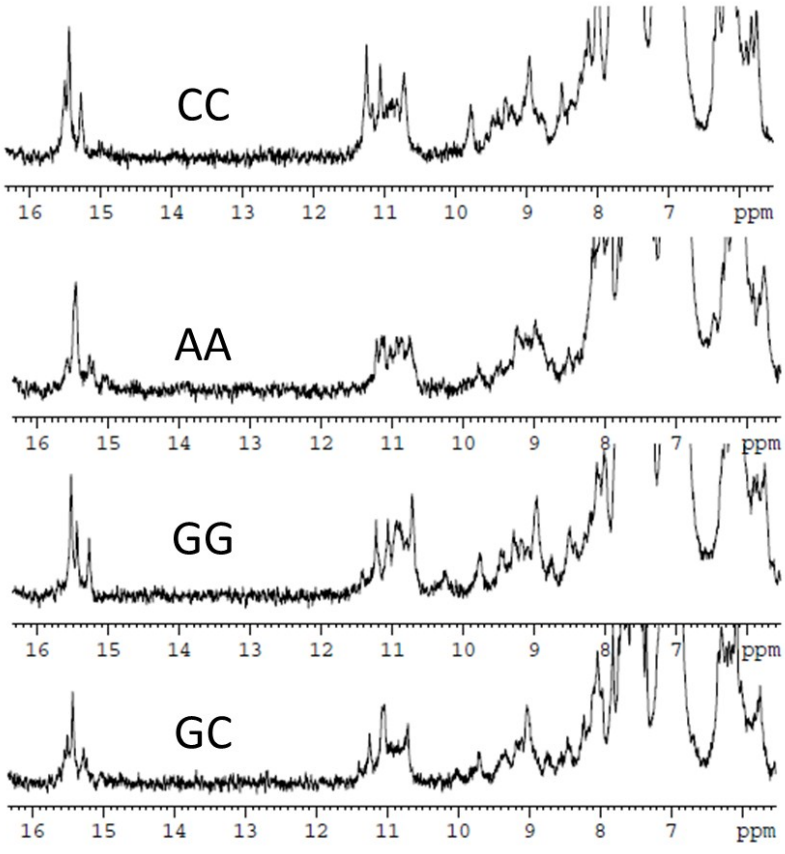
Normalized CD spectrum = experimental CD spectrum /(maximum(ellipticity) – minimum(ellipticity)).



