Data in Brief 35 (2021) 106829

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Dataset on pyrene-labelled apolipoprotein A-I, model development and fitting to monitor oligomeric species of its lipid-free form

Wilson A. Tárraga[a], Horacio A. Garda[a,c], Lisandro J. Falomir Lockhart[a,b,1], Marina C. Gonzalez[a,c,1,*]

[a] *Instituto de Investigaciones Bioquímicas de La Plata "Rodolfo Brenner" (INIBIOLP), Centro Científico Tecnológico-La Plata, Calle 60 y 120 (1900), La Plata, Argentina*
[b] *Departamento de Ciencias Biológicas, Facultad de Ciencias Exactas, Universidad Nacional de La Plata, La Plata, Argentina*
[c] *Facultad de Ciencias Médicas, Universidad Nacional de La Plata, La Plata, Argentina*

ARTICLE INFO

Article history:
Received 10 January 2021
Revised 28 January 2021
Accepted 29 January 2021
Available online 13 February 2021

Keywords:
Multi-parametric fluorescent probes
Protein oligomerisation
Fluorescence spectroscopy

ABSTRACT

This article contains data for the self-association of pyrene-labelled single Cys-mutants of apolipoprotein A-I (apoA-I). Mathematical models were developed to characterise the self-association events at different cysteine positions on apoA-I obtained as a function of protein concentration based on the multi-parametric spectrum of pyrene, particularly *P*-value and excimer emissions. The present work complements data related to the article entitled "Analysis of pyrene-labelled apolipoprotein A-I oligomerisation in solution: Spectra deconvolution and changes in *P*-value and excimer formation" Tárraga et al. [1].

DOI of original article: 10.1016/j.abb.2020.108748

* Corresponding author at: Instituto de Investigaciones Bioquímica de La Plata "Rodolfo Brenner" (INIBIOLP), Centro Científico Tecnológico-La Plata, Calle 60 y 120 (1900), La Plata, Argentina.
E-mail address: marinacego@med.unlp.edu.ar (M.C. Gonzalez).

[1] Contributed equally to this work.

https://doi.org/10.1016/j.dib.2021.106829

Specifications Table

Subject	Biochemistry, Genetics and Molecular Biology–Biophysics
Specific subject area	Protein interactions by fluorescence spectroscopy
Type of data	Tables Figures
How data were acquired	Fluorescence spectra were acquired in an SLM Fluorimeter with an Olis S.R.L. update. Absorbance spectra and discrete wavelength measurements were obtained in an Agilent 8453 UV/Vis diode array spectrophotometer. Spectra were manually processed in MS Excel and model fitting were performed with Solver Add-in. Models for analysis were developed in Wolfram Mathematica 11.3.
Data format	Raw and Analysed
Parameters for data collection	Recombinant proteins were purified and labelled *in vitro* with pyrenyl-N maleimide. Efficiency of labelling and structural controls are reported. Fluorescence spectra were recorded with a 1 nm interval resolution, employing 4/4 nm slit-widths, and a variable gain to adjust sensibility for samples that ranged from 0.25 to 100 µM pyrene-labelled protein.
description of data collection	Absorbance spectra were registered at 25 °C in quartz cuvettes between 190 and 900 nm. Fluorescence spectra were collected at 25 °C in quartz cuvettes between 350 and 650 nm after thermal equilibrium was reached preparing samples by serial dilution from a concentrated stock. Analysed spectra were background and Raman subtracted and normalized by integrated total intensity in order to associate relative changes in peaks´intensity to oligomerisation state changes. Excimer and *p*-value signal were calculated based on spectral windows consisting of 1 nm centred at pi (376 nm), piii (384 nm), piv (395 nm) and pe (460 nm).
Data source location	Instituto de Investigaciones Bioquímicas de La Plata "Rodolfo Brenner" (INIBIOLP). Facultad de C. Médicas, Calle 60 y 120 s/n (1900) La Plata, Prov. Buenos Aires, Argentina.
Data accessibility	The raw data are available with this article and in the Mendeley Data, link: http://dx.doi.org/10.17632/v8yz654vn8.1
Related research article	Wilson A. Tárraga, Lisandro J. Falomir-Lockhart, Horacio A. Garda and Marina C. González. Analysis of pyrene-labelled apolipoprotein A-I oligomerisation in solution: Spectra deconvolution and changes in *P*-value and excimer formation. Tárraga W. 699 (2021). http://dx.doi.org/10.1016/j.abb.2020.108748.

Value of the Data

- The data presented in this article would be important to a better understanding the self-association of apoA-I, a complex system where more than one equilibrium was reported.
- The simple experimental design, the mathematical models developed based on properties of the fluorescence probe and its combination with spectra deconvolution of the data acquired could serve as an example for the application of similar methods to the study of other complex systems of protein interactions.
- These data provided could be helpful for other researchers to the development of more complex mathematical models in order to gain more information from the self-association of apoA-I.

1. Data Description

1.1. Purification and labelling of single apoA-I cysteine variants

The wild-type (WT) and seven human apoA-I Cysteine (Cys) variants were produced recombinantly, purified and labelled with N-(1-Pyrenyl) maleimide (NPM) according to the method described below. Here, we show a representative sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) in which there is a batch of NPM-labelled apoA-I variants (Fig. 1) included WT apoA-I.

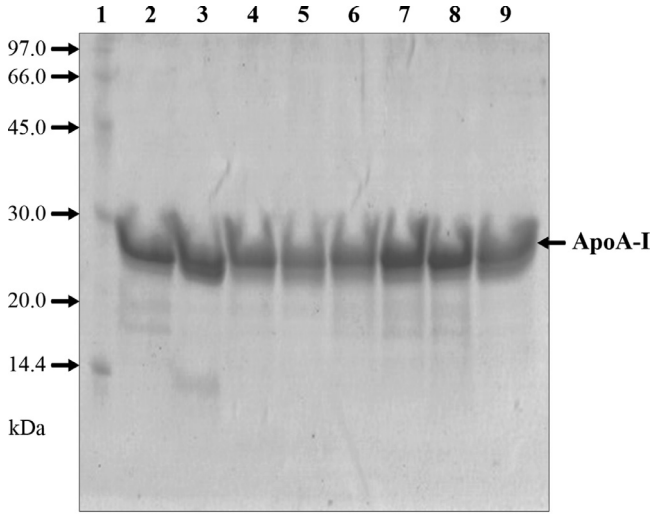

Fig. 1. SDS-PAGE analysis of purified recombinant apoA-I variants labelled with pyrene. Proteins were separated on 16% acrylamide gel and stained with Commassie Blue. Lane 1, MW standards; lane 2, WT apoA-I control; lane 3, K107C; lane 4, K133C; lane 5, K226C; lane 6, F104C; lane 7, L137C; lane 8, F225C; and lane 9, +244C. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Thermodynamic parameters of apoA-I Cys-variants labelled with pyrene.

Mutant (N)	ΔG_{H20}(Kcal*mol^{-1})	m(Kcal*mol^{-1}*M$_{GdHCl}$$^{-1}$)
WT (3)	3.1 ± 0.2	3.5 ± 0.3
WT-Pyr (3)	3.28 ± 0.03	3.67 ± 0.02
F104C-Pyr (3)	1.74 ± 0.02	2.52 ± 0.05
K107C-Pyr (3)	2.60 ± 0.01	3.16 ± 0.08
K133C-Pyr (3)	3.25 ± 0.03	3.70 ± 0.08
L137C-Pyr (3)	2.57 ± 0.01	3.0 ± 0.1
F225C-Pyr (3)	2.77 ± 0.07	3.32 ± 0.02
K226C-Pyr (3)	2.48 ± 0.06	2.52 ± 0.06

The free energy of unfolding $\Delta G_{H20}°$ and cooperativity parameter m were determined by fitting a model based on a N-U equilibrium, where GdHCl-induced unfolding was monitored by Trp-fluorescence. Data are presented as average ± standard error of the mean. N represents the number of independent batches of labelled protein (replicates) averaged for each mutant.

