Proof-of-concept analytical instrument for label-free optical de-convolution of protein species in a mixture

John E. Hales , Samir Aoudjane , Gabriel Aeppli , Paul A. Dalby

PII: S0021-9673(21)00092-3
DOI: https://doi.org/10.1016/j.chroma.2021.461968
Reference: CHROMA 461968

To appear in: *Journal of Chromatography A*

Received date: 23 October 2020
Revised date: 18 January 2021
Accepted date: 1 February 2021

Please cite this article as: John E. Hales , Samir Aoudjane , Gabriel Aeppli , Paul A. Dalby , Proof-of-concept analytical instrument for label-free optical de-convolution of protein species in a mixture, *Journal of Chromatography A* (2021), doi: https://doi.org/10.1016/j.chroma.2021.461968

Highlights

- Real-time optical deconvolution of two proteins.
- Label-free chromatography monitoring and quantitation of co-eluting proteins.
- Protein species have signature intrinsic fluorescence lifetime.
- Potential to use for process analytical technology (PAT).

Proof-of-concept analytical instrument for label-free optical de-convolution of protein species in a mixture

Author names and affiliations

John E. Hales[a]

Samir Aoudjane[a]

Gabriel Aeppli[b, 1]

Paul A. Dalby[a]

[a]Department of Biochemical Engineering, University College London, Bernard Katz Building, Gower Street, London, WC1E 6BT, UK.

[b]London Centre for Nanotechnology, 17-19 Gordon Street, London, WC1H 0AH, UK

Corresponding authors

John E. Hales:

Department of Biochemical Engineering, University College London, Bernard Katz Building, Gower Street, London, WC1E 6BT, UK.

Contact telephone number: +44 (0)20 7679 9611

Email address: john.hales@ucl.ac.uk

Paul A. Dalby:

Department of Biochemical Engineering, University College London, Bernard Katz Building, Gower Street, London, WC1E 6BT, UK.

Contact telephone number: +44 (0)2076799566

Email address: p.dalby@ucl.ac.uk

Present address

The present addresses of Gabriel Aeppli are:

[1]Paul Scherrer Institut, Villigen, PSI CH-5232, Switzerland.

[1]Department of Physics, ETH Zürich, Zürich, CH-8093, Switzerland.

[1]Institut de Physique, EPFL, Lausanne, CH-1015, Switzerland.

A**bstract**

The adoption of process analytical technologies by the biopharmaceutical industry can reduce the cost of therapeutic drugs and facilitate investigation of new bioprocesses. Control of critical process parameters to retain critical product quality attributes within strict bounds is important for ensuring a consistently high product quality, but developing the sophisticated analytical technologies required has proven to be a major challenge. Here, we demonstrate a new optical technique for continuous monitoring of protein species as they are eluted from a chromatographic column, even when they fully co-elute with other protein species, without making any assumption about or peak-fitting to the elution profile. To achieve this, we designed and constructed a time-resolved intrinsic fluorescence lifetime chromatograph, and established an analytical framework for deconvolving and quantifying distinct but co-eluting protein species in real time. This proof-of-concept technology has potentially useful applications as a process analytical technology and more generally as an analytical technique for label-free quantification of proteins in mixtures.

Keywords

Analytical instrumentation,

Biophysics,

Co-eluting proteins,

Decay-associated chromatogram,

Process analytical technology,

Time-resolved intrinsic fluorescence.

1. Introduction

The effective chromatographic separation of protein-based biopharmaceuticals from other biomolecules, based on their size, charge, hydrophobicity or binding affinity, is essential at manufacturing scales for achieving consistently high product quality. Chromatography was also one of the first bioprocessing steps to be amenable to continuous manufacturing,[1] requiring process analytical technology (PAT), where process parameters are continuously monitored and fed back for control of the process itself. A major barrier to a more effective PAT approach for chromatography is the lack of real-time and robust, in-line or on-line sensors, that can distinguish the product from biophysically similar proteins that tend to co-elute with it. Such proteins may derive from the host cell, or could be variants of the product protein including those that result from proteolysis, alternative post-translational modifications, oxidation, misfolding, aggregation, and deamidation.[2]

On-line UV-absorption or refractive-index measurements are often used to monitor chromatography, but if there are two or more overlapping peaks in the chromatogram then these techniques cannot readily be used to quantify the relative amounts of the co-eluting proteins. Co-eluting protein species can be resolved to some degree by measuring absorption spectra and analysing the data with advanced tools such as partial least squares regression.[3-8] However, this typically involves extensive off-line calibration and the spectral shifts between proteins tend to be subtle, making it difficult to resolve more than two co-eluting proteins in real time.

Other common detection modalities, such as fluorescence intensity and electrical conductivity, are also typically univariate, and therefore suffer the same drawback as UV-absorption measurements. Off-line analytical methods such as protein gel electrophoresis, capillary electrophoresis, orthogonal analytical chromatography modes, or mass-spectrometry cannot iden-

tify co-eluting contaminants in real time, and so would not provide adequate feedback to control the manufacturing process.

Time-resolved fluorescence lifetime detection (TRF) is a well-established spectroscopic technique that can identify changes to protein conformations or their association with ligands, based on the time-profile of their intrinsic fluorescence.[9-12] The lifetime and intensity of this fluorescence is dependent on the number, relative location and dipolar microenvironment of the aromatic residues tryptophan and tyrosine, which are present in 99 % of all proteins. This technique has not been previously combined with liquid chromatography for the real-time detection of proteins, although there is precedent for distinguishing polycyclic aromatic hydrocarbon compounds (PAHs) that co-elute by HPLC, based on their fluorescence lifetimes.[13-17] However, the analytical instrumentation used to measure the lifetimes of PAHs is ineffective for distinguishing proteins because the difference in the fluorescence decay times between similar proteins can be less than 0.1 ns, which is a major challenge to resolve in real time. In addition, a shift to high-bandwidth detection electronics is required because proteins have fluorescence lifetimes typically in the region of 0.5 ns to 6 ns, which is much shorter than the 20 ns or greater that is typical for PAHs. Finally, UV-transparent optics must be used for protein detection as they absorb light at wavelengths between 250 nm to 290 nm. Aside from PAHs, the lifetimes of two dyes migrating at different rates through a capillary electrophoresis system have been measured on-the-fly[18] and time-filtered detection of organic molecules has been shown to help reduce the background signal.[19]