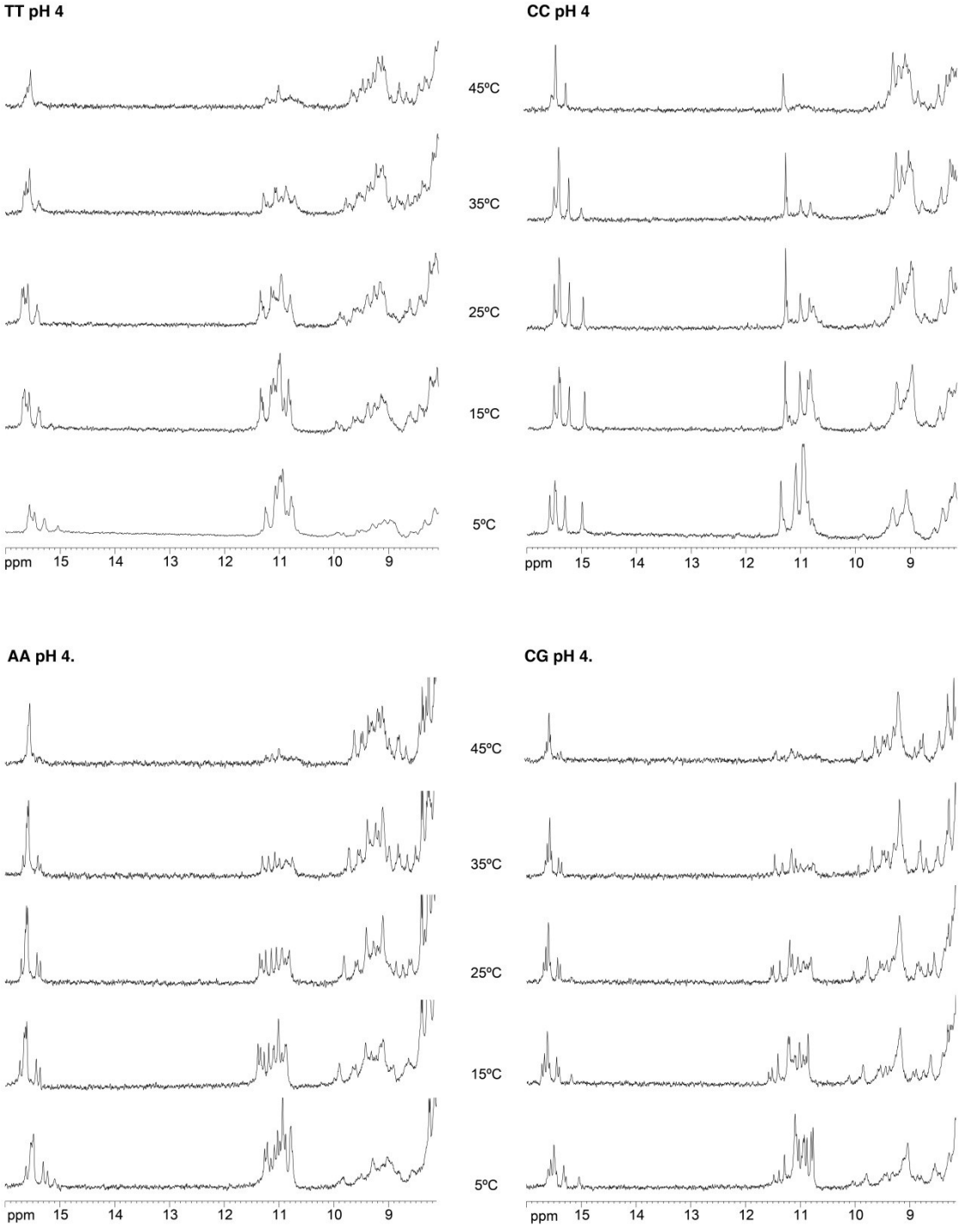
S2. NMR spectra

2.1. NMR spectra measured at pH 5.2, 20mM sodium phosphate buffer, 100mM KCl and 5°C. C_{DNA} = 600 μM.

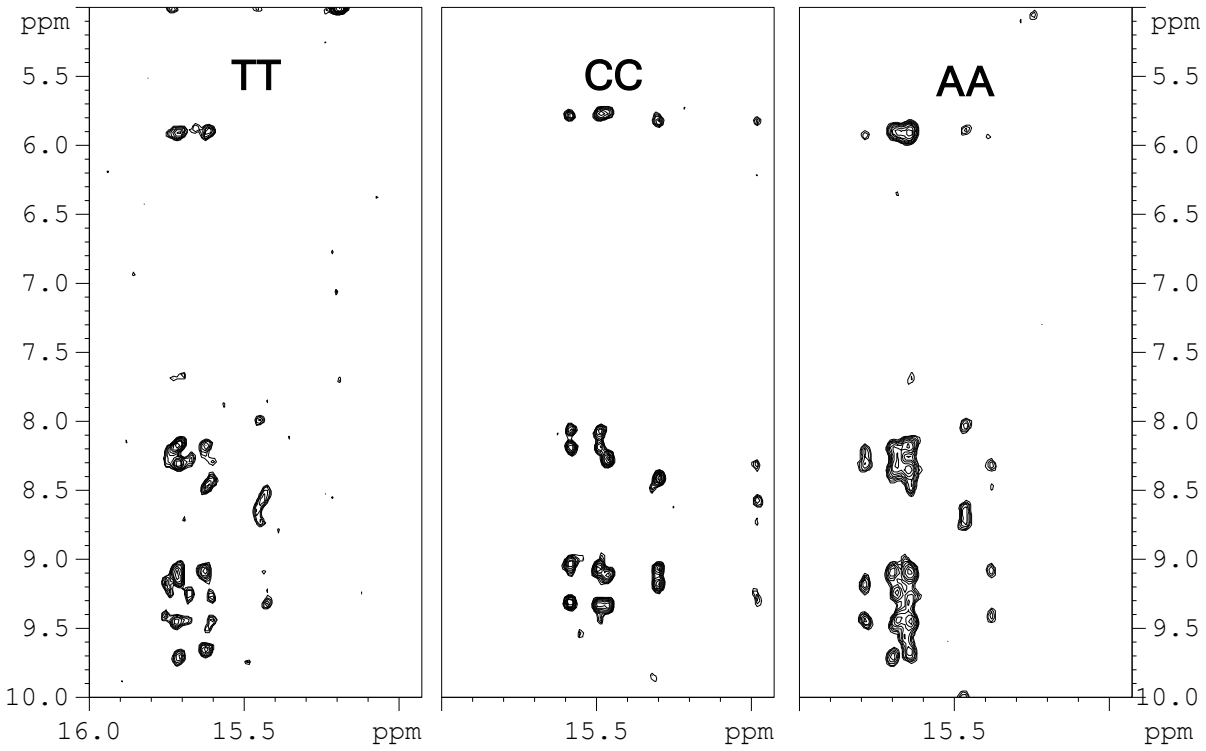




2.2. Exchangeable protons regions of the ^{1}H NMR spectra of TT (a), CC (b), AA (d) and CG (d) at different temperatures. Same buffer as in the previous figure.

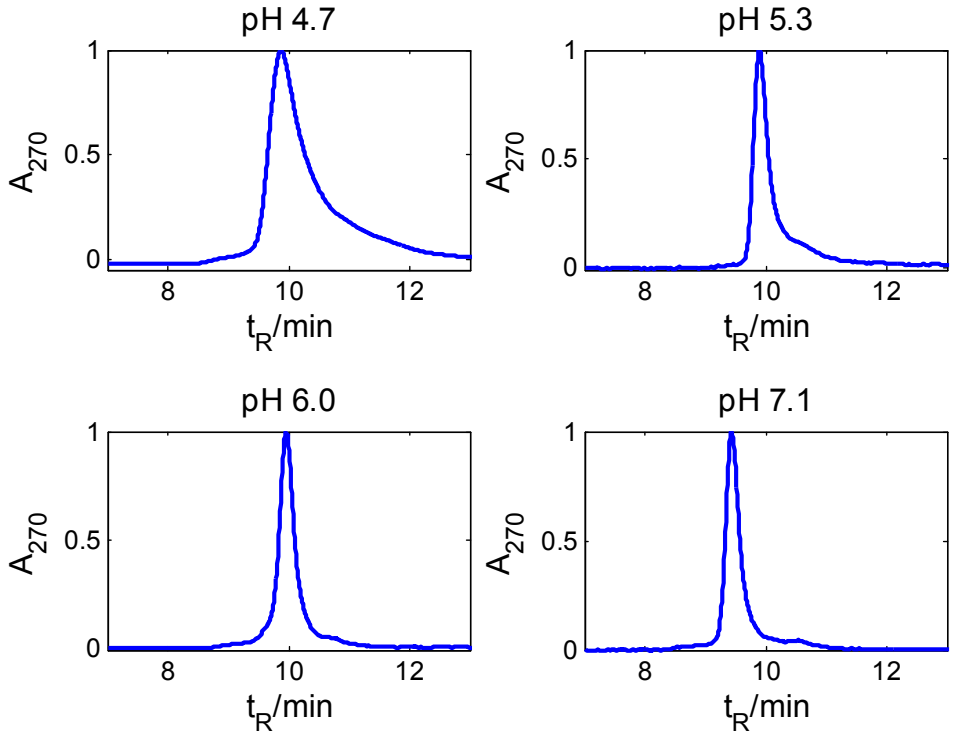


2.3. Cytosine imino/amino region of NOESY spectra of TT, CC and AA at pH 3.5 and T = 5°C, 90:10 H_2O/D_2O (mixing time: 150 ms). Buffer conditions were 20 mM potasium phosphate, 100 mM KCl. Differences in cytosine chemical shifts indicate that the loop sequence affects significantly the overall spectra.