1.2. Structural analyses of apoA-I Cys-variants

We have studied and compared WT apoA-I stability and folding with apoA-I Cys-variants, and we analysed the proteins intrinsic fluorescence and denaturation curves results in the next section.

1.2.1. Tryptophan intrinsic fluorescence

ApoA-I has four tryptophan (Trp) residues in positions 8, 50, 72, and 108. The apoA-I intrinsic fluorescence was employed to analyse the effect of apoA-I Cys-NMP variants on the overall structure and the stability in solution, and to evaluate the sensitive to environment changes induced by mutations and labelling. Changes in the centre of mass (CM) of Trp-fluorescence emission through the Guanidinium hydrochloride (GdHCl) denaturing curves are shown for all variants and WT apoA-I (Fig. 2). The free energy of unfolding and the cooperativity parameter have been reported (Table 1). Data from GdHCl denaturing curves (Fig. 1) and thermodynamic

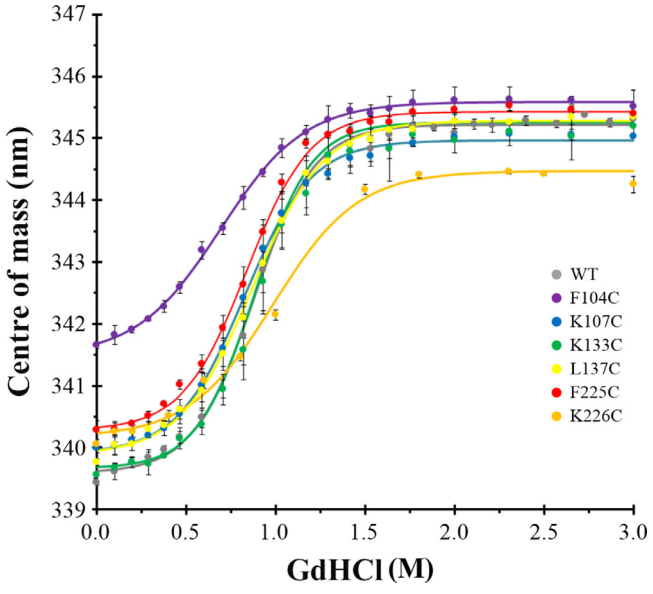

Fig. 2. GdHCl–induced unfolding of apoA-I variants labelled with pyrene was monitored by Trp-fluorescence emission (λ_{exc} 295 nm). Protein concentration was 3.5 µM in buffer PBS, pH 7.4. Unlabelled WT apoA-I (grey), F104C (purple), K107C (blue), K133C (green), L137C (yellow), F225C (red), and K226C (orange). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

parameters (Table 1) for all labelled apoA-I variant are available in Figure_2.pdf, Figure_2.xlsx and Table_1.xlsx files in the online repository.

1.2.2. Reconstitution of high density lipoproteins (rHDL)

The reconstitution of rHDL particles is a method used for structural and biological characterisation of apoA-I variants. The reconstituted rHDL with apoA-I were prepared by the cholate dialysis method, using 1-Palmitoyl-2-oleoylphosphatidylcholine (POPC), apoA-I, and sodium cholate in molar ratios of 35:1:80 and 95:1:150, which generate particles of 7.8 nm and 9.6 nm diameters, respectively. Initially, the desired amount of lipids was mixed in $CHCl_3$, dried under N_2, and dispersed in buffer TBS (10 mM Tris-HCl, 150 mM NaCl, pH 7.4) during 1 h at 37 °C. Then, sodium cholate was added and the solution was kept at 4 °C until the lipid dispersion cleared due to formation of micelles. Proteins at 5 µM were incubated 30 min with the micelles and then dialysed extensively against buffer TBS at 4 °C to eliminate the sodium cholate. Finally, the size of the rHDL was checked by non-denaturing gradient (8–25%) PAGE (Fig. 3). Picture is available in the Figure_3.pdf file in the online repository.

1.3. Fluorescence emission spectra of lipid-free apoA-I mutants labelled with pyrene

Fluorescence emission spectra of apoA-I labelled with pyrene were measured to evaluate pyrene spectral changes in water solution. Proteins were prepared in buffer 25 mM $NaHPO_4$ and 150 mM NaCl (pH 7.4) at 25 °C. Fluorescence emission spectra were acquired in an SLM4800 spectrofluorometer (Applied Photophysics) updated by Olis (ISS Inc., Champaign, IL) exciting at 345 nm and collecting emission from 350 to 650 nm with a 1 nm interval resolution and 4 nm slits width. The apoA-I concentration was calculated by BCA (Bicinchoninic Acid) Protein Assay (Thermo Scientific). Absorbance spectra were acquired from 190 to 900 nm in an 8453 UV-Visible diode-array spectrophotometer (Agilent) with a 1 nm interval resolution. After subtraction of

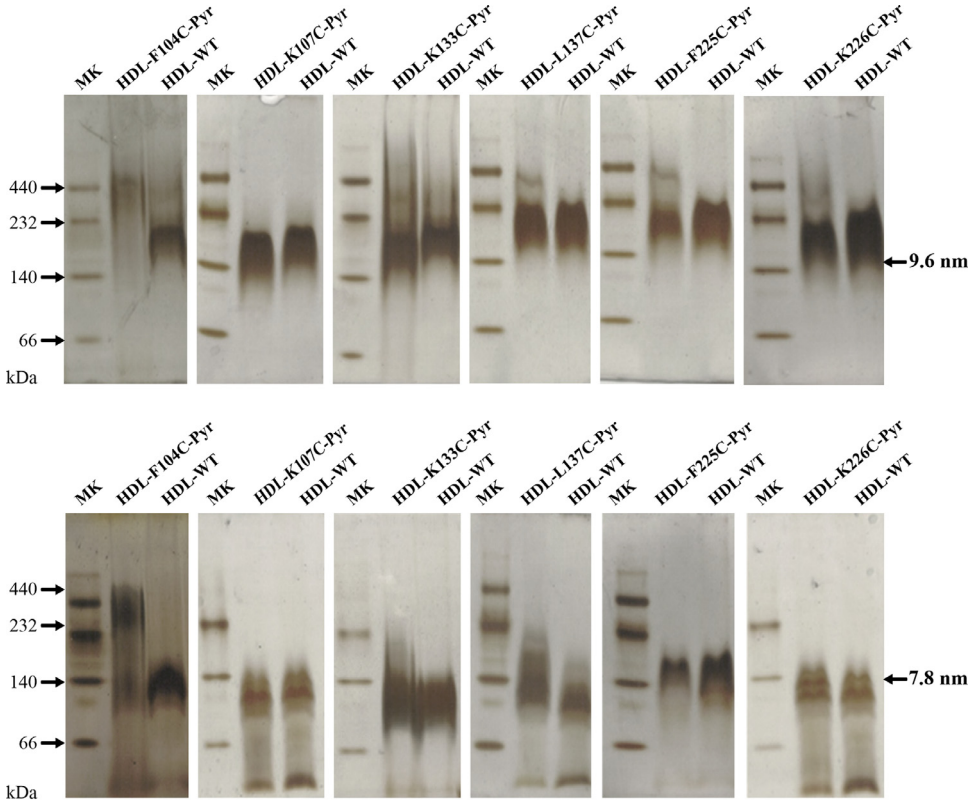

Fig. 3. Non-denaturing gel electrophoresis analysis of rHDL particles. The rHDL were prepared following the standard method of cholate dialysis of a mixture of POPC/apoA-I/Cholate (35:1:80) and (95:1:150) with labelled apoA-I variantes. Proteins were separated on 8 to 25% acrylamide gradient gels and stained with silver. For each mutant, lane 1 contains MW standards; lane 2, rHDL-variant; and lane 3, WT apoA-I rHDL control.