The most common technique for measuring intrinsic fluorescence lifetimes is time-correlated single photon counting (TCSPC) but, unless analysing very dilute solutions, TCSPC is susceptible to the “pile-up” problem: if the fluorescence is too intense, then the distribution of counted photons becomes statistically biased towards shorter times, compromising the accu-

racy of the decay measurement and also its ability to resolve multiple decay components.[20] Consequently, TCSPC is poorly suited to continuous measurements of both the fluorescence intensity, which is indicative of the protein concentration and therefore essential for PAT applications, and the fluorescence lifetimes. High bandwidth digital oscilloscopes have also been used to directly measure the fluorescence lifetimes of PAHs but would not be ideal for the detection of proteins due to their shorter decay times.[16] Frequency-domain fluorescence lifetime detection has instead been the preferred technique for PAHs, and could be adapted for proteins,[15] but would require continuous-wave instead of pulsed-excitation sources, and therefore either significantly more sensitive detectors, or much higher excitation powers for signal recovery, and also additional approaches that can account for DC noise sources. Compounding this, proteins have a relatively low quantum yield and nanosecond decay times so the detectors would need to be both very sensitive and have a high bandwidth.

Here, we have constructed an early-stage experimental prototype to monitor chromatographic eluants using TRF that enables the real-time quantification of the contributions from multiple protein species. The TRF approach uses an alternative technique to TCSPC that negates the pile-up problem, enabling simultaneous measurement of changes in both the time decay and the fluorescence intensity, at intervals of under 5 s, on protein solutions at a wide range of concentrations relevant to bioprocessing and drug formulation. We developed algorithms that can fit the data concurrently with data acquisition, enabling real-time product monitoring and pooling decisions. We also developed analytical tools that can quantify individual protein species as they are eluted from the column, even if they fully co-elute with another protein species, and without making any assumptions about the elution profile, or any requirements for peak-fitting to Gaussians for example.

2. Materials and methods

We designed and constructed a time-resolved intrinsic fluorescence lifetime detector for chromatography, consisting of a liquid flow cell, optical and electronic sub-assemblies (TRF chromatograph) (Figure 1), linked directly to a size-exclusion FPLC (SEC) column (GE Healthcare, Superdex 200 Increase 10/300GL).

Bovine serum albumin (BSA) and ovalbumin were supplied by Sigma Aldrich UK (Gillingham, UK). Buffer constituents, Sodium Phosphate di/monobasic and Sodium Chloride were supplied by Sigma Aldrich UK. Ultrapure water (>18MΩ.cm) was used in the preparation of all buffers. Solutions were degassed by helium sparging (BOC group) for a minimum of 10 min[21] and filtered by 0.22 µm “Stericup” vacuum filtration units supplied by Merck & Co. (Kenilworth, New Jersey, USA).

Excitation pulses with a pulse width of 1 ns at a wavelength of 266 nm emitted by a diode-pumped Q switched solid state laser (Crylas, FQSS 266-50) were focused towards an optically transparent window within a polymer-coated UV-fused silica capillary with a 300 µm bore (Molex, TSP300665) via beam-shaping optics, where they could excite the aromatic residues of proteins eluted into the capillary from the SEC connected to a modified FPLC pump (GE Healthcare, AKTA AKTAfplc). The window was exposed by removing the polymer coating with a blue flame. The focus of the laser was not coincident with the flow path as focusing the laser onto the window damages the capillary. The pulse energy was nominally 60 µJ and the repetition rate was 70 Hz but there was pulse to pulse variability in the excitation energy and period which did not detract from the findings.

The intrinsic fluorescence from the protein was reflected by an ellipsoidal reflector (Newport, Ellip-2) towards a > 2.0 GHz ultrafast photodiode (Alphalas, UPD-200-UP) through a 300 nm long-pass filter (Semrock, FF01-300/LP-25) and then a 400 nm short-pass filter (Edmund Optics, 84-702) to block light from the UV laser at 266 nm and longer wavelengths. The current

output from the ultrafast photodiode was amplified by a 2.2 GHz transimpedance amplifier (Femto, HSA-X-I-2-40) to generate a voltage which fed a 12 GHz sampling oscilloscope (Pico Technology, PicoScope 9211A). This created a digital representation of the intrinsic fluorescence lifetime of the proteins with a time resolution of 120 ps and voltage resolution of 16 bit from 256 excitation pulses in typically less than 5 s. The sampling oscilloscope was triggered by the output from a photodiode (Thorlabs, DET10A/M) aligned to a partial reflection of the excitation pulse from a beam-sampling window (Thorlabs, BSF10-UV).

LabVIEW 2015 (National Instruments) was used to control the experimental equipment and provide real-time analysis of the variations in fluorescence intensity and lifetime. For each experiment, a new fluorescence intensity and lifetime measurement would begin as soon as the previous measurement finished, as the sample was continuously eluted from the chromatographic column. The width of the elution peak was typically much wider than the ~ 31 µL eluted between measurements. Each lifetime measurement was tagged with a record of the chromatographic elution time at which the measurement was taken.

3. Calculation

We developed an analytical framework for extracting information on-the-fly from the data regarding the identity and quantity of co-eluting proteins, which does not require the global analysis of an emission decay surface.[24] In this framework, the input to the system is the excitation pulse and the output is the digitised signal recorded by the sampling oscilloscope. The system is assumed to be linear: increasing the intensity of the input pulse would increase the amplitude of the digitised signal by the same factor, and the signals resulting from adding additional inputs would not interact with each other.[25] Consequently, the output is the convolution of the input and the impulse response of the system, which is the convolution of the impulse responses of the sample and the detection electronics. The impulse response of the de-

tection electronics can be determined beforehand, and the impulse response of the sample contains information on both the quantity and identity of the proteins that have been excited by the input pulse, which is the relevant information.

Since convolution is commutative, the digital signal $I(t)$ can be modelled as

(1) $$I(t) = \int_{-\infty}^{\infty} R(t - t')\, S(t')\, dt'$$

where $R(t)$ is the convolution of the excitation pulse and the impulse response of the detection electronics, and $S(t)$ is the impulse response of the sample. $R(t)$ can be determined experimentally by measuring the time-profile of the excitation pulse using the same optical configuration and detection electronics (Supplementary Figure 1). Since the filters positioned before the ultrafast photodiode block the excitation light from entering the ultrafast photodiode, the filter assembly would be substituted for a neutral density filter (Thorlabs, NDUV30A) to enable enough light to impinge on the ultrafast photodiode for the signal to be detectable. $R(t)$ was modelled by a function consisting of the sum of two Gaussian functions

(2) $$R(t) = \sum_{i=1}^{2} \frac{\alpha_i}{\sigma_i \sqrt{2\pi}} e^{-\frac{(t-\varphi_i)^2}{2\sigma_i^2}}$$

where α is the amplitude, σ is the standard deviation and φ is the temporal offset of the digital representation of the excitation pulse.

