University of Lethbridge

US-SOMO HPLC-SAXS module: dealing with capillary fouling and extraction of pure component patterns from poorly resolved SEC-SAXS data

Emre Brookes, Patrice Vachette, Mattia Roccoc and Javier Pe´rezd

Presented by: Zachary Robinson

Overview

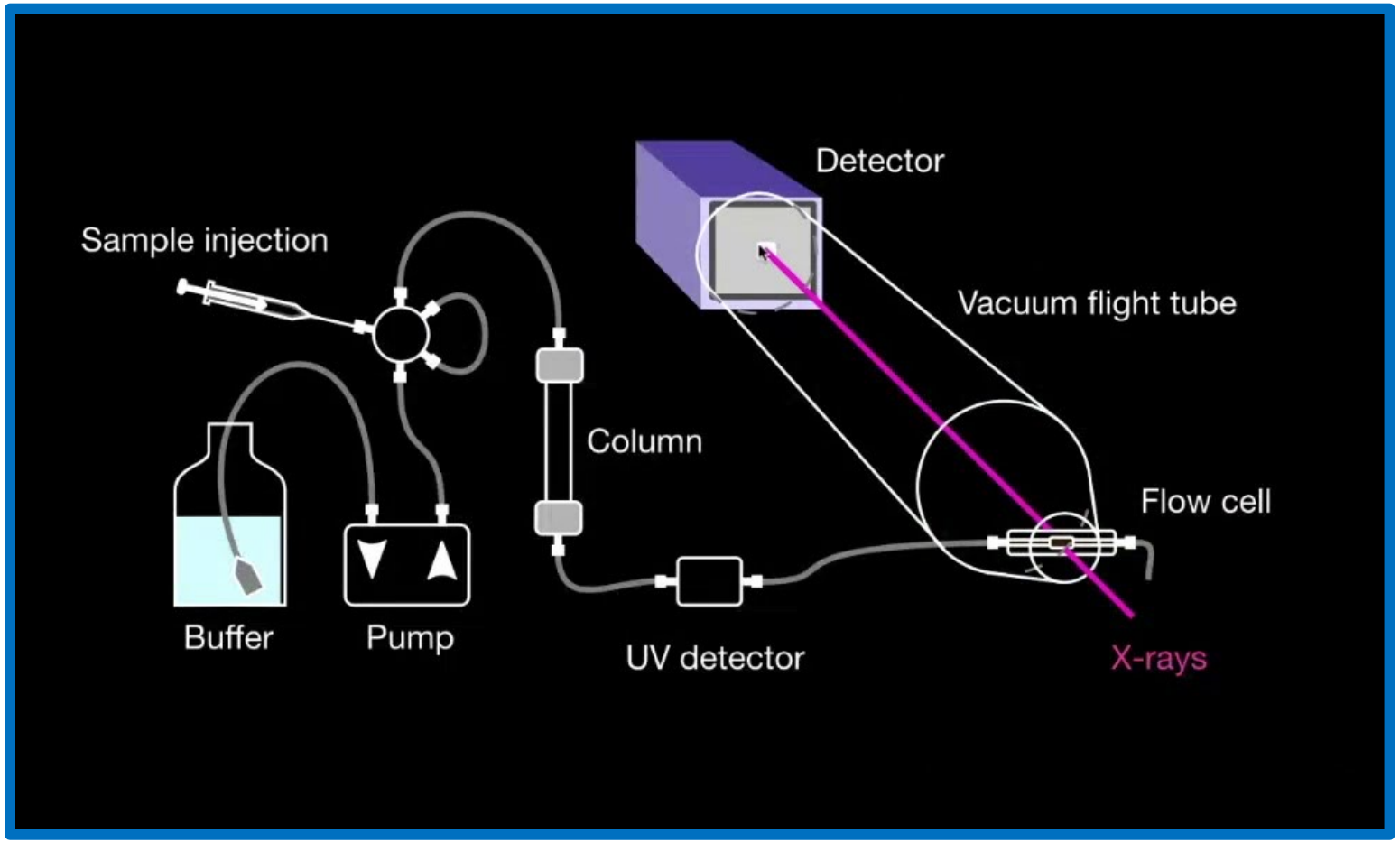
University of Lethbridge

- SAXS Background
- Capillary Fouling
- Proposed Solutions
- Poor SEC Resolution
- Proposed Solutions
- Future Directions

SEC-SAXS

University of Lethbridge

- Sample is pushed through a size-exclusion column to separate out different species before it is passed through a capillary for SAXS measurements to be taken.