Raman signal, spectra were corrected by inner-inter effect (IFE) which affects the fluorescence intensity from 350 to 375 nm (overlap with absorbance spectrum). An example of this is shown for K133C-Pyr (Fig. 4). The corrected spectra for all mutants are shown in Fig. 5. Data from IFE correction (Fig. 4) and fluorescence emission (Fig. 5) for all labelled apoA-I variants are available in Figure_4.xlsx and Figure_5.xlsx files in the online repository.

1.4. Corrected and deconvolution spectra

The pyrene probe fluorescent properties are useful to obtain information regarding protein structure and its dynamics. Spatial proximity between pyrenes is detected by excited dimer (excimer) formation, evidenced by appearance of a broad emission band centred at 460 nm. Moreover, pyrenes microenvironment polarity changes can be monitored by the ratio between 375 and 385 nm intensities (pI and pIII) from its fine structure emission [2]. However, the exact multi-component spectra information for quantitative analysis is still a challenge, because the fine structure emission depends on aqueous or organic environment and experimental conditions [3]. In particular, five emission signals from 370 to 410 nm, designed as peaks/bands I, II, III, IV, and V, are well distinguishable in apolar and aprotic solvents; while other signals can be

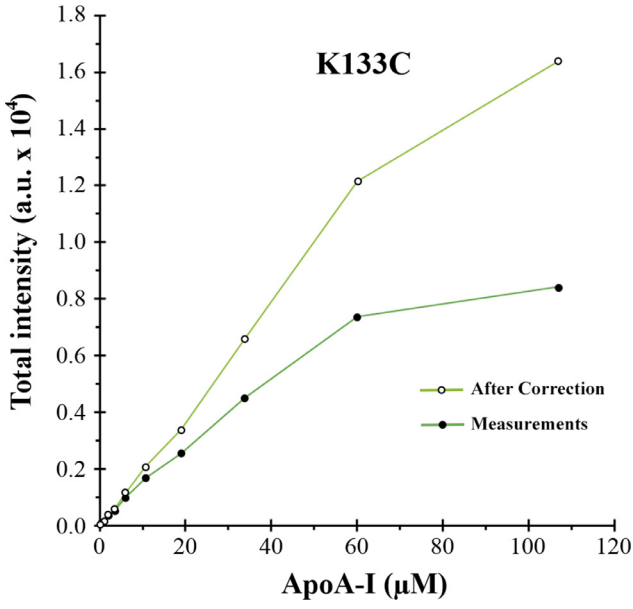

Fig. 4. Inner filter correction. Typical data after correction by inner filter effect. Example of this correction is showed for K133C labelled with pyrene.

observed at longer wavelengths (415–425 nm), designated in this study with decimal numbers (i.e. band 6). We propose a spectral deconvolution method for fluorescence emission spectra to resolve excimer formation and polarity changes (*P*-value). The method is a spectral decomposition in which is considered the individual sum of its components and this is described mathematically by a sum of different types of distributions.

There are two spectral distributions commonly used to describe the fluorescence emission signals, the *Gaussian* and the *log-type* distributions [4]. The *Gaussian* distribution is defined by 3 parameters: μ (mean), σ (variance) and Amp (amplitude) [5].

$$\text{Gaussian distribution}: \; I(\lambda) = \frac{\text{Amp}}{\sqrt[2]{2\pi}} e^{-\left(\frac{(\lambda-\mu)^2}{2\sigma^2}\right)} \tag{1}$$

The log distribution has been proposed as the function that appropriately describes the fluorescence emission spectra shape in complex molecules [6]. This is defined by 4 parameters: μ (mean), α (asymmetry), Amp (amplitude) and λ_0 (starting point). The log distribution is asymmetric, and the fluorescence signal is distributed to one side or the other of the maximum fluorescence emission.

$$\text{Log distribution}: \mathrm{I}(\lambda) = \begin{cases} \text{Amp} * e^{\left(\frac{\text{Ln}(2)}{(\text{Ln}(\alpha))^2} * \left(\text{Ln}\left(\frac{\frac{1}{\lambda_0}-\frac{1}{\lambda}}{\frac{1}{\lambda_0}-\frac{1}{\mu}}\right)\right)^2\right)}, & \lambda > \lambda_0 \\ 0, & \lambda \le \lambda_0 \end{cases} \tag{2}$$

In this work, because some bands are not distinguishable in aqueous buffers, we have employed combinations of 5 or 4 distributions, with or without excimer emission, to describe pyrene fluorescence spectra. The bands were named according to their maxima location as peaks I, III and IV, band 6 and band E. All spectra were fitted using the non-linear fitting tool of Wolfram Mathematica 11.3, software package (https://www.wolfram.com/mathematica/). Example of this deconvolution is shown in an related article (Fig. 3, Tárraga et al. [1]). The deconvolution