3. SEC analysis in non-denaturing conditions

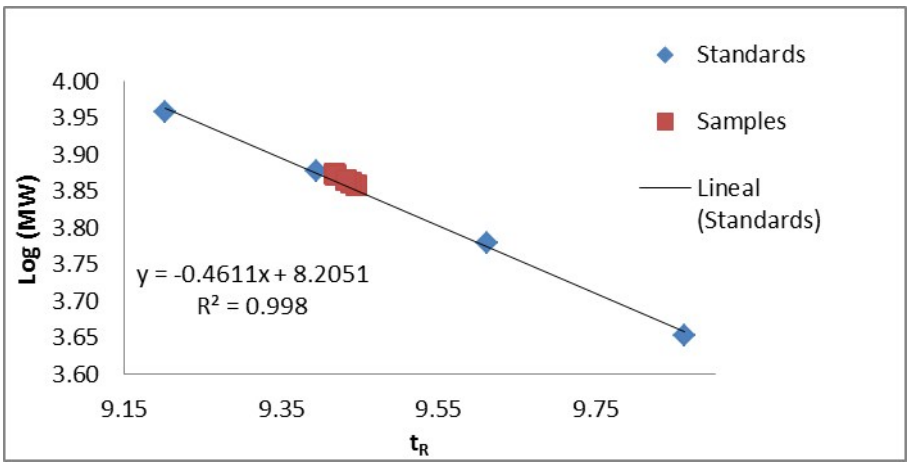
3.1. SEC chromatograms for GC sequence at several pH values



3.2. SEC data and calibration plot at pH 7.1 (phosphate buffer).

Standards	MW (g/mol)	log (MW)	Measured t_R (min)
T15	4501.0	3.6533	9.86
T20	6022.0	3.7797	9.61
T25	7542.9	3.8775	9.39
T30	9063.9	3.9573	9.20

Samples	MW ($g \cdot mol^{-1}$)	log(MW)	Measured t_R	Calculated t_R	Proposed structure
TT	7362.8	3.867	9.44	9.43	Unfolded sequence
CC	7332.8	3.865	9.42	9.43	Unfolded sequence
AA	7380.8	3.868	9.42	9.43	Unfolded sequence
GG	7412.8	3.870	9.44	9.42	Unfolded sequence
GC	7372.8	3.867	9.43	9.43	Unfolded sequence
TA	7372.0	3.867	9.44	9.43	Unfolded sequence
CA	7357.0	3.866	9.44	9.43	Unfolded sequence
TG	7388.0	3.868	9.45	9.44	Unfolded sequence

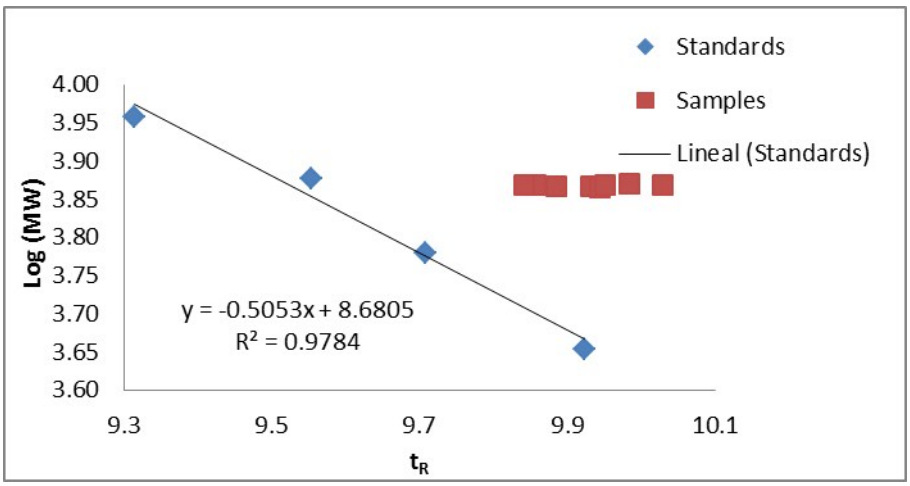


This graph represents the dependence of log(MW) with t_R for a series of four standards (T_{15}, T_{20}, T_{25}, and T_{30}). These standards do not show a folded structure in these experimental conditions. The variation of log(MW) with t_R shows a linear dependence. The samples eluted at t_R values equal (or very similar) to those calculated from the regression line, which means that these samples are not folded in this experimental conditions.

3.3. SEC data and calibration plot at pH 5.3 (acetate buffer).

Standards	MW (g·mol^{-1})	log(MW)	Measured t_R
T15	4501.0	3.6533	9.92
T20	6022.0	3.7797	9.71
T25	7542.9	3.8775	9.55
T30	9063.9	3.9573	9.31

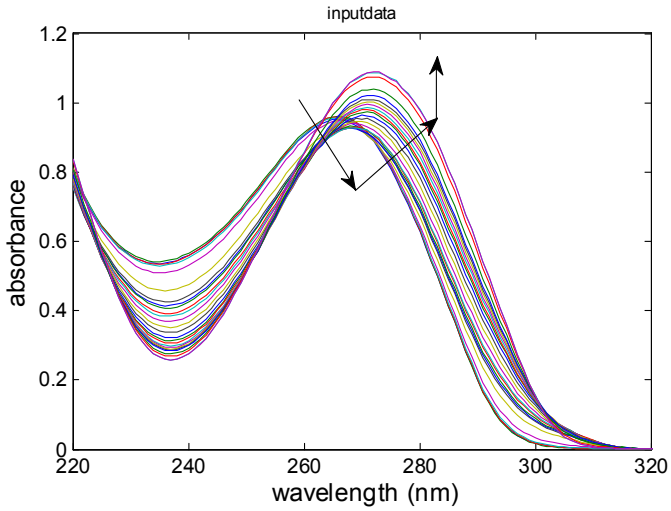
Samples	MW (g·mol^{-1})	log(MW)	Measured t_R	Calculated t_R	Proposed structure
TT	7362.8	3.867	9.93	9.53	*i*-motif
CC	7332.8	3.865	9.94	9.53	*i*-motif
AA	7380.8	3.868	10.03	9.52	*i*-motif
GG	7412.8	3.870	9.98	9.52	*i*-motif
GC	7372.8	3.867	9.96	9.52	*i*-motif
TA	7372.0	3.867	9.86	9.43	*i*-motif
CA	7357.0	3.866	9.89	9.43	*i*-motif
TG	7388.0	3.868	9.84	9.42	*i*-motif