$S(t)$ was modelled by a function consisting of one exponential decay or a sum of two exponential decays. These models contain one or two decay components, respectively.

(3) $$S(t) = \begin{cases} 0 & if\ t < 0 \\ \frac{\beta_1}{\tau_1} e^{-\frac{t}{\tau_1}} + \frac{\beta_2}{\tau_2} e^{-\frac{t}{\tau_2}} & if\ t \geq 0 \end{cases}$$

where β_j where $j = 1, 2$ is the pre-exponential factor and τ_j is the fluorescence decay time for each exponential decay component. From (1), the digital signal can then be modelled as

$$(4) \qquad I(t) = \frac{1}{2} \sum_{i=1}^{2} \alpha_i \sum_{j=1}^{2} \left\{ \frac{\beta_j}{\tau_j} e^{\frac{\sigma_i^2 - 2\tau_j(t-\varphi_i)}{2\tau_j^2}} \left[1 - \operatorname{erf}\left(\frac{\sigma_i^2 - \tau_j(t-\varphi_i)}{\sqrt{2}\,\sigma_i\,\tau_j} \right) \right] \right\}$$

If the second pre-exponential factor is set to $\beta_2 = 0$, then $S(t)$ is modelled by a function consisting of just one decay component, and the digital signal can then be modelled as,

$$(5) \qquad I(t) = \frac{1}{2} \sum_{i=1}^{2} \frac{\alpha_i\,\beta_1}{\tau_1} e^{\frac{\sigma_i^2 - 2\tau_1(t-\varphi_i)}{2\tau_1^2}} \left[1 - \operatorname{erf}\left(\frac{\sigma_i^2 - \tau_1(t-\varphi_i)}{\sqrt{2}\,\sigma_i\,\tau_1} \right) \right]$$

If the impulse response of each protein species in a sample can be modelled as a sum of two exponential decays, then $S(t)$ can be modelled as,

$$(6)\ S(t) = \begin{cases} 0 & if\ t < 0 \\ \sum_{j=1}^{\mathrm{N}} \gamma_j \left(\frac{\beta_j}{\tau_{j1}} e^{-\frac{t}{\tau_{j1}}} + \frac{1-\beta_j}{\tau_{j2}} e^{-\frac{t}{\tau_{j2}}} \right) & if\ t \geq 0 \end{cases}$$

where $\beta = \frac{\beta_{j1}}{\beta_{j1}+\beta_{j2}}$ corresponds to the contribution of the first decay component and $\gamma_j = \beta_{j1} + \beta_{j2}$ is the sum of the pre-exponential factors of the j^{th} protein species being addressed by the excitation pulse in the capillary. In this notation, β is a property of the protein species being addressed and γ is proportional to the quantity of that species. When a single exponential decay model is applied, $\beta = 1$. Each protein species has characteristic τ_1, τ_2 and β that can be used to identify that species as it is eluted from the chromatography column (Figure 2).

Consequently, the digital signal for N proteins species can be modelled as,

$$(7) \qquad I(t) = \int_0^{\infty} \sum_{i=1}^{2} \frac{\alpha_i}{\sigma_i \sqrt{2\pi}} e^{-\frac{(t-t'-\varphi_i)^2}{2\,\sigma_i^2}} \sum_{j=1}^{\mathrm{N}} \gamma_j \left(\frac{\beta_j}{\tau_{j1}} e^{-\frac{t'}{\tau_{j1}}} + \frac{1-\beta_j}{\tau_{j2}} e^{-\frac{t'}{\tau_{j2}}} \right) dt'$$

Since convolution is distributive, this can be solved:

(8) $$I(t) = \frac{1}{2}\sum_{i=1}^{2}\alpha_i \sum_{j=1}^{N}\gamma_j \left\{ \frac{\beta_j}{\tau_{j1}}\, e^{\frac{\sigma_i^2 - 2\tau_{j1}(t-\varphi_i)}{2\tau_{j1}^2}} \left[1 - \operatorname{erf}\left(\frac{\sigma_i^2 - \tau_{j1}(t-\varphi_i)}{\sqrt{2}\,\sigma_i\,\tau_{j1}}\right)\right] + \frac{1-\beta_j}{\tau_{j2}}\, e^{\frac{\sigma_i^2 - 2\tau_{j2}(t-\varphi_i)}{2\tau_{j2}^2}} \left[1 - \operatorname{erf}\left(\frac{\sigma_i^2 - \tau_{j2}(t-\varphi_i)}{\sqrt{2}\,\sigma_i\,\tau_{j2}}\right)\right] \right\}$$

The fluorescence intensity is the integral of the digital signal plus a baseline offset c which accounts for electrical noise. The baseline was calculated by taking the median of the first 13 points in the digital signal, which corresponds to a region of time before the excitation pulse has arrived at the capillary. The fluorescence intensity was calculated by numerically integrating the signal using the trapezoid rule. The contribution of the j^{th} protein species $I_j(t)$ to the fluorescence intensity was calculated from the fitted parameters by noting that

(9) $$I_j(t) = \int_{-\infty}^{\infty} R(t) * S_j(t)\, dt = \int_{-\infty}^{\infty} R(t)\, dt \int_{-\infty}^{\infty} S_j(t)\, dt$$

Therefore, if the time window of the sampling oscilloscope were infinitely wide, then

(10) $$I_j(t) = \gamma_j \sum_{i=1}^{2} \alpha_i$$

Of course, if the time window of the sampling oscilloscope were infinite, then the contribution of the background $I_{\text{background}}(t)$ would tend to infinity, since

(11) $$I_{\text{background}}(t) = \int_{-\infty}^{\infty} c\, dt \rightarrow \infty$$

If the width of the time window of the sampling oscilloscope is given by θ, then,

(12) $$I_j(t) = \int_0^{\theta} R(t) * S_j(t)\, dt = \int_0^{\theta}\int_{-\infty}^{\infty} R(t-t')\, S(t')\, dt'\, dt$$