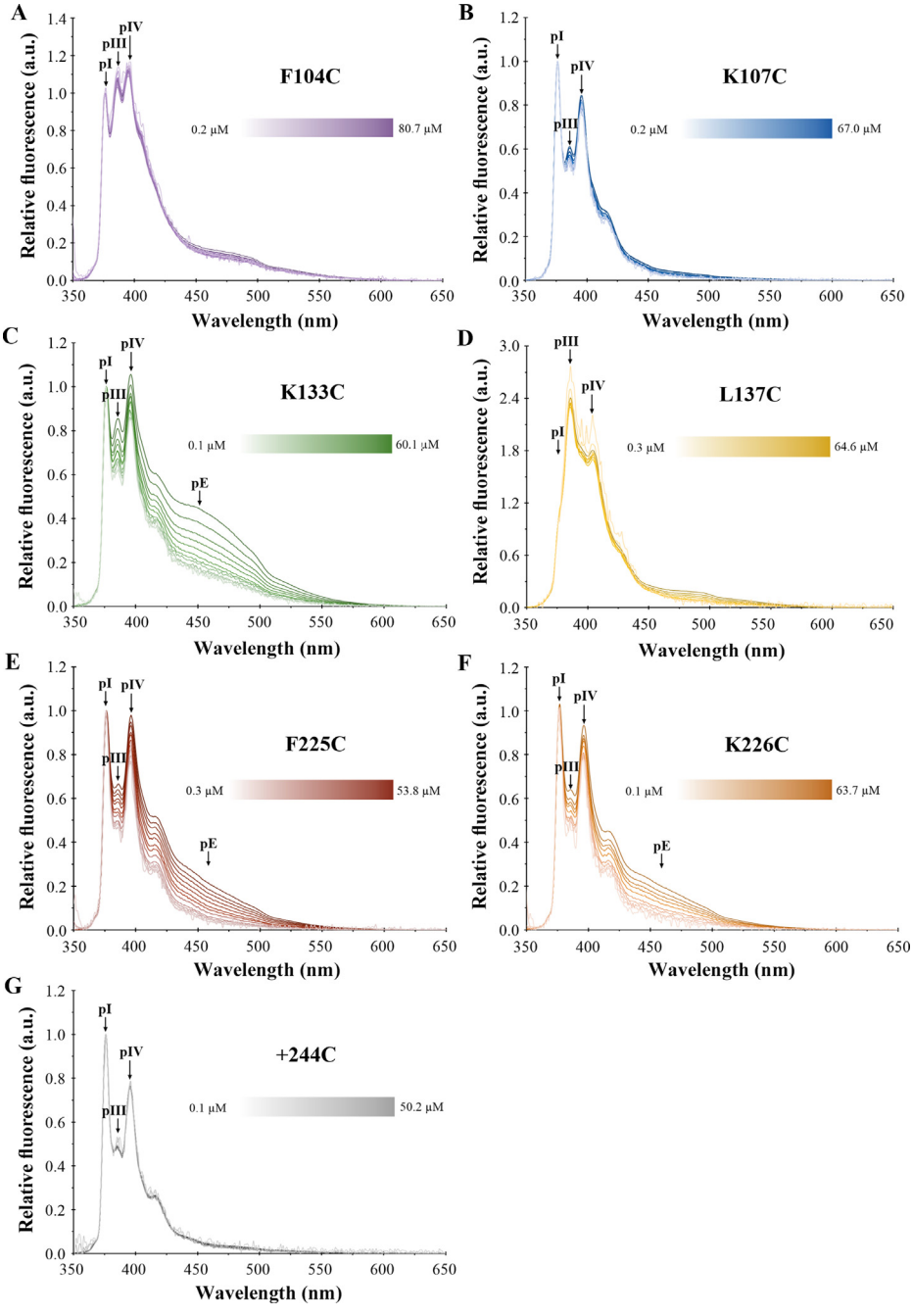

Fig. 5. Fluorescence emission spectra of lipid-free apoA-I mutants labelled with pyrene. Panels A)-F), spectra were obtained at different apoA-I concentrations (colour degree scale) for different apoA-I variants labelled with pyrene. For clarity, the spectra were relativised to 376 nm emission (pI). Peaks are shown with arrows: 376 nm (pI), 384 nm (pIII), 395 nm (pIV) and 460 nm excimer (pE). Colour code: F104C (purple), K107C (blue), K133C (green), L137C (yellow), F225C (red), K226C (orange) and +244C (black). Fig. 5B and C appears in the original research article that accompanies this data article [1]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

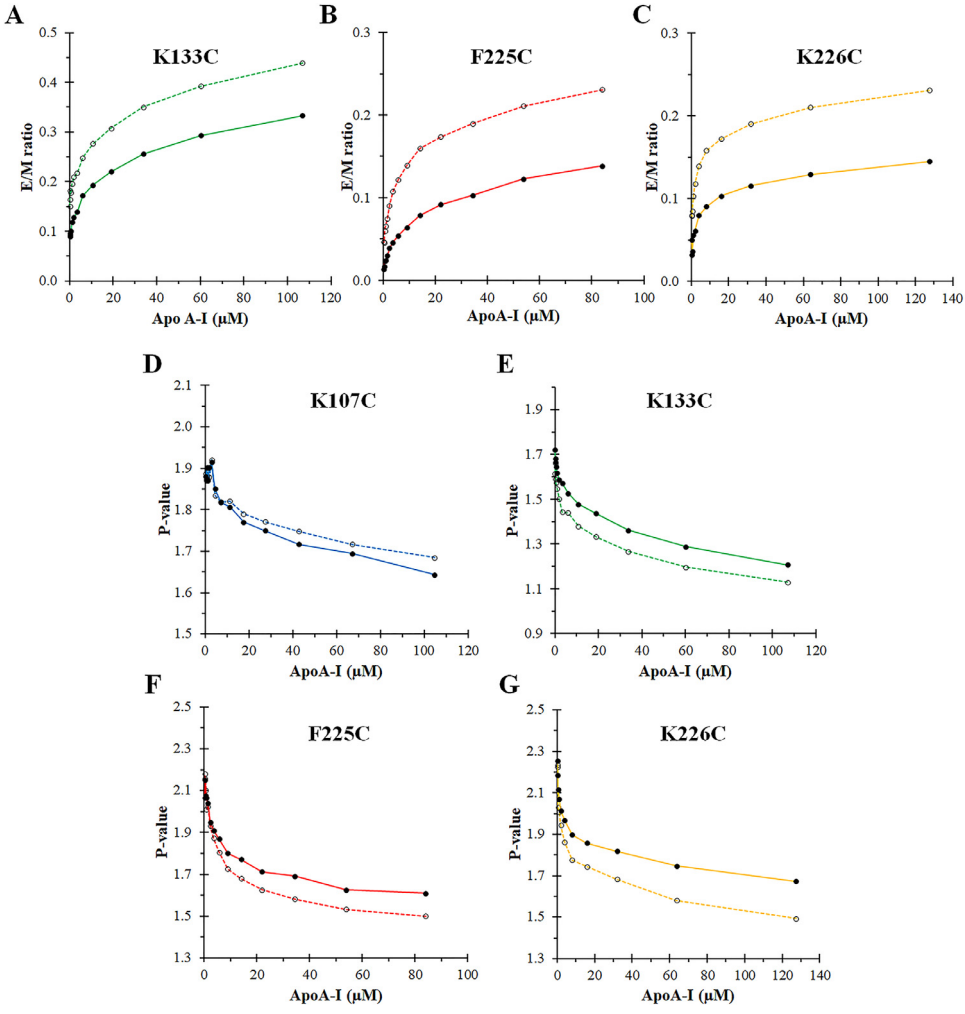

Fig. 6. E/M ratio and *P*-value data. Corrected (dashed line) and Corrected and deconvoluted data (solid line) are presented for apoA-I variant labelled with pyrene which showed changes on E/M panels A)-C) and *P*-value panel D)-G), respectively. Colour code: K107C (blue), K133C (green), F225C (red) and K226C (orange). Fig. 6A and E appears in the original research article that accompanies this data article [1]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

effect on excimer and *P*-value data are shown in Fig. 6. Data from deconvolution effect on E/M and *P*-value signals for labelled apoA-I variants that showed pyrene spectral changes are available in Figure_6.xlsx file. In addition, the Mathematica script used for deconvolution spectra is available as Mathematica_Script.nb with a sample file Sample_K133C_Pyr.xlsx for typical fluorescence emission spectra titration. Files are accessible in the online repository.