Size Exclusion Chromatography coupled SAXS (SEC-SAXS), (BioCAT, 2021)

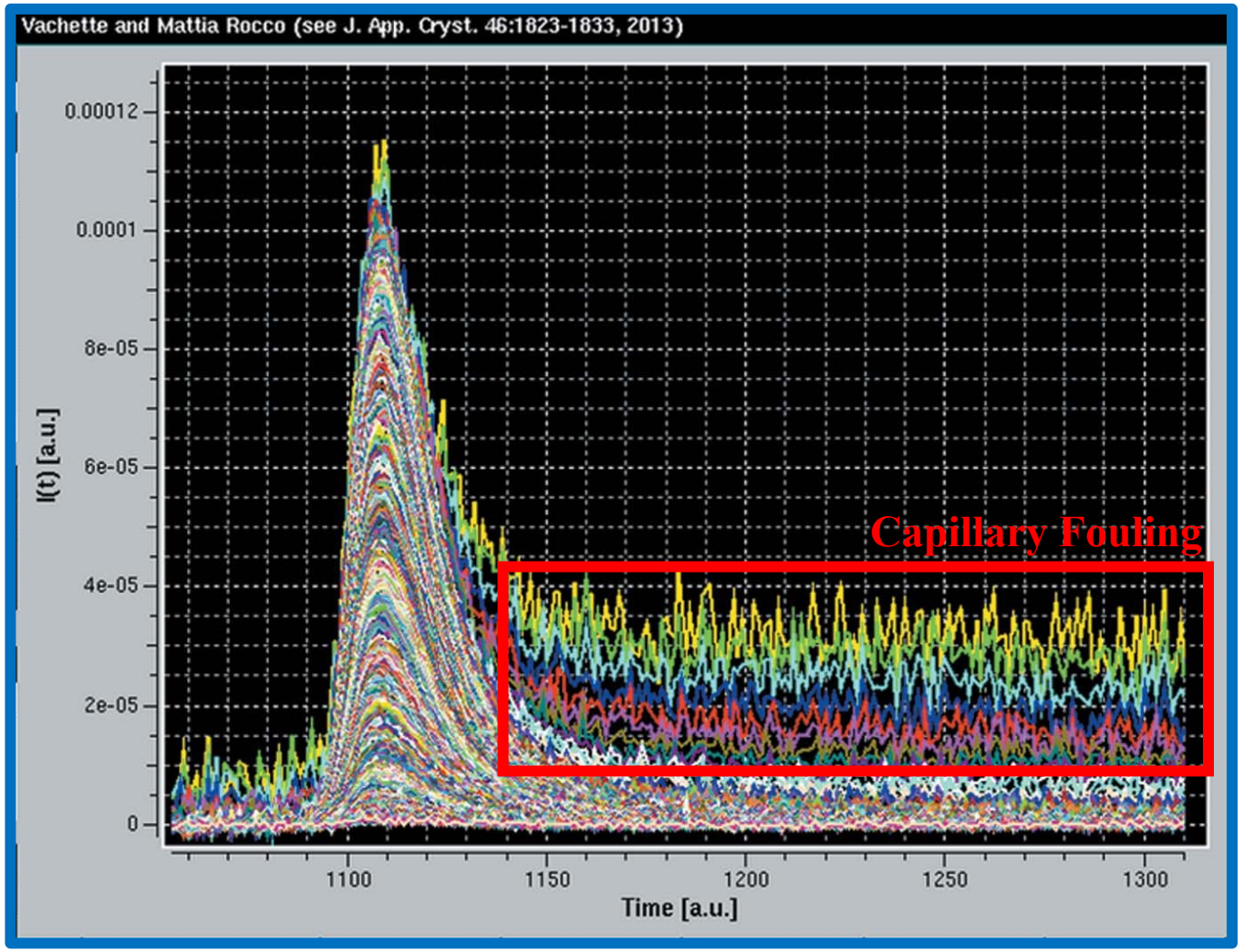
US SOMO HPLC SAXS Module

University of Lethbridge

FIAT LUX

Step 1

- Plot m$I_q(t)$ vs t
- t corresponds to time
- m corresponds to the number of q values at that time
- $I_q(t)$ is scattering intensity as a function of time.
- **This plot allows for detection of capillary fouling.**