This can be solved by approximating the shape of the excitation pulse as a delta function:

(13) $R(t) \sim \sum_{i=1}^{2} \alpha_i \, \delta(t - \varphi_i)$

This assumption is valid because the width of the excitation pulse is much narrower than the width of the time window. The contribution of the j^{th} protein species $I_j(t)$ to the fluorescence intensity measured across the time window is then

(14) $I_j(t) \sim \sum_{i=1}^{2} \alpha_i \, \gamma_j \left[\beta_j \left(1 - e^{\frac{\varphi_i - \theta}{\tau_{j1}}}\right) + (1 - \beta_j)\left(1 - e^{\frac{\varphi_i - \theta}{\tau_{j2}}}\right)\right]$

However, if the size of the time window is chosen such that $\theta \gg \tau_{j1}, \tau_{j2}, \varphi_i$, then

(15) $I_j(t) \sim \gamma_j \sum_{i=1}^{2} \alpha_i$

The contribution of the background is

(16) $I_{\text{background}}(t) = c\,\theta$

This analytical framework provides the foundation for processing the measured signal to identify and quantify the eluted proteins in real time.

4. Results and Discussion

To evaluate the system and the analytical framework, BSA and ovalbumin were eluted separately from a SEC column, and a series of fluorescence time decays were acquired which were then analysed in real-time (Figure 3) to generate a series of chromatograms.

The total fluorescence chromatogram (TFC) traces the fluorescence intensity against elution volume (Figure 3a). The fluorescence intensity is related to the total quantity of protein being addressed by the excitation pulse and depends on the extinction coefficients and fluorescence yields of the proteins. TFCs are similar to univariate UV-Vis absorption chromatograms. The fluorescence intensity was calculated as the integral of the digital signal, and the

elution volume was the time elapsed between the acquisition of the measurement and the start of the chromatographic run multiplied by the flow rate.

The decay chromatogram (DC) traces the fluorescence decay times τ_1 and τ_2 against elution volume. The decay times are related to the identity of the protein being addressed by the excitation pulse, and not to the quantity of protein. The DC was generated by fitting lifetime measurements and plotting the fitted decay times against the elution volume. DC-1s and DC-2s employ models with either one or two decay components, respectively. For on-the-fly analysis, DCs were prepared by fitting the lifetime data using either Eq. (5) or Eq. (4) depending on whether a DC-1 or DC-2 was requested and fits were only attempted if there was sufficient signal. In this example, two decay components were used to fit the data generating DC-2 chromatograms consisting of decay times τ_1 (Figure 3b) and τ_2 (Figure 3c) plotted against elution volume. See Supplementary Figure 2 for examples of on-the-fly model fits of lifetime data.

The data analysis was streamlined and automated so that, in the example shown, the TFC and DC-2 were generated in real-time as the proteins were eluted from the column. From inspection of the DC-2, it is clear that BSA and ovalbumin have contrasting fluorescence lifetime profiles which can be used to identify that the leading peak is BSA and the lagging peak is ovalbumin, which could not be determined from inspecting the TFC alone.

The resolving power of the TRF chromatograph was evaluated by measuring the eluate when BSA and ovalbumin were deliberately co-eluted from a SEC column. The elution volume of each protein species was controlled by injecting them into the column at different times, with a fixed flow rate of 0.4 mL min^{-1}. We subjected the experimental data to analysis offline but there is no technical reason why this same analysis could not also be performed on-the-fly.

As for the on-the-fly analysis, DC-1s and DC-2s were prepared by fitting the lifetime data using either Eq. (5) or Eq. (4) but fits were attempted for every lifetime measurement even when there was very low or unmeasurable signal. For DC-2s the weighting β of the first decay can be plotted against elution volume to generate a first-decay component chromatogram (FDCC).

When BSA was injected 4 mL before ovalbumin, the TFC contained two distinct peaks preceded by a smaller peak in each case (Figure 4a) which could be identified as BSA or ovalbumin from the accompanying DC-1 trace, which contained two flat traces at the characteristic decay time of each protein species (Figure 4b). The DC-1 trace was flat across the elution volume for each TFC peak, indicating that each peak contained pure protein. The widths of the elution peaks in the TFCs resulted in loss of base line separation when the volume separating the injection of each protein species was reduced to 2.0 mL. The DC-1 traces transitioned between the two characteristic decay constants, appearing sloped where the TFC peaks partially overlapped, indicating that mixtures of the two proteins were being eluted. The DC-1 traces converged to flat lines when the peaks ceased to overlap, demonstrating that the DC-1 traces can be used to assess protein purity.

When BSA was injected 0.8 mL before ovalbumin, the two protein species co-eluted almost exactly, such that only a single peak was observed in the TFC. The DC-1 trace contained a sharp change in the decay time, which indicated that the single TFC peak did not contain a single, pure protein species, and that it contained at least two poorly-resolved proteins.

The goodness of the fit to lifetime measurements is typically improved by using a two-decay component model instead of a single-decay component model, so the combination of the fluorescence decay times τ_1 and τ_2 (Figure 4c and d) alongside the weighting β of the first decay

component (Figure 4e) can be used to characterise the time-profile of the intrinsic fluorescence of different protein species. The parameters τ_1, τ_2, and β can be used to uniquely characterise different protein species, as they are dependent on only the optical properties of the aromatic residues within their particular protein environments. This can be utilised to construct decay-associated chromatograms (DACs) which trace the contribution of each of j protein species to the TFC. DACs require prior information on the τ_{j1}, τ_{j2} and β_j parameters for each of the j protein species eluted from the chromatographic column, but do not make any assumptions about the elution profile of each species. DACs were prepared by fitting the lifetime measurements with Eq. (8) with β_j, τ_{j1} and τ_{j2} fixed, and the parameters γ_j free to vary. The contribution of the j^{th} protein species could then be calculated using Eq. (14), which was then plotted against elution volume.