1.5. Mathematical models

Corrected and deconvoluted data were fitted to two mathematical models, Excimer/Monomer ratio (E/M ratio) and *P*-value, using the Solver Add-in functions from MS Excel. A weighted

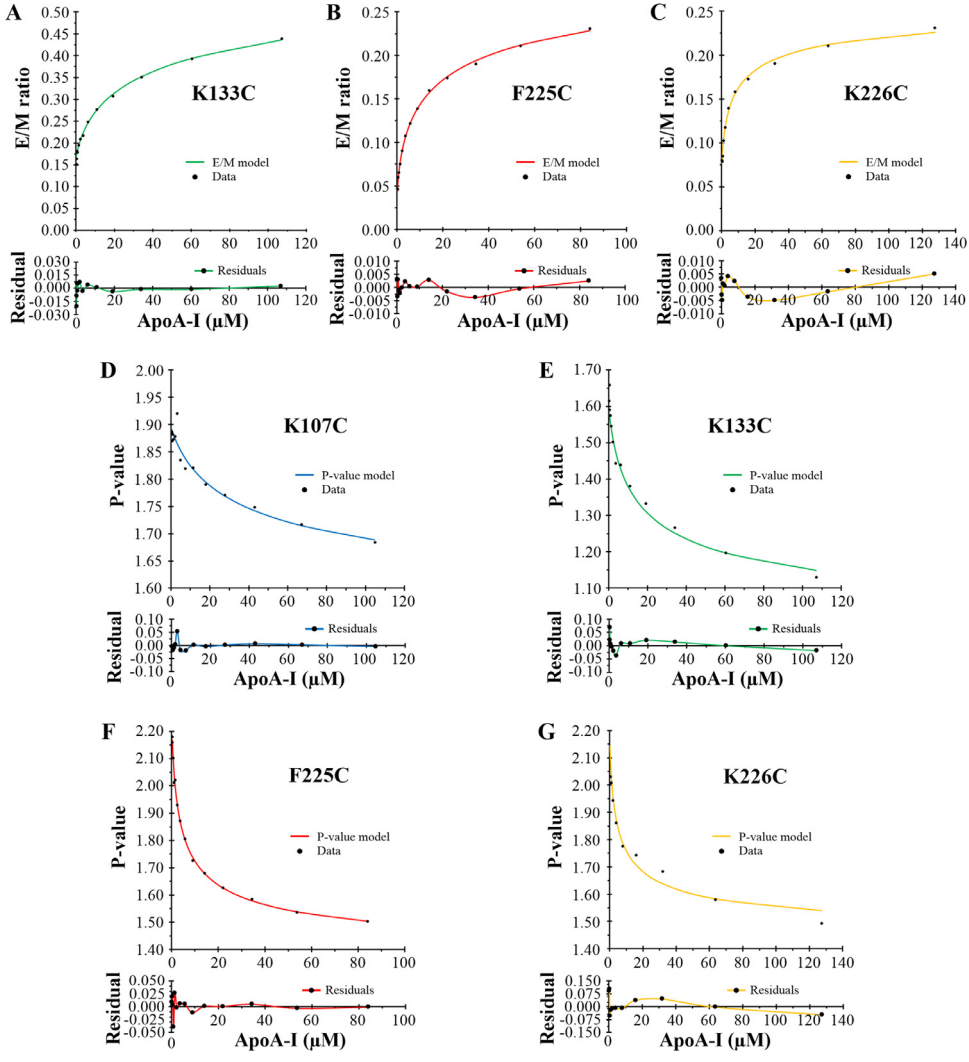

Fig. 7. Fitting examples with corrected spectra. E/M and *P*-value models (colour line) were fitted to data from corrected spectra (black circles) on panels A)-C) and D)-G), respectively. Colour code: K107C (blue), K133C (green), F225C (red) and K226C (orange). Figs. 7A and 6E appear in the original research article that accompanies this data article [1]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

fitting was applied at low concentrations (bellow 1 μM) to reduce errors associated with low signal/noise ratio of spectra of diluted samples. Constrains were defined and fixed according to the experimental design so that fitting parameters can only take values with physical meaning based on protein concentration and fluorescence intensity. The fitting and the residues distribution for K107C, K133C, F225C and K226C are presented in Figs. 7 and 8 with data from fitting curves available in Figure_7.xlsx and Figure_8.xlsx files. Otherwise, the calculated parameters are shown in Tables 2 and 3 with data in Table_7.xlsx and Table_8.xlsx files. Files are accessible in the online repository.

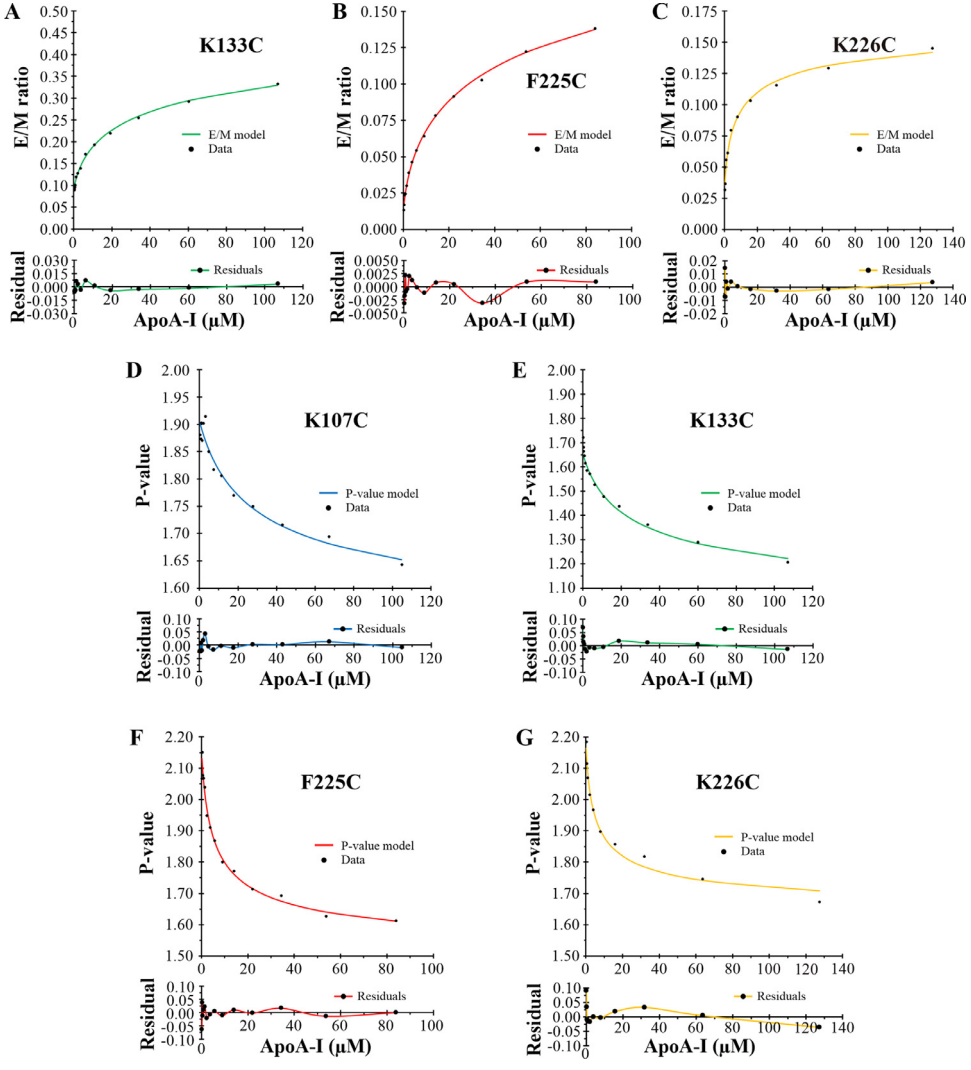

Fig. 8. Fitting examples with corrected and deconvoluted spectra. E/M and *P*-value models (colour line) were fitted to data from corrected and deconvoluted spectra (black circles) on panels A)-C) and D)-G), respectively. Colour code: K107C (blue), K133C (green), F225C (red) and K226C (orange). Fig. 8A and E appears in the original research article that accompanies this data article [1]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2. Experimental Design, Materials and Methods

2.1. Expression and purification of apolipoprotein A-I cysteine variants

The purification protocol used is that described by Prieto [7] and Ryan [8], with some modifications. Briefly, *Escherichia coli* strains BL21 (DE3) were transformed with pET30a(+) human apoA-I plasmids carrying WT or mutants cDNA. All plasmids carry an E2D mutation that generates an acid-labile site for easier purification. Over-expression of proteins in Luria Bertani (LB) medium at 37 °C was induced with 0.4 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) for 3-4 h. Cells were collected, lysed by sonication and clarified in a buffer containing 25 mM

Table 2
Parameters of the apoA-I self-association equilibrium calculated with models following E/M ratio and *P*-value changes using corrected spectra.