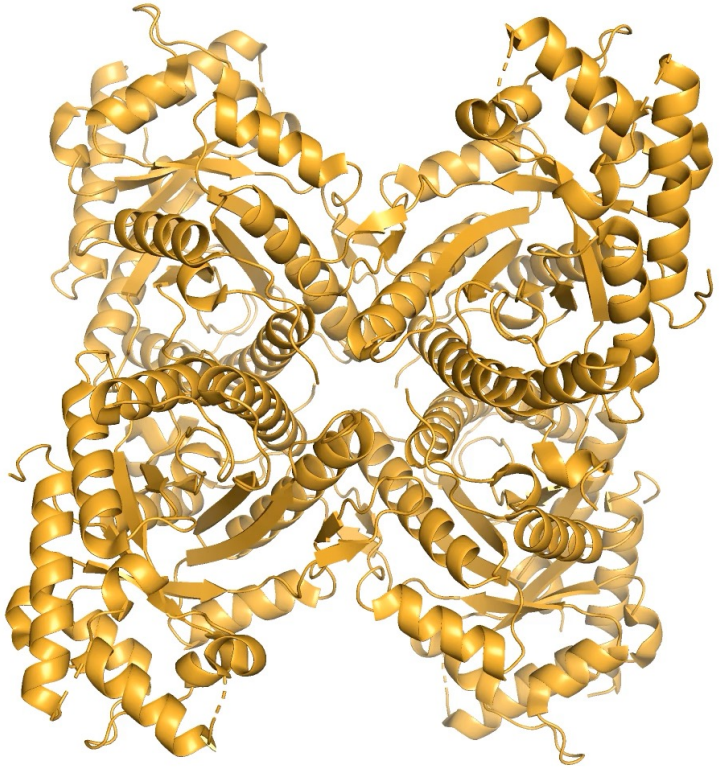


Brookes et al., 2016

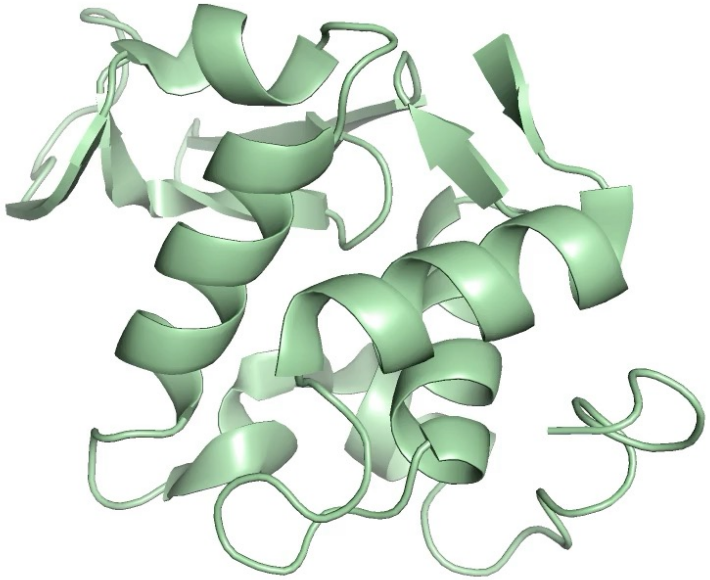
Experimental Methods

University of Lethbridge

- Buffer conditions previously shown to exhibit high capillary fouling rates in lysozyme were chosen.



(Wisniewski *et al, Unpublished*)



(Harata *et al,* 1993)

Assumptions

University of Lethbridge

1. Capillary fouling deposits are proportional to sample concentration while exposed to the beam.

2. The buffer nor make/model of the instrument used contribute to the extent of fouling that occurs.

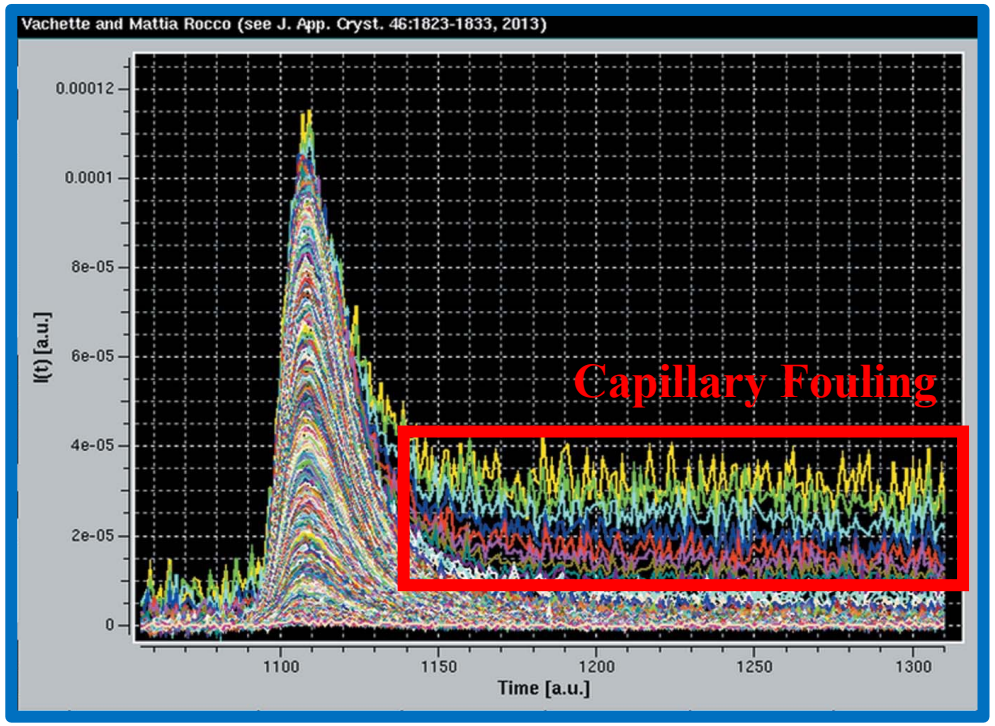
3. No “cleaning” occurs while the sample is eluting. (First approximation.)