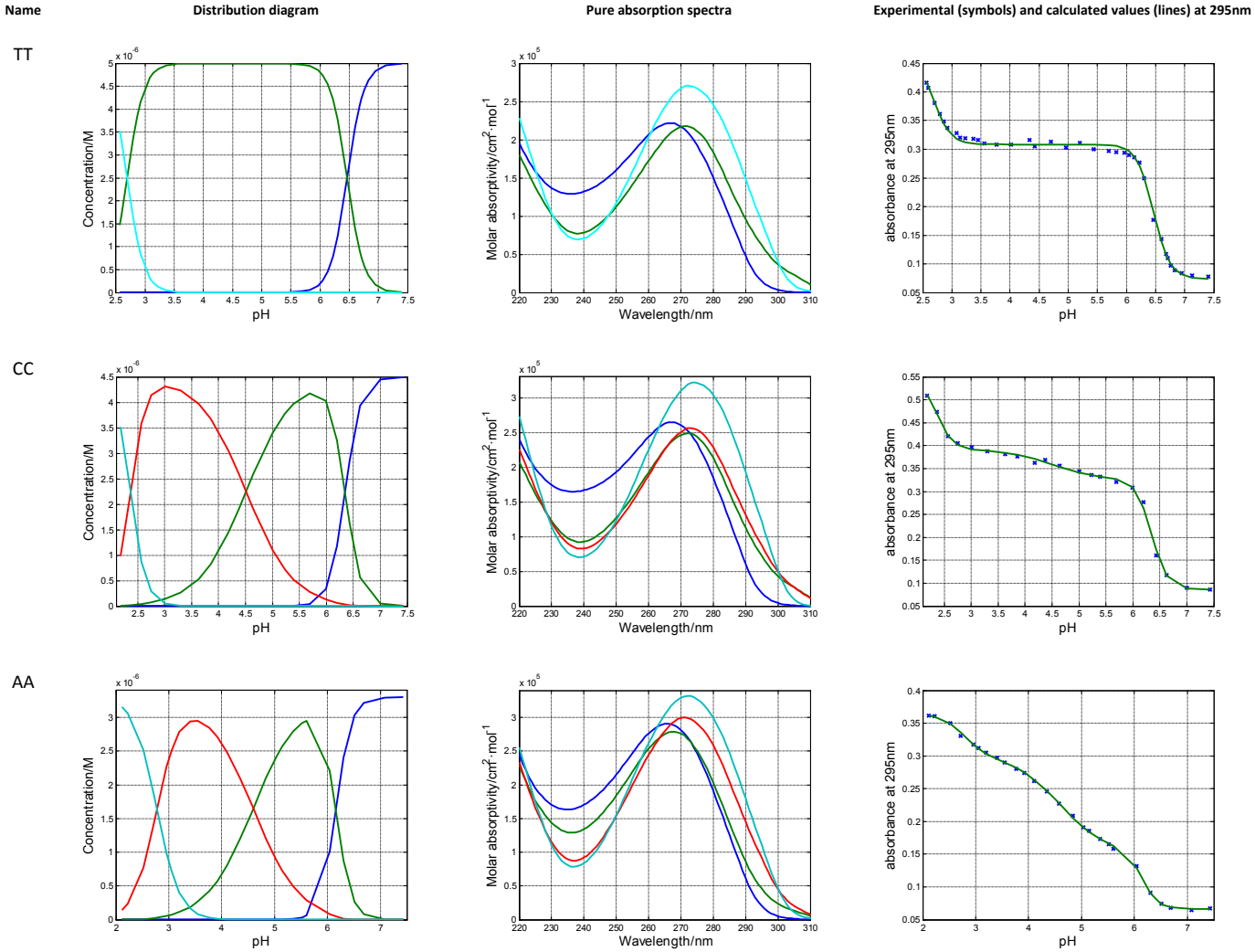
As in the previous graph (pH 7.1), the dependence of log(MW) with t_R for a series of four standards (T_{15}, T_{20}, T_{25}, and T_{30}) is shown. Again, these standards do not show a folded structure in these experimental conditions, and the variation of log(MW) with t_R shows a linear dependence. However, at pH 5.3, the samples eluted at t_R values greater than those expected for unfolded strands, which means that the hydrodynamic volume of samples is smaller than at pH 7.1. This fact is related to the folding into i-motif structures.

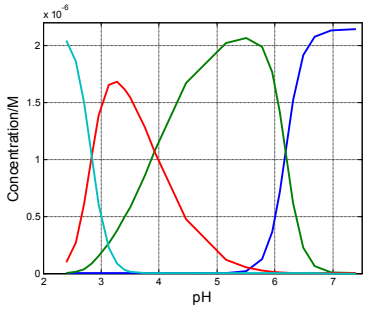
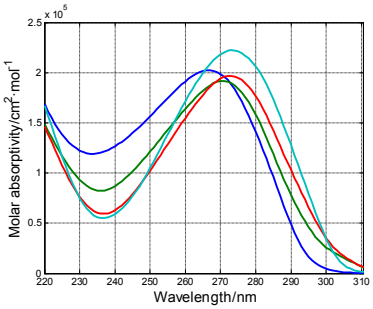
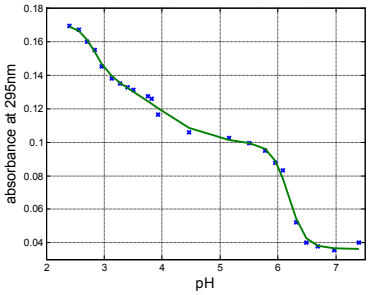
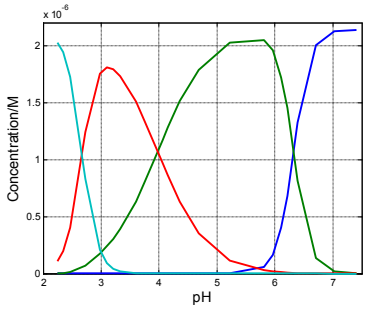
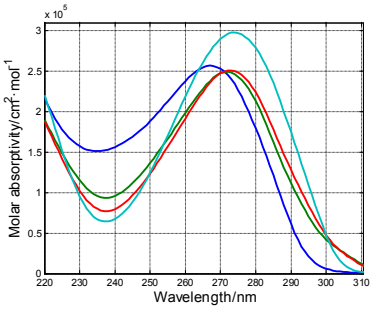
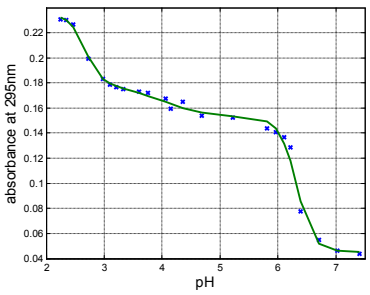
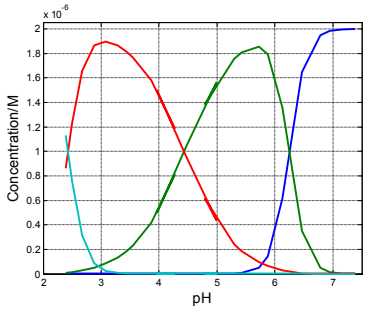
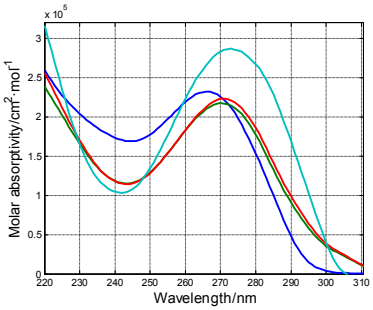
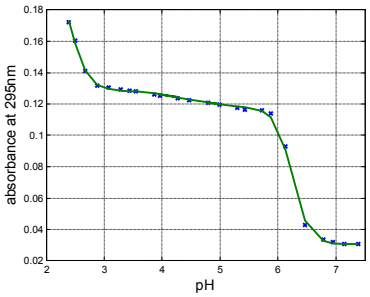
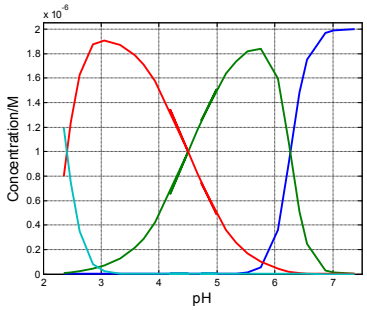
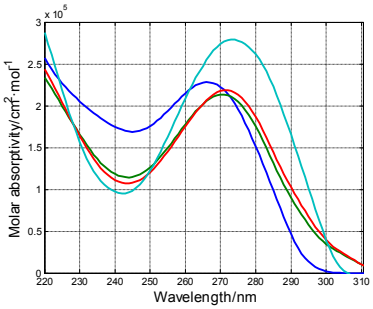
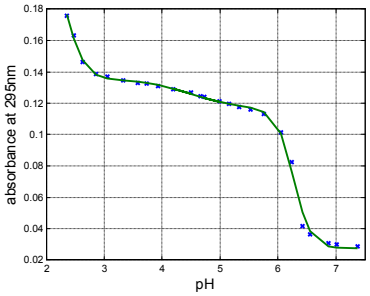
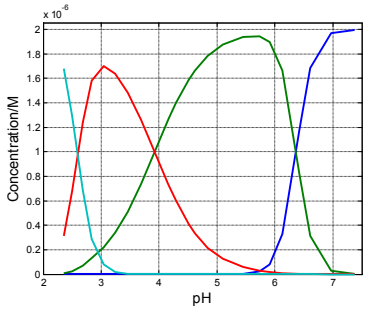
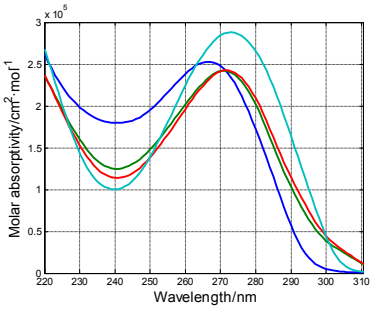
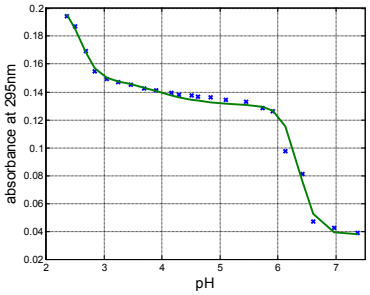
S4. Acid-base titrations