Parameter	K133C E/M	F225C E/M	K226C E/M	K107C *P*-value	K133C *P*-value	F225C *P*-value	K226C *P*-value
K_A	0.2 ± 0.2	0.4 ± 0.2	0.31 ± 0.02	0.017 ± 0.005	0.04 ± 0.01	0.19 ± 0.03	0.1 ± 0.1
h2	1 ± 0	0.92 ± 0.08	1 ± 0	(n.a.)	(n.a.)	(n.a.)	(n.a.)
Yd	1016 ± 452	2372 ± 756	1882 ± 456	(n.a.)	(n.a.)	(n.a.)	(n.a.)
Ym	3684 ± 947	10,462 ± 1905	7867 ± 34	(n.a.)	(n.a.)	(n.a.)	(n.a.)
Yḿ	259 ± 31	335 ± 26	409 ± 28	(n.a.)	(n.a.)	(n.a.)	(n.a.)
YpIA	(n.a.)	(n.a.)	(n.a.)	4754 ± 347	4209 ± 252	8016 ± 1502	4526 ± 1254
YpIA2	(n.a.)	(n.a.)	(n.a.)	3071 ± 344	1612 ± 149	3029 ± 562	2953 ± 970
YpIIIA	(n.a.)	(n.a.)	(n.a.)	2777 ± 503	2519 ± 219	3582 ± 612	2695 ± 107
YpIIIA2	(n.a.)	(n.a.)	(n.a.)	2777 ± 503	2519 ± 219	3582 ± 612	3394 ± 1096
YṕIA	(n.a.)	(n.a.)	(n.a.)	516 ± 104	630 ± 55	895 ± 153	341 ± 443
YṕIA2	(n.a.)	(n.a.)	(n.a.)	140 ± 70	0 ± 0	0 ± 0	0 ± 0
YṕIIIA	(n.a.)	(n.a.)	(n.a.)	268 ± 198	183 ± 99	488 ± 296	381 ± 178
YṕIIIA2	(n.a.)	(n.a.)	(n.a.)	307 ± 180	197 ± 90	488 ± 296	396 ± 199
R^2	> 0.992	> 0.998	> 0.996	> 0.891	> 0.952	> 0.983	> 0.952

Parameter values are averages ± standard errors of the mean between titrations of independent preparations of proteins, including purification and labelling: K107C (4); K133C (3); F225C (3); K226C (2). Minimum R^2 is also reported for each mutant. The K_A values appear in the original research article that accompanies this data article [1]. (n.a.): not applicable.

Table 3
Parameters of the apoA-I self-association equilibrium calculated with models following E/M ratio and *P*-value changes using corrected and deconvoluted spectra.

Parameter	K133C E/M	F225C E/M	K226C E/M	K107C *P*-value	K133C *P*-value	F225C *P*-value	K226C *P*-value
K_A	0.09 ± 0.02	0.16 ± 0.08	0.27 ± 0.09	0.014 ± 0.003	0.021 ± 0.006	0.12 ± 0.02	0.07 ± 0.04
h2	0.9 ± 0.1	0.9 ± 0.1	1 ± 0	(n.a.)	(n.a.)	(n.a.)	(n.a.)
Yd	1890 ± 64	964 ± 260	1098 ± 1214	(n.a.)	(n.a.)	(n.a.)	(n.a.)
Ym	5453 ± 1810	7191 ± 3617	5521 ± 5850	(n.a.)	(n.a.)	(n.a.)	(n.a.)
YpIA	(n.a.)	(n.a.)	(n.a.)	3737 ± 410	2451 ± 362	5255 ± 565	3739 ± 328
YpIA2	(n.a.)	(n.a.)	(n.a.)	2903 ± 130	1064 ± 123	3132 ± 798	2443 ± 569
YpIIIA	(n.a.)	(n.a.)	(n.a.)	2569 ± 232	2030 ± 334	3226 ± 582	2615 ± 824
YpIIIA2	(n.a.)	(n.a.)	(n.a.)	2823 ± 232	2030 ± 334	3438 ± 669	2615 ± 824
R^2	> 0.997	> 0.998	> 0.975	> 0.866	> 0.948	> 0.965	> 0.962

Parameter values are averages ± standard errors of the mean between titrations of independent preparations of proteins, including purification and labelling: K107C (4); K133C (3); F225C (3); K226C (2). Minimum R^2 is also reported for each mutant. The K_A values appear in the original research article that accompanies this data article [1]. (n.a.): not applicable.

$KHPO_4$, 500 mM NaCl, 5 mM imidazole, 10 mM β-mercaptoethanol (BME) and 3 M GdHCl at pH 7.4 (buffer EQB, equilibrium buffer with 3 M GdHCl). The soluble fusion proteins were purified using HisTrap columns by nickel affinity (IMAC, Immobilised Metal Ion Affinity Chromatography) with an FPLC system, washing and eluting with 35 and 100 mM imidazole. The purified fusion proteins were treated with 45% formic acid during 5 hs at 55 °C (BME and GdHCl were maintained) to release the target proteins without the first two N-terminal amino acids. Proteins were thoroughly dialysed in EQB buffer during 24 hs at 4 °C to complete formic acid removal. A second run through the HisTrap column removed the His-tag fragment and the pure target proteins was eluted from the column with only 5 mM imidazole. The target protein was dialysed against 30 mM Tris-HCl, pH 8.0, concentrated using Centricon filters (10 kDa) and loaded onto a MonoQ (GE Healthcare) anion exchange chromatography column to be eluted with a 1 M NaCl gradient from 10 to 35%. Finally, the fractions containing target proteins were pooled, concentrated and stored with 3 M GdHCl and 10 mM BME at −20 °C. After the first IMAC step, the protein concentration was monitored and quantified by absorbance at 280 nm ($A_{280\ nm}$) in the presence of 3 M GdHCl, using a molar extinction coefficient (ε^{280nm}) of 37,930 and 32,430 $cm^{-1}M^{-1}$ for uncleaved and cleaved apoA-I [7], respectively, and their purity was determined by SDS-PAGE. WT apoA-I purification required no BME.