4. The proportionality constant is species independent.

Steady-State End Signal

University of Lethbridge

- First, calculate average intensity for a region post-elution.
- If the value is close to 0, then no further steps are necessary. If lower than zero, other experimental issues.
- If greater than zero, capillary fouling is likely.



Brookes et al., 2016

$$I_{BL}(q) = \sum_{k=1,m} \frac{I(q, t_{sk})}{m}$$

Iterative Approach

University of Lethbridge

Step #1: Set the initial baseline to zero: $B_0(q, t) = 0$

Step #2: Loop from i = 0 to the desired number of iterations.

$$D(q, t_k) = \gamma(q)[I(q, t_{k-1}) - B(q, t_{k-1})]$$

Iterative Approach 2

University of Lethbridge

$$I_{TOT_i}(q) = \sum_{k=1}^{n-m} I(q, t_k) - B(q, t_k) \qquad I_{BL}(q) = \sum_{k=1,m} \frac{I(q, t_{sk})}{m}$$

$$\gamma_i(q) = \frac{I_{BL}(q)}{I_{TOT_i}(q)} \qquad |\gamma_i(q) - \gamma_{i-1}(q)| < \varepsilon$$

Iterative Approach 3

University of Lethbridge

$$\gamma_i(q) = \frac{I_{BL}(q)}{I_{TOT_i}(q)}$$

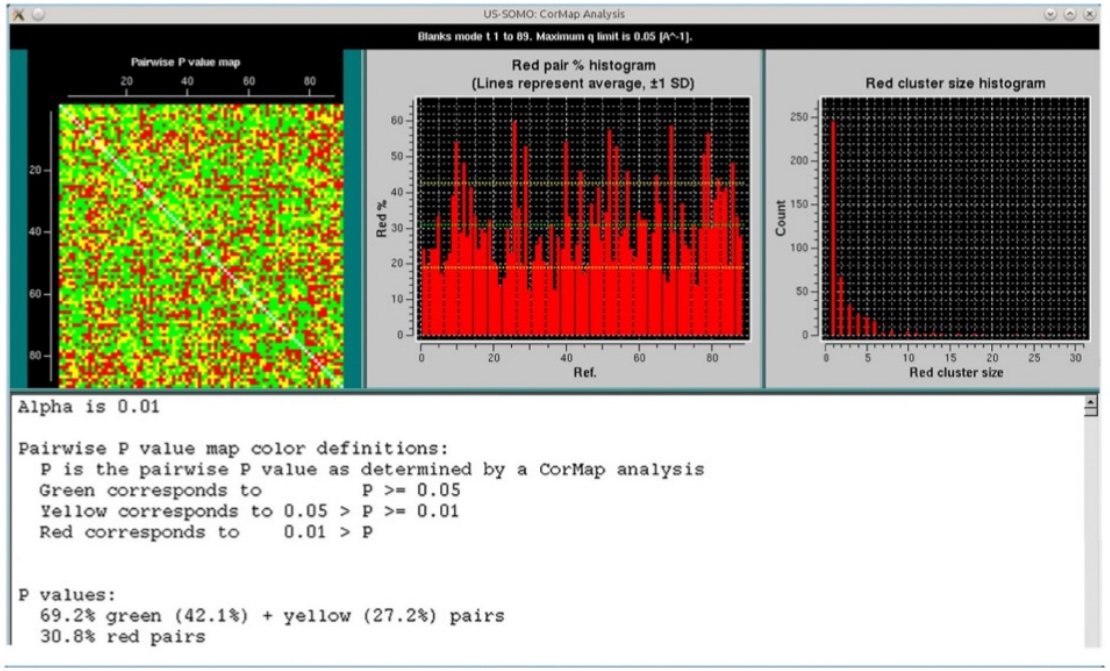