4.1. Experimental spectra recorded throughout the titration of the AA sequence. Arrows indicate the absorbance changes when pH is lowered from pH 7 to 2.5 at 25ºC.

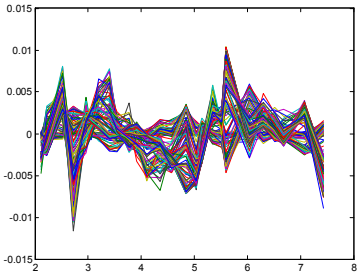
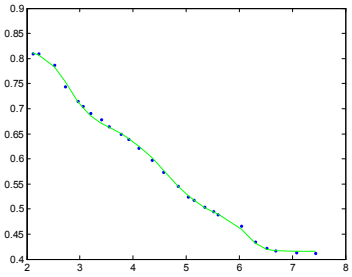
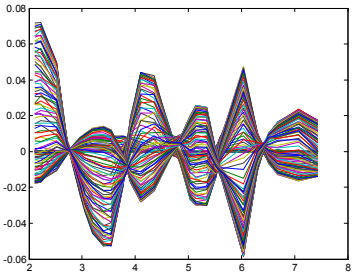
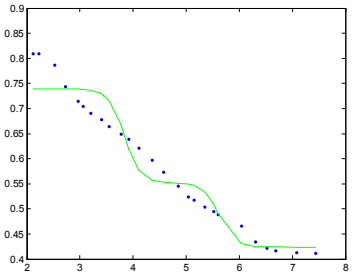
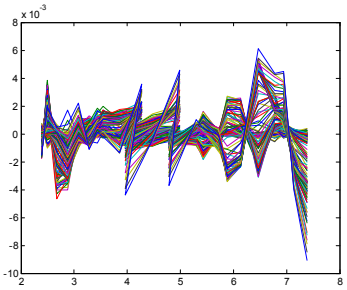
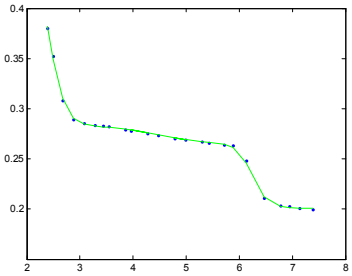
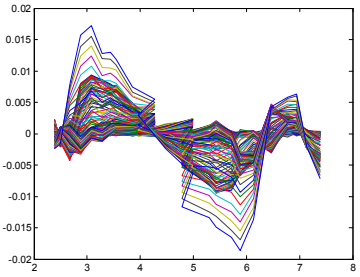
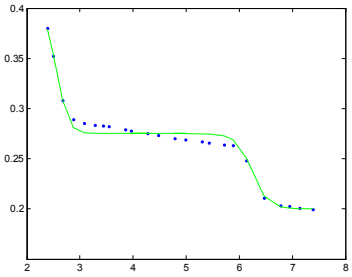


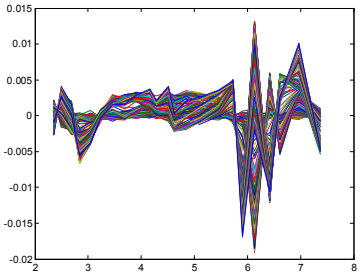
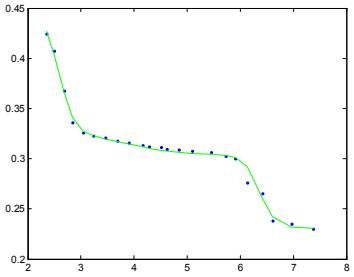
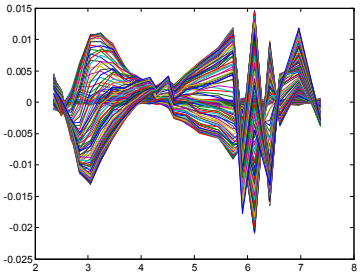
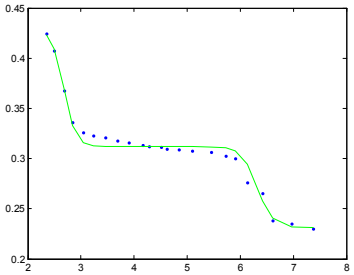
4.2. Calculated distribution diagrams and pure spectra from multivariate analysis of data recorded along acid-base titrations of the considered sequences at 25ºC.