The data set from the experiment described in Figure 3 was re-analysed to determine the τ_1, τ_2 and β parameters for BSA and ovalbumin since there was no overlap in their elution profiles in this run (Supplementary Figure 3). The parameters used for the DAC analysis were the fit parameters for the decay measurements closest to the peak positions in the TFC, which were determined by fitting the TFC using an exponentially modified Gaussian function. The numbers input were, for BSA: $\tau_1 = 0.520327$ ns, $\tau_2 = 7.10122$ ns, and $\beta = 0.03012$; and for ovalbumin: $\tau_1 = 0.900353$ ns, $\tau_2 = 5.65585$ ns, and $\beta = 0.17753$. The decay-associated chromatogram (DAC) shows the variation in the total fluorescence intensity associated with each individual protein species, as a function of the elution volume (Figure 4f). In this example, the DAC was calculated after all of the decay curves had been recorded, but if the characteristic parameters of the protein species were known in advance, by measuring standard materials for example, then the DAC could also be generated in real-time.

The DAC trace, which represents the amplitudes for each decay in the signal, was able to resolve the two proteins regardless of whether there was full, partial or no overlap between the TFC peaks, providing in fine detail, the elution profile of each protein species that contributed to the overall elution profile shown in the TFC.

There were secondary peaks in the elution profiles of both individual protein preparations, which were most likely non-covalently formed dimers known to occur in both BSA and ovalbumin.[26,27] These species could be monitored accurately with DAC but not by DC or TFC. For instance, when BSA was injected 0.8 mL after ovalbumin, the DAC trace shows that the secondary peak of ovalbumin fully co-elutes with BSA (Figure 4f, column 5), whereas this could not be deduced from the TFC or DC traces.

We tested whether DAC could accurately determine the concentration of each co-eluting protein species by injecting different concentrations of BSA and ovalbumin into the chromatograph, such that they always fully co-eluted from the SEC column. The same effect could have been achieved by injecting the samples at the same time, but disconnecting the chromatography column. For each set of injections, only a single primary peak was visible in the TFC (Figure 5a and c) but the elution of each protein species could be tracked in the accompanying DAC (Figure 5b and d).

Since the DAC traces deconvolve the TFC into separate elution profiles for each protein species without making assumptions about the elution profile, it is possible to further fit the DACs using peak shape functions to quantify each protein species. The DAC peaks were fitted using an exponentially modified Gaussian function (Supplementary Figure 4). The quantity of each protein species is proportional to the area under the peak, and varied linearly with the protein concentration injected (Figure 6). The gradient depends on the number of aromatic residues per protein and their quantum yield and extinction coefficient with pumping at 266

nm, and this can be calibrated to report on actual protein concentration using the same standard samples when determining the τ_1, τ_2 and β parameters.

For PAT applications using SEC, protein fragments and aggregates would need to be tracked at concentrations that are 1 % to 5 % of the product, and ideally lower. This study demonstrated 10 % and with further technical development we anticipate that even lower levels could be achieved.

5. Conclusions

TRF can be used to monitor the elution of proteins during chromatography and as a standalone technique for quantifying proteins in mixtures. By generating DACs in real-time, the elution profile of two different protein species can be monitored independently, even when they fully or partially co-elute. This overcomes a major shortcoming of chromatograms measured by UV-absorption or intrinsic fluorescence intensity, as these can only monitor the total amount of protein eluted. Whilst the capital cost of a TRF chromatograph would be greater than for a UV-absorption chromatograph due to the UV laser and the high bandwidth electronics, the cost of operation would be comparable. TRF-based chromatography is an emerging technology for monitoring the product concentration in continuous manufacturing processes, or for making accurate peak-cutting and fraction pooling decisions in batch purification processes. Future work will include a demonstration of the utility of TRF-based chromatography in more complex systems, and in other modes of chromatography, such as protein A affinity chromatography.

Author statement

John E. Hales: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization.

Samir Aoudjane: Conceptualization, Methodology, Validation, Investigation, Writing - Review & Editing. Gabriel Aeppli: Conceptualization, Writing - Review & Editing, 4 Funding acquisition. Paul A. Dalby: Conceptualization, Methodology, Validation, Writing - Review & Editing, Supervision, Project administration, Funding acquisition.

Acknowledgements

This work was supported by the Bioprocessing Research Industry Club (BBSRC, BB/E005942/1), and EPSRC EP/N025105/1. The funding sources had no involvement in the study design, in the collection, analysis and interpretation of data, in the writing of the report, or in the decision to submit the article for publication.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: J.E.H., S.A., G.A., and P.A.D. are co-authors on a patent application which could increase in value if the methods and ideas described in this paper find widespread application.

References

1. A. Jungbauer, Continuous downstream processing of biopharmaceuticals, Trends Biotechnol. 31 (2013) 479–492. https://doi.org/10.1016/j.tibtech.2013.05.011

2. M.C. Manning, K. Patel, R.T. Borchardt, Stability of protein pharmaceuticals, Pharm. Res. 6 (1989) 903–918. https://doi.org/10.1023/A:1015929109894

3. S. Arase, K. Horie, T. Kato, A. Noda, Y. Mito, M. Takahashi, T. Yanagisawa, Intelligent peak deconvolution through in-depth study of the data matrix from liquid chromatography

coupled with a photo-diode array detector applied to pharmaceutical analysis, J. Chromatogr. A 1469 (2016) 35–47. https://doi.org/10.1016/j.chroma.2016.09.037

4. N. Brestrich, A. Sanden, A. Kraft, K. McCann, J. Bertolini, J. Hubbuch, Advances in inline quantification of co-eluting proteins in chromatography: Process-data-based model calibration and application towards real-life separation issues, Biotechnol. Bioeng. 112 (2015) 1406–1416. https://doi.org/10.1002/bit.25546

5. N. Brestrich, M. Rüdt, D. Büchler, J. Hubbuch, Selective protein quantification for preparative chromatography using variable pathlength UV/Vis spectroscopy and partial least squares regression, Chem. Eng. Sci. 176 (2018) 157–164. https://doi.org/10.1016/j.ces.2017.10.030

6. T. Hahn, T. Huuk, A. Osberghaus, K. Doninger, S. Nath, S. Hepbildikler, V. Heuveline, J. Hubbuch, Calibration-free inverse modeling of ion-exchange chromatography in industrial antibody purification, Eng. Life Sci. 16 (2016) 107–113. https://doi.org/10.1002/elsc.201400248

7. T. Hahn, P. Baumann, T. Huuk, V. Heuveline, J. Hubbuch, UV absorption-based inverse modeling of protein chromatography, Eng. Life Sci. 16 (2016) 99–106. https://doi.org/10.1002/elsc.201400247