Previous to labelling apoA-I cysteine variants with NPM the proteins were dialysed for 12 hs in buffer PBSE (25 mM $NaHPO_4$, 150 mM NaCl and 7 mM EDTA, pH 6.5) to complete BME elimination. The working pH was maintained at 6.5 to avoid the tendency of NPM to undergo aminolysis at pH $\geq$ 7.5, especially in the presence of free amines [9]. Initially, the proteins were incubated with Tris(2-carboxyethyl) phosphine (TCEP) in PBSE with GdHCl (3 M) and dimethyl sulfoxide (30% DMSO) for 1 h at 25 °C to reduce disulphide bonds. This process was followed by incubation with a five-fold molar excess (over apoA-I) of NPM (dissolved in DMSO) during 3 hs at 25 °C with the addition of 1/3 of NPM after each hour and renewing the TCEP. Then, five-fold molar excess reduced Glutathione (GSH) (over NPM) was added and incubated for 1 h to block the unreacted NPM. Then, the labelled samples were passed through an exclusion chromatography column (Sephadex G25) to remove excess of TECP, GdHCl and GS-NPM and protein was eluted with PBS buffer supplemented with 4 M Urea at pH 6. Finally, proteins were concentrated before being dialysed in PBSE + BME buffer and stored with 3 M GdHCl at −20 °C. The degree of labelling was determined using the extinction coefficients of 36,000 $M^{-1}cm^{-1}$ at 343 nm for pyrene [10], and Bicinchoninic Acid Protein Assay (Thermo Scientific) for apoA-I. For each experiment, fresh protein samples were thawed and dialysed against PBS buffer prior to use.

2.2. The tryptophan intrinsic fluorescence

The tryptophan (Trp) intrinsic fluorescence was employed to analyse apoA-I cysteine variants folding and stability in solution. Protein samples were dissolved in PBS buffer at 3.5 µM. Fluorescence emission spectra were acquired in an SLM4800 spectrofluorometer (Applied Photophysics) updated by Olis (ISS Inc., Champaign, IL) exciting at 295 nm and collecting emission from 300 to 500 nm with 1 nm resolution at increasing concentrations of GdHCl, from 0 to 3 M. Centre of mass (CM) was used to follow changes in Trp-fluorescence (λ_{exc} 295 nm) with a limited integration range from 310 to 370 nm to avoid most pyrene emission (from 360 to 600 nm). Wavelength Maximum Fluorescence (WMF_0) was also reported at 0 M GdHCl for each variant. The denaturation free energy in water ($\Delta G°_{H20}$, 25 °C and 1 atm, buffer PBS pH 7.4) and the cooperativity parameter (m) were obtained by fitting the N-U (Native-Unfolded) model [11] represented by the following equation:

$$CM(x) = \frac{YN_{H20} + e^{\frac{-\Delta G_{H20} + mx}{RT}} YU_{H20}}{1 + e^{\frac{-\Delta G_{H20} + mx}{RT}}} \quad (3)$$

The *x* represents the GdHCl concentration, CM is the spectrum centre of mass position; YN_{H2O} and YU_{H2O} are the specific signals for native and unfolded states in aqueous buffer, respectively, R is the Boltzmann's constant and T the temperature set at 298° K (25 °C).

2.3. Models considered for labelled apoA-I self-association

2.3.1. E/M ratio model of association-dissociation equilibrium for apoA-I

The apoA-I association process was represented as follow:

$$A + A \leftrightarrow A2 \; K_A = \frac{[A2]}{[A]^2} \quad (M1\text{–}1)$$

where an association constant (K_A) was defined.

Essentially, the fluorescence to each wavelength from a pyrene spectrum can be represented as a sum of their contributions, pure and bleed-through from neighbouring peaks. The general definition was applied to the monomer (Fm) and excimer (Fd) fluorescence described in the following equations:

$$Fm = Ym * [A] + Y'd * [A2] \quad (M1\text{–}2)$$

$$Fd = Yd * [A2] + Y'm * [A] \quad \text{(M1–3)}$$

where [A] and [A2] were the monomer and dimer species concentrations; Ym, Yd, Yḿ and Yd́ were the specific fluorescence from monomer and dimer in their pure and cross-talk (Y and Ý) forms.

We have also considered the *degree of labelling* (g) to estimate the occupancy of labelled monomers within all species according to a random distribution. For a monomer-dimer equilibrium:

$$[A] = g * [A] + (1 - g) * [A] \quad \text{(M1–4)}$$

$$[A2] = g^2 * [A2] + 2 * g * (1 - g) * [A2] + (1 - g)^2 * [A2] \quad \text{(M1–5)}$$

where the last components of each specie did not emit fluorescence, and the 2*g*(1-g)*[A2] could only emit as a monomer.

Appling these distributions on Eqs. (M1-2) and (M1-3) and considering that the monomer fluorescence did not change due to the association, we have obtained equations (Eqs. (M1-6) and (M1-7)):

$$Fm = Ym * (g * [A]) + Y'd * (g^2 * [A2]) + Ym * (2 * g * (1 - g) * [A2]) \quad \text{(M1–6)}$$

$$Fd = Y'm * (g * [A]) + Yd * (g^2 * [A2]) + Y'm * (2 * g * (1 - g) * [A2]) \quad \text{(M1–7)}$$

Finally, we have also considered that the excimer formation could not be permanent and that the excimer emission would only be present for a short time due to less optimal orientation of the probes. We have introduced a parameter called *efficiency of excimer formation* ($0 < h2 < 1$), it was only applied to the specie $g^2 * [A2]$ to consider the fraction emitting as an excimer, while the other fraction would emit also as a monomer:

$$Fd(g^2 * [A2]) = h2 * Yd * [A2] + (1 - h2) * (2 * Y'm) * [A2] \quad \text{(M1–8)}$$

$$F'd(g^2 * [A2]) = h2 * Y'd * [A2] + (1 - h2) * (2 * Ym) * [A2] \quad \text{(M1–9)}$$

where the parameters (Ym and Yḿ) were expressed as pyrene unit.

Altogether, if we applied the fluorescence ratio between all species that emit as excimer or as monomer (Eqs. (M1-6) and (M1-7)) and considering equations M1-8 and M1-9 we obtained the following equation:

$$\frac{Fd}{Fm} = \frac{(h2 * Yd * (g^2 * [A2]) + (1 - h2) * (2 * Y'm) * (g^2 * [A2]) + Y'm * (2 * g * (1 - g) * [A2]) + Y'm * (g * [A]))}{(h2 * Y'd * (g^2 * [A2]) + (1 - h2) * (2 * Ym) * (g^2 * [A2]) + Ym * (2 * g * (1 - g) * [A2]) + Ym * (g * [A]))} \quad \text{(M1–10)}$$

Finally, when the equation M1-10 was combined with the positive solution of equations of mass balance ($[At] = [A] + 2 * [A2]$), and the association constant definition (Eq. (M1-1)) and simplifying, we obtained an equation that is function of the total protein concentration ([At]):

$$\frac{Fd}{Fm}([At]) = \frac{g * h2 * \left(-1 - 4 * [At] * K_A + \sqrt{1 + 8 * [At] * K_A}\right) * (Yd - 2 * Y'm) - 8 * [At] * K_A * Y'm}{g * h2 * \left(-1 - 4 * [At] * K_A + \sqrt{1 + 8 * [At] * K_A}\right) * (Y'd - 2 * Ym) - 8 * [At] * K_A * Ym} + c1 \quad \text{(M1–11)}$$

where c1 is an error constant that contains all the approximations we considered.

Knowing that the excimer bleed-through into the monomer emission (Yd́) is negligible due to the asymmetry of pyrene fluorescence and deconvolution data, we simplified the model by

eliminating this contribution, also aiming to decrease the number of parameters. Therefore, final model to be fitted to Excimer/Monomer fluorescence emission ratio was:

$$\frac{Fd}{Fm}([At]) = \frac{g*h2*\left(-1-4*[At]*K_A+\sqrt{1+8*[At]*K_A}\right)*\left(Yd-2*Y'm\right)-8*[At]*K_A*Y'm}{g*h2*\left(-1-4*[At]*K_A+\sqrt{1+8*[At]*K_A}\right)*(-2*Ym)-8*[At]*K_A*Ym}+c1 \quad \text{(M1–12)}$$

2.3.2. P-*value model of association-dissociation equilibrium for apoA-I*

Modelling *P*-value changes as a function of total protein concentration followed the same rational as described previously.