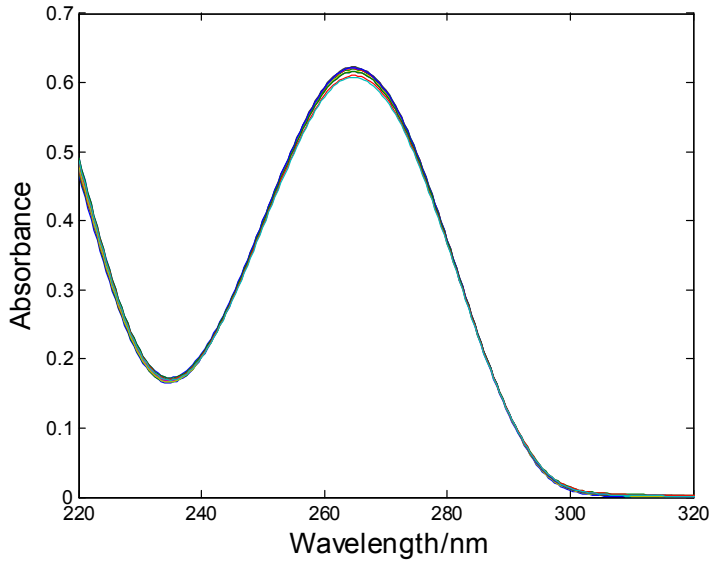
Name	Distribution diagram	Pure absorption spectra	Experimental (symbols) and calculated values (lines) at 295nm
GG			
GC			
TA			
CA			
TG			

4.3. Validation of the existence of the additional transition for several acid-base titrations carried out at 25ºC

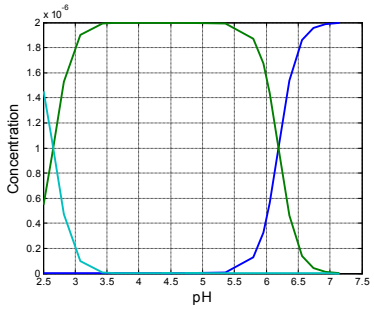
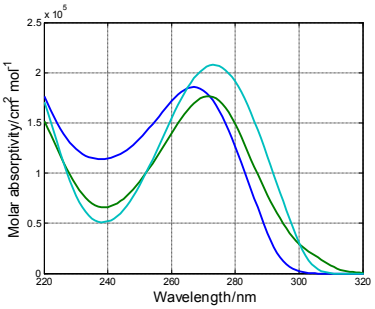
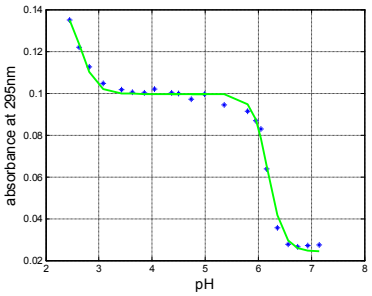
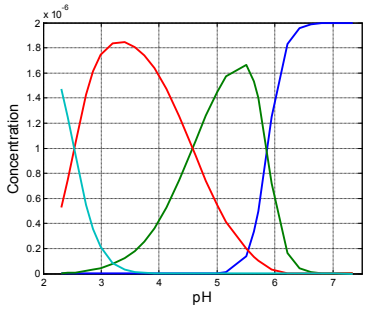
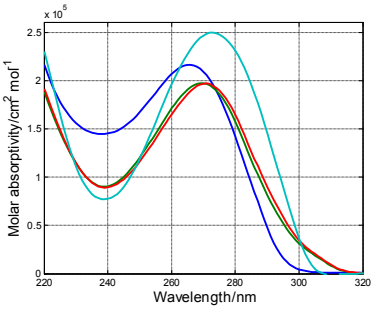
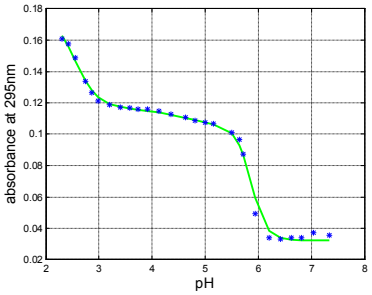
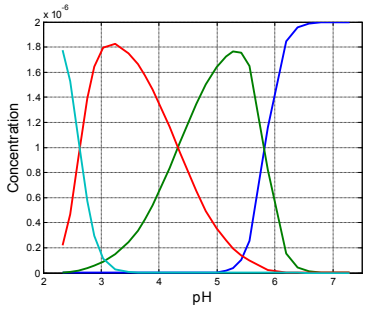
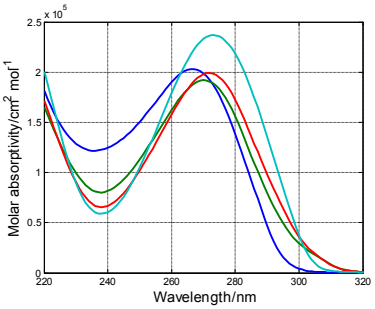
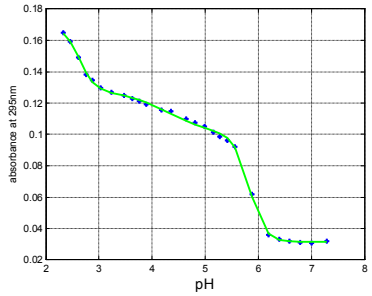
Sequence	Number of components (*nc*) considered in the model	Sum of squares in matrix E	Plot of absorbance data not explained by the proposed model (matrix E)	Fits at 295nm Symbols represent the experimental absorbance values. Lines represent the calculated absorbance values with the proposed model
AA	4	0.020		
	3	0.933		
TA	4	0.006		
	3	0.028		

Sequence	Number of components (*nc*) considered in the model	Sum of squares in matrix E	Plot of absorbance data not explained by the proposed model (matrix E)	Fits at 295nm Symbols represent the experimental absorbance values. Lines represent the calculated absorbance values with the proposed model
TG	4	0.035		
	3	0.063		

4.4. Absorbance spectra of T_{25} sequence measured from pH 7.4 to pH 2.6 at 25°C. The shape of the spectrum remains unaltered, and only the intensity shows a small variation due to dilution.



4.5. Calculated distribution diagrams and pure spectra from multivariate analysis of data recorded along acid-base titrations of the considered sequences at 37°C.

Name	Distribution diagram	Pure absorption spectra	Experimental (symbols) and calculated values (lines) at 295nm
TT			
AA			
CA			

Calculated parameters from the multivariate analysis of spectra recorded along spectrophotometrically-monitored acid-base titrations at 37°C.

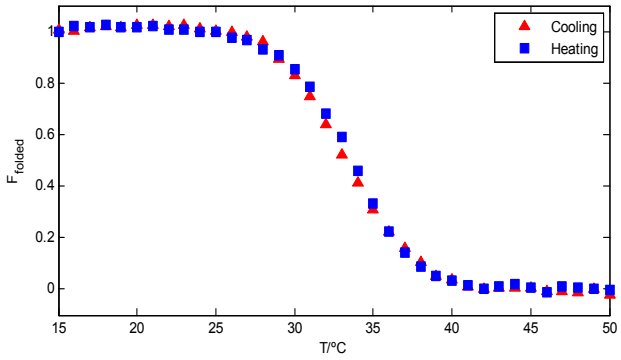
Name	pH-transition midpoint (*p* parameter [a])	pH range predominance of the *i*-motif structure [b]
TT	6.2±0.04 (3) 2.7±0.1 (3)	3.5
AA	5.86±0.04 (3) 4.6±0.2 (1) 2.5±0.3 (3)	3.4
CA	5.83±0.02 (3) 4.3±0.1 (1) 2.6±0.1 (3)	3.2

[a] As commented before, this parameter describes the cooperativity of the transition.
[b] Calculated from the difference between the first and last pH-transition midpoints.

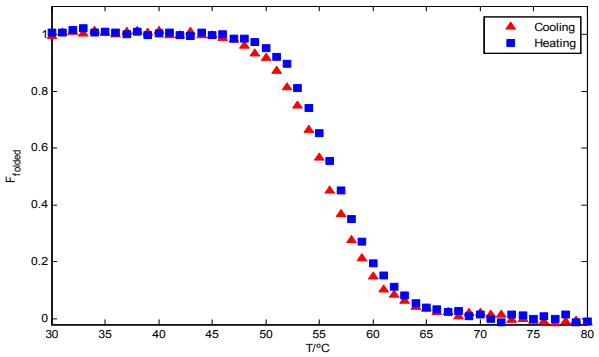
S5. Thermal stability studies

5.1. Normalized heating and cooling curves. Absorbance spectra were recorded at 1°C intervals with a hold time of 3 min at each temperature, which yielded an average heating rate of approximately 0.3°C min^{-1}. Buffer solutions were 20 mM phosphate or acetate, and 150 mM KCl.

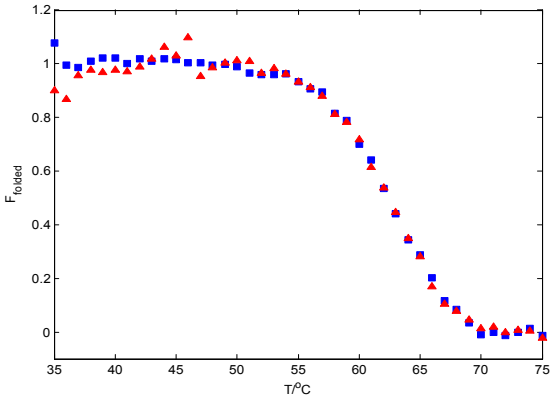
TT sequence at pH 6.2



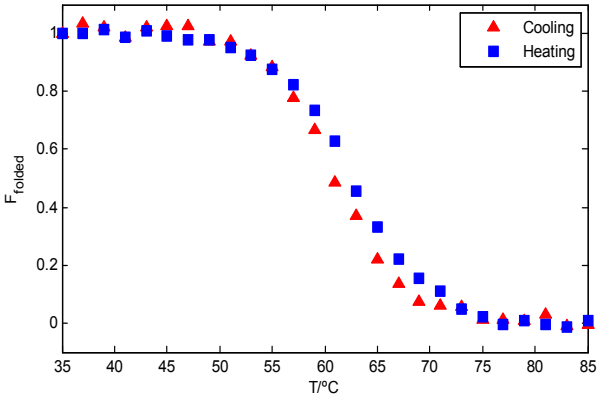
GC sequence at pH 5.1



TT sequence at pH 4.6



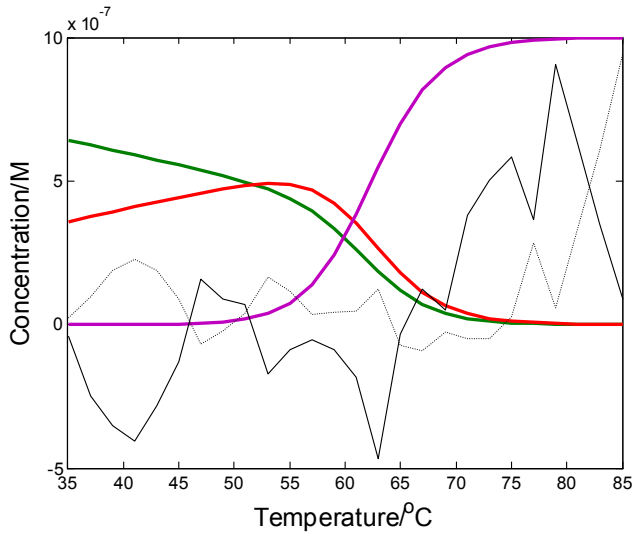
AA sequence at pH 4.6

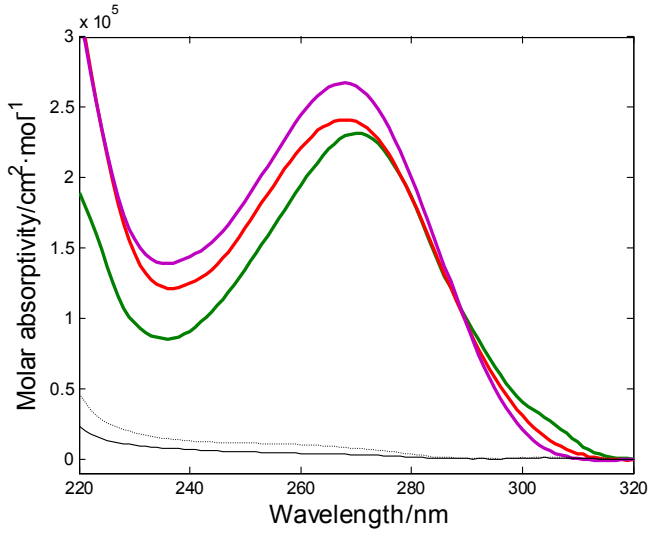


5.2. Multivariate analysis of melting data by means of hybrid modelling.

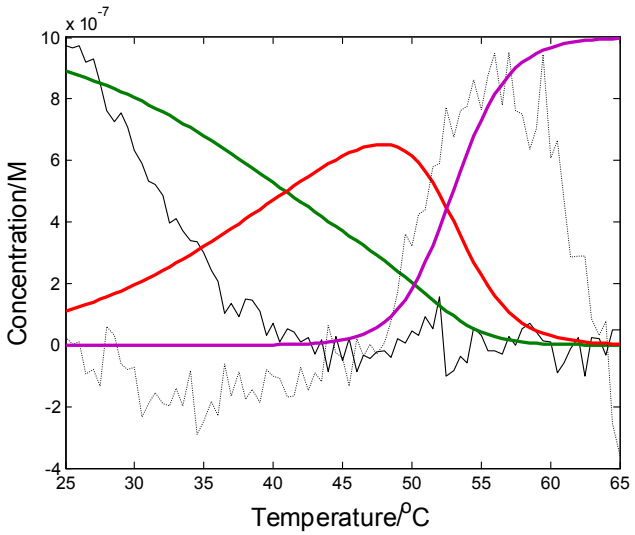
Green and red thick lines correspond to acid-base species in acid-base distribution diagrams. Magenta thick line represents the unfolded strand. Thin black lines correspond to mathematical components not related with the unfolding process, just explaining noise and baseline drifts.

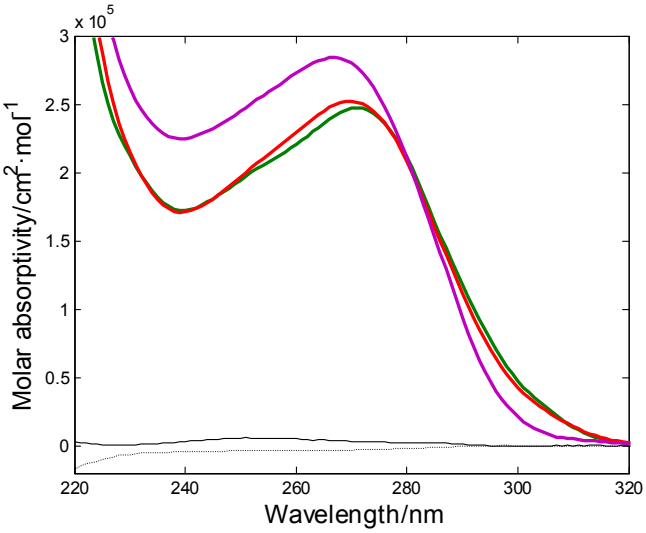
Distribution diagram for the melting of AA sequence at pH 4.6





Distribution diagram for the melting of CA sequence at pH 5.1

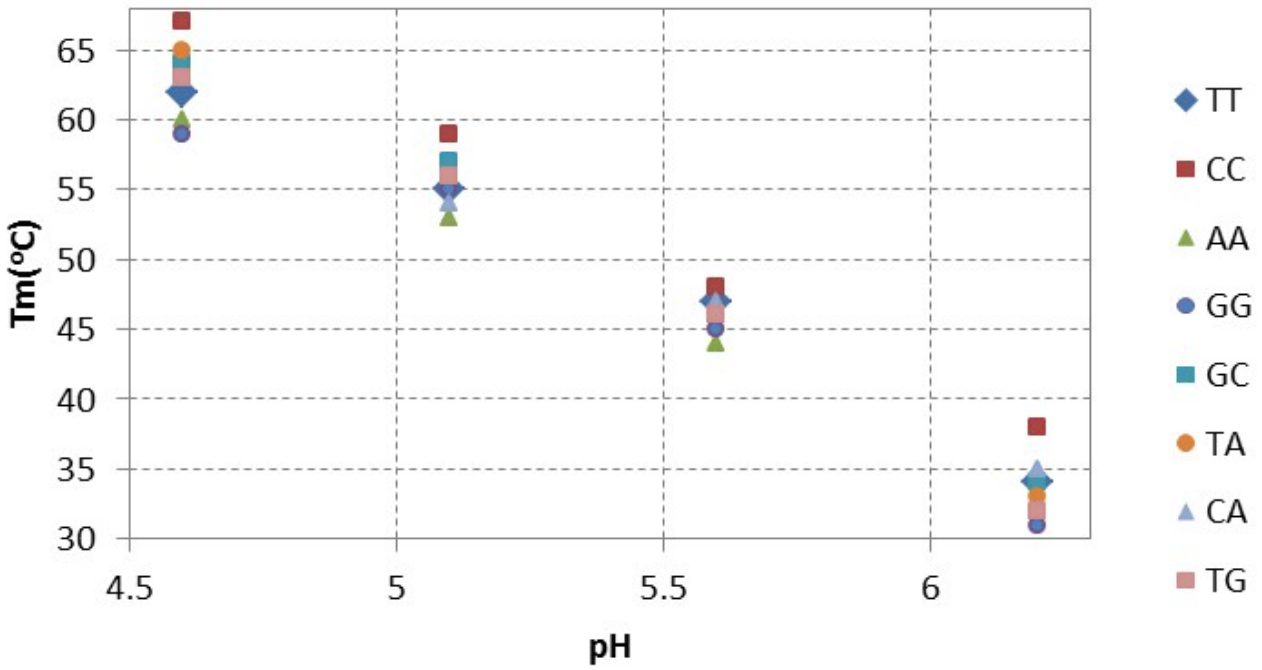




Calculated parameters for the unfolding process:

System	First transition			Second transition		
	ΔH^0 (kcal·mol^{-1})	ΔS^0 (cal·K^{-1}·mol^{-1})	T_m (ºC)	ΔH^0 (kcal·mol^{-1})	ΔS^0 (cal·K^{-1}·mol^{-1})	T_m (ºC)
AA at pH 4.6	6.9±0.1	21±1	52±2	71.6±0.5	215±2	60.7±0.8
CA at pH 5.1	24.5±0.2	78±1	40.9±0.7	100.9±0.6	310±4	52.5±0.7

5.3. Melting temperatures corresponding to Table 1 (main text).



5.4. ΔG^0 dependence upon pH. Thermal stability studies were carried out at 20mM phosphate or acetate buffer and 150mM KCl.