8. N. Field, S. Konstantinidis, A. Velayudhan, High-throughput investigation of single and binary protein adsorption isotherms in anion exchange chromatography employing multivariate analysis, J. Chromatogr. A 1510 (2017) 13–24. https://doi.org/10.1016/j.chroma.2017.06.012

9. J.M. Beechem, L. Brand, Time-resolved fluorescence of proteins, Annu. Rev. Biochem. 54 (1985) 43–71. https://doi.org/10.1146/annurev.bi.54.070185.000355

10. K. Sudhakar, W.W. Wright, S.A. Williams, C.M. Phillips, J.M. Vanderkooi, Phenylalanine fluorescence and phosphorescence used as a probe of conformation for cod parvalbumin, J. Fluoresc. 3 (1993) 57–64. https://doi.org/10.1007/BF00865318

11. I. Gryczynski, M. Eftink, J.R. Lakowicz, Conformation heterogeneity in proteins as an origin of heterogeneous fluorescence decays, illustrated by native and denatured ribonuclease T_1, Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol. 954 (1988) 244–252. https://doi.org/10.1016/0167-4838(88)90079-9

12. B.C.R. Zhu, R.A. Laine, M.D. Barkley, Intrinsic tryptophan fluorescence measurements suggest that polylactosaminyl glycosylation affects the protein conformation of the gelatin-binding domain from human placental fibronectin, Eur. J. Biochem. 189 (1990) 509–516. https://doi.org/10.1111/j.1432-1033.1990.tb15516.x

13. D.J. Desilets, P.T. Kissinger, F.E. Lytle, Measurement of fluorescence lifetimes during liquid chromatography, Anal. Chem. 59 (1987) 1830–1834. https://doi.org/10.1021/ac00141a020

14. S. Landgraf, Time-resolved fluorescence HPLC detection using semiconductor light sources: principles and applications, Chem. Eng. Technol. 39 (2016) 175–182. https://doi.org/10.1002/ceat.201500152

15. W.T. Cobb, L.B. McGown, On-line fluorescence lifetime detection for chromatographic peak resolution, Anal. Chem. 62 (1990) 186–189. https://doi.org/10.1021/ac00201a019

16. M.A. Dvorak, G.A. Oswald, M.H. Van Benthem, G.D. Gillispie, On-the-fly fluorescence lifetime determination with total emission detection in HPLC, Anal. Chem. 69 (1997) 3458–3464. https://doi.org/10.1021/ac961095j

17. M.B. Smalley, J.M. Shaver, L.B. McGown, On-the-fly fluorescence lifetime detection in HPLC using a multiharmonic Fourier transform phase-modulation spectrofluorometer, Anal. Chem. 65 (1993) 3466–3472. https://doi.org/10.1021/ac00071a022

18. L.-C. Li, L.B. McGown, On-the-fly frequency-domain fluorescence lifetime detection in capillary electrophoresis, Anal. Chem. 68 (1996) 2737–2743. https://doi.org/10.1021/ac960160m

19. K.J. Miller, F.E. Lytle, Capillary zone electrophoresis with time-resolved fluorescence detection using a diode-pumped solid-state laser, J. Chromatogr. A 648 (1993) 245–250. https://doi.org/10.1016/0021-9673(93)83307-E

20. J. Arlt, D. Tyndall, B.R. Rae, D.D.-U. Li, J.A. Richardson, R.K. Henderson, 2013. A study of pile-up in integrated time-correlated single photon counting systems. Rev. Sci. Instrum. 84, 103105. https://doi.org/10.1063/1.4824196

21. O.S. Degenhardt, B. Waters, A. Rebelo-Cameirao, A. Meyer, H. Brunner, N.P. Toltl, 2004. Comparison of the effectiveness of various deaeration techniques, Dissolut. Technol. 11. https://doi.org/10.14227/DT110104P6

22. A. Bujacz, Structures of bovine, equine and leporine serum albumin, Acta Crystallogr. D Struct. Biol. 68 (2012) 1278–1289. https://doi.org/10.1107/S0907444912027047

23. P.E. Stein, A.G.W. Leslie, J.T. Finch, R.W. Carrell, Crystal structure of uncleaved ovalbumin at 1·95 Å resolution, J. Mol. Biol. 221 (1991) 941–959. https://doi.org/10.1016/0022-2836(91)80185-W

24. M. Maeder, K.J. Molloy, Lifetime resolved fluorescence detection in chromatography. Global analysis of emission decay surfaces, Chemometr. Intell. Lab. Syst. 28 (1995) 99–107. https://doi.org/10.1016/0169-7439(95)80043-9

25. S.W. Smith, Digital Signal Processing, Newnes, Boston, 2003.

26. P.G. Squire, P. Moser, C.T. O'Konski, The hydrodynamic properties of bovine serum albumin monomer and dimer, Biochemistry 7 (1968) 4261–4272. https://doi.org/10.1021/bi00852a018

27. A.A. Ansari, A. Salahuddin, Purification and some properties of rabbit anti-ovalbumin, Biochem. J. 135 (1973) 705–711. https://doi.org/10.1042/bj1350705

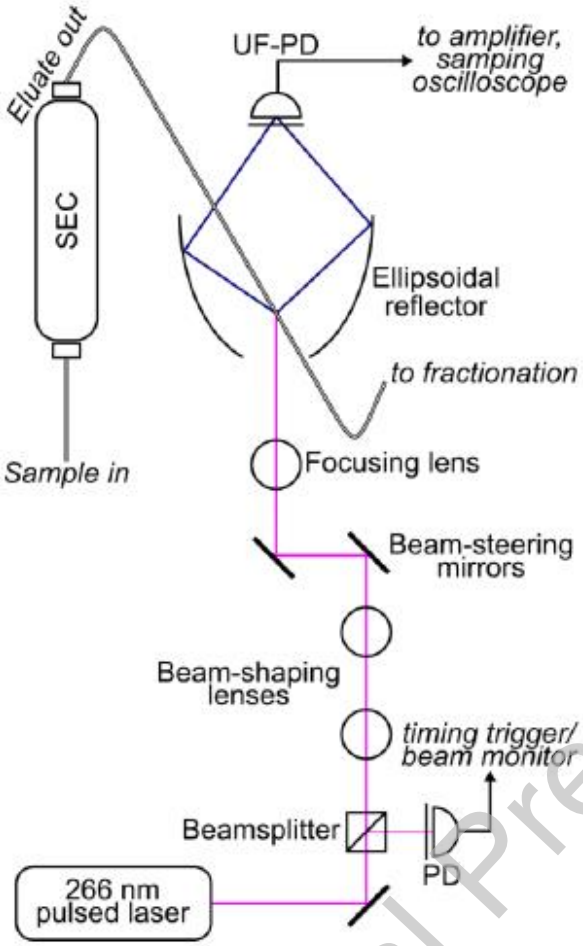

Figure 1. Design of the time resolved intrinsic fluorescence lifetime chromatograph. The magenta line indicates the beam path of the 266 nm nanosecond laser pulses towards a capillary through which the eluate from a size exclusion chromatographic column is flowed. The navy line represents the intrinsic fluorescence emission, which is reflected by the ellipsoidal reflector towards an ultra-fast photodiode connected to an electronic sub-assembly. The resulting voltage is then recorded and the signal is processed to determine the identity and quantity of the eluted proteins.