The maximum intensities of peak I and III (FpI and FpIII) were represented as a sum of pure and bleed-through contributions:

$$FpI = YpIA*[A] + YpIA2*[A2] + Y'pIIIA*[A] + Y'pIIIA2*[A2] \quad \text{(M2–1)}$$

$$FpIII = YpIIIA*[A] + YpIIIA2*[A2] + Y'pIA*[A] + Y'pIA2*[A2] \quad \text{(M2–2)}$$

where pyrene excimer was not considered. The parameters YpIA, YpIIIA, YpIA2, YpIIIA2 and YṕIA, YṕIIIA, YṕIA2, YṕIIIA2 are the specific and their cross-talk (Ý) fluorescence signals.

Considering a random distribution between labelled and unlabelled protein (Eqs. (M1-4) and (M1-5), E/M model), we obtained the following equations:

$$\begin{aligned} FpI = {} & (YpIA*(g*[A]) + (2*YpIA2)*\left(g^2*[A2]\right) + YpIA2*(2*g*(1-g)*[A2]) \\ & + Y'pIIIA*(g*[A]) + (2*Y'pIIIA2)*(g^2*[A2]) + Y'pIIIA2*(2*g*(1-g)*[A2])) \end{aligned} \quad \text{(M2–3)}$$

$$\begin{aligned} FpIII = {} & YpIIIA*(g*[A]) + (2*YpIIIA2)*\left(g^2*[A2]\right) + YpIIIA2*(2*g*(1-g)*[A2]) \\ & + Y'pIA*(g*[A]) + \left(2*Y'pIA2\right)*\left(g^2*[A2]\right) + Y'pIA2*(2*g*(1-g)*[A2]) \end{aligned} \quad \text{(M2–4)}$$

where the monomer fluorescence in dimeric species with two molecules of pyrene we expressed as a pyrene unit.

The *P*-value function is obtained making the relation between pI and pIII signals:

$$\begin{aligned} FpI/FpIII = {} & (YpIA*(g*[A]) + (2*YpIA2)*(g^2*[A2]) + YpIA2*(2*g*(1-g)*[A2]) \\ & + Y'pIIIA*(g*[A]) + (2*Y'pIIIA2)*(g^2*[A2]) + Y'pIIIA2*(2*g*(1-g)*[A2]))/ \\ & (YpIIIA*(g*[A]) + (2*YpIIIA2)*(g^2*[A2]) + YpIIIA2*(2*g*(1-g)*[A2]) \\ & + Y'pIA*(g*[A]) + (2*Y'pIA2)*(g^2*[A2]) + Y'pIA2*(2*g*(1-g)*[A2])\ ) \end{aligned} \quad \text{(M2–5)}$$

If we simplified (Eq. (M2-5)), we obtained (Eq. (M2-6)):

$$P\ value = \frac{\left(YpIA + Y'pIIIA\right)*[A] + \left(YpIA2 + Y'pIIIA2\right)*2*[A2]}{(YpIIIA + Y'pIA)*[A] + (YpIIIA2 + Y'pIA2)*2*[A2]} \quad \text{(M2–6)}$$

As we described previously, merging the positive solution of mass balanced equation and association constant definition we obtain the equation (Eq. (M2-7)), that we used as model to be fitted to *P*-value data.

$$P-value([At]) = \frac{\left(YpIA + Y'pIIIA\right)*\left(\frac{-1+\sqrt{1+8*[At]*K_A}}{4*K_A}\right) + \left(YpIA2 + Y'pIIIA2\right)*2*\left(\frac{1}{8}*\left(4*[At] + \frac{1}{K_A} - \frac{\sqrt{1+8*[At]*K_A}}{K_A}\right)\right)}{(YpIIIA + Y'pIA)*\left(\frac{-1+\sqrt{1+8*[At]*K_A}}{4*K_A}\right) + (YpIIIA2 + Y'pIA2)*2*\left(\frac{1}{8}*\left(4*[At] + \frac{1}{K_A} - \frac{\sqrt{1+8*[At]*K_A}}{K_A}\right)\right)} + c2 \quad \text{(M2–7)}$$

where c2 is an error constant that contains all the approximations we considered.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET, PUE 22920160100002); Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT) PICT 2016-0915 and PICT 201-0613; Universidad Nacional de La Plata (UNLP) M183.

References

[1] W.A. Tárraga, H. Falomir-Lockhart, H.A. Garda, M.C. González, Analysis of pyrene-labelled apolipoprotein A-I oligomerization in solution: spectra deconvolution and changes in P-value and excimer formation, Arch. Biochem. Biophys. (2021) 699, doi:10.1016/j.abb.2020.108748.
[2] G. Bains, A.B. Patel, V. Narayanaswami, Pyrene: a probe to study protein conformation and conformational changes, Molecules 16 (2011) 7909–7935, doi:10.3390/molecules16097909.
[3] K. Kalyanasundaram, J.K. Thomas, Environmental effects on vibronic band intensities in pyrene monomer fluorescence and their application in studies of micellar systems, J. Am. Chem. Soc. 99 (1977) 2039–2044, doi:10.1021/ja00449a004.
[4] H. Langhals, A re-examination of the line shape of the electronic spectra of complex molecules in solution: log-normal function versus Gaussian, Spectrochim. Acta A 56 (2000) 2207–2210, doi:10.1016/S1386-1425(00)00274-2.
[5] E.A. Burstein, V.I. Emelyanenko, Log-normal description of fluorescence spectra of organic fluorophores, Photochem. Photobiol. 64 (1996) 316–320, doi:10.1111/j.1751-1097.1996.tb02464.x.
[6] E.A. Burstein, S.M. Abornev, Y.K. Reshetnyak, Decomposition of protein tryptophan fluorescence spectra into log-normal components. I. Decomposition algorithms, Biophys. J. 81 (2001) 1699–1709, doi:10.1016/S0006-3495(01)75823-7.
[7] E.D. Prieto, N. Ramella, L.A. Cuellar, M.A. Tricerri, H.A. Garda, Characterization of a human apolipoprotein a-I construct expressed in a bacterial system, Protein J. 31 (2012) 681–688, doi:10.1007/s10930-012-9448-z.
[8] R.O. Ryan, T.M. Forte, M.N. Oda, Optimized bacterial expression of human apolipoprotein A-I, Protein Expr. Purif. 27 (2003) 98–103, doi:10.1016/s1046-5928(02)00568-5.
[9] C.W. Wu, L.R. Yarbrough, F.Y.H. Wu, N-(1-pyrene)maleimide: a fluorescent cross-linking reagent, Biochemistry 15 (1976) 2863–2868, doi:10.1021/bi00658a025.
[10] J.I. Gallea, M.S. Celej, Structural insights into amyloid oligomers of the Parkinson disease-related protein α-synuclein, J. Biol. Chem. 289 (2014) 26733–26742, doi:10.1074/jbc.M114.566695.
[11] D.W. Bolen, M.M.S., Unfolding free energy changes determined by the linear extrapolation method . 1 . Unfolding of phenylmethanesulfonyl alfa-chymotrypsin using different denaturants, Am. Chem. Soc. (1988) 8063–8068, doi:10.1021/bi00421a014.