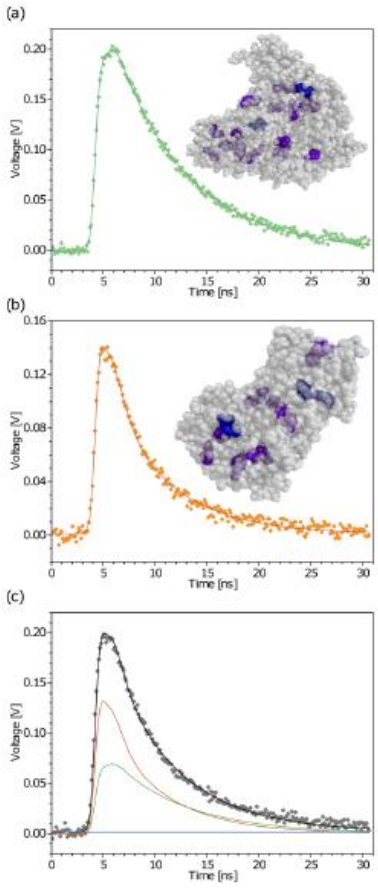

Figure 2. Examples of the distinct digital signals (scatter points) generated when (a) BSA or (b) ovalbumin are eluted, flowed into the detection volume, and then data fitted (lines) to determine the characteristic decay components and their respective contributions. Insets, atomic structures of BSA (PDB: 4F5S)[22] and ovalbumin (PDB: 1OVA)[23] where the tryptophan residues have been coloured blue and the tyrosine residues have been coloured violet. (c) Example of a digital signal (scatter points) generated from a mixture of BSA and ovalbumin. Knowledge of the characteristic decay signature of each protein is used to deconvolve the signal into the contributions made by BSA (green) and ovalbumin (dark orange), and the fitted curve (black) is the sum of these contributions and the baseline (dark blue).

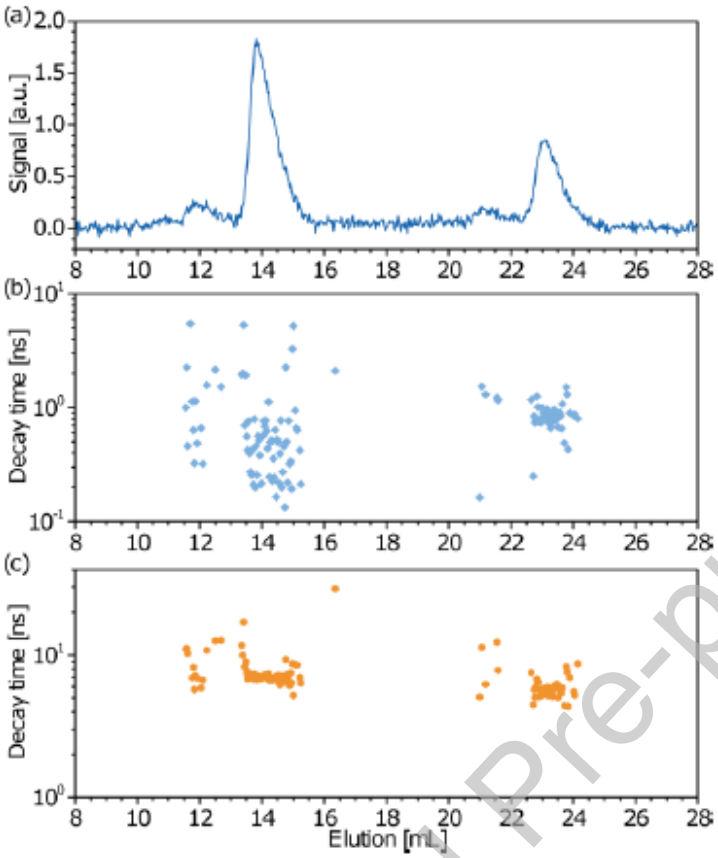

Figure 3. Real-time time resolved intrinsic fluorescence lifetime chromatography of BSA and ovalbumin separated on a size-exclusion chromatography column. BSA and ovalbumin were injected separately to increase the separation of the elution peaks by 8 mL. (a) The TFC reports the quantity of protein – similar to a univariate UV absorption measurement – and cannot distinguish between different proteins species. The decay times (b) τ_1 and (c) τ_2 from the on-the-fly decay chromatogram analysis using two decay components are different for each protein species so the lead and lag peaks can be identified as BSA and ovalbumin, respectively.

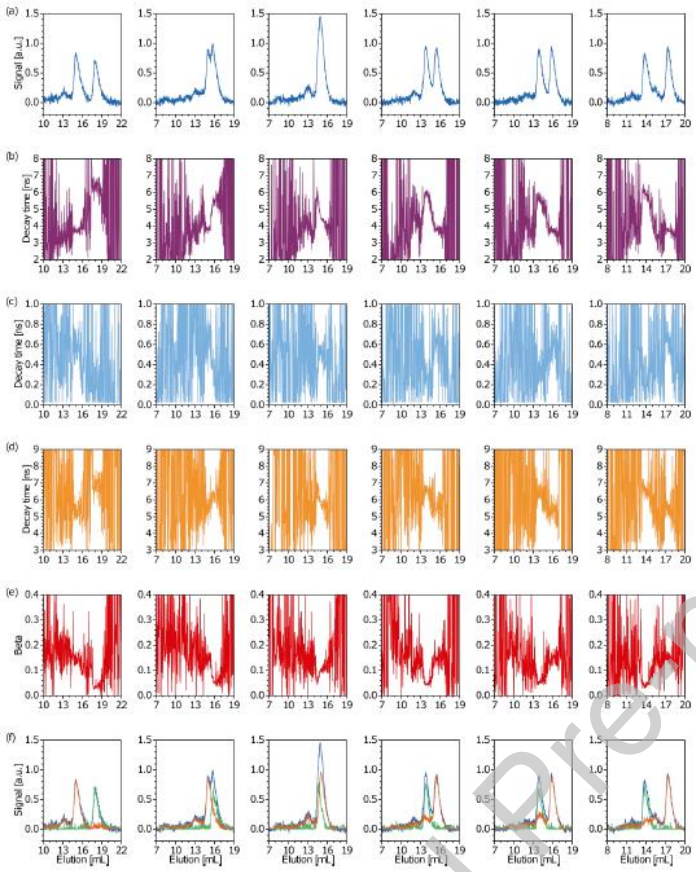

Figure 4. Quantification of partially co-eluting proteins species. BSA and ovalbumin were separately injected into the SEC column at different times to control the degree of overlap of their elution profiles. From left to right, the BSA injected 4 mL, 2 mL and 0.8 mL after ovalbumin, and then ovalbumin injected 0.4 mL, 0.8 mL and 2.4 mL after BSA. The flow rate was 0.4 mL/min. The full set of chromatograms have been reported, including the (a) TFC, (b) DC-1, (c) τ_1, (d) τ_2, (e) FDCC, and (f) DAC for each experiment. For the DACs, the dark blue trace represents the TFC minus the background contribution as determined by Eq. (16), and the green and dark orange traces represent the elution profiles of BSA and ovalbumin, respectively.

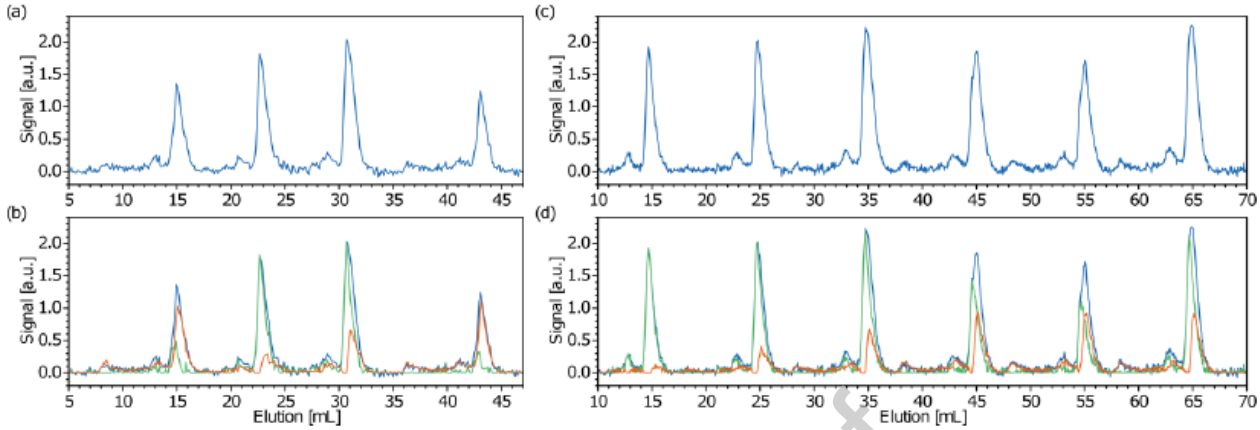

Figure 5. Quantification of co-eluting protein species. In two experiments, a series of BSA and ovalbumin samples at different concentrations were injected into the SEC column such that the BSA and ovalbumin elution profiles were fully overlapping. In the TFCs of the (a) first and (c) second experiment, the separate elution profiles are indistinguishable, but in the DACs (b) and (d) traces for BSA (green) and ovalbumin (dark orange) can be deconvolved. In the first experiment from left to right, the BSA concentrations were 1 mg/mL, 5 mg/mL, 5 mg/mL and 0.5 mg/mL, and the ovalbumin concentrations were 5 mg/mL, 1 mg/mL, 3 mg/mL and 5 mg/mL. For the second experiment, the BSA concentrations were 5 mg/mL, 5 mg/mL, 5 mg/mL, 4 mg/mL, 3 mg/mL and 5 mg/mL, and the ovalbumin concentrations were 0.5 mg/mL, 2 mg/mL, 4 mg/mL, 5 mg/mL, 5 mg/mL and 5 mg/mL.

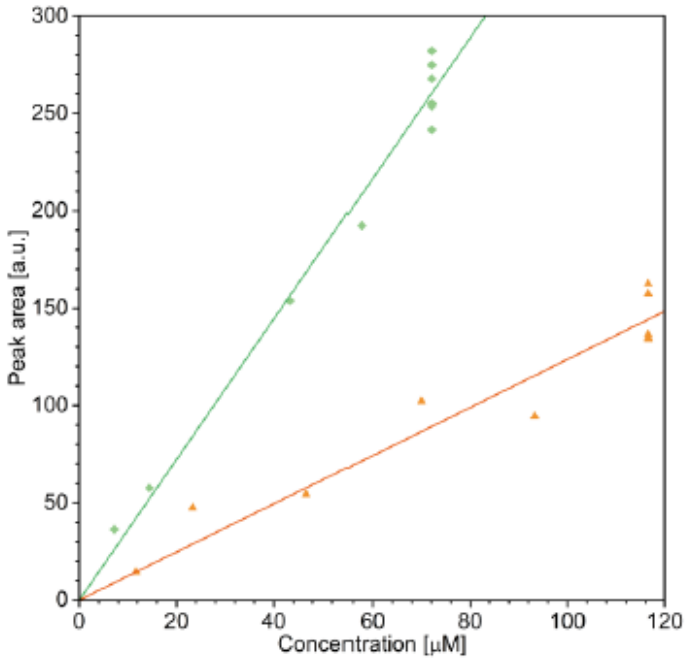

Figure 6. Time resolved intrinsic fluorescence lifetime chromatography as a PAT. The area under each of the primary peaks for the DAC traces for BSA (light green scatter points) and ovalbumin (orange scatter points) in Figure 5 have been plotted against the known concentration of the injected sample. Linear fits for the BSA (green line) and ovalbumin (dark orange) data show that the peak area is proportional to the concentration of each protein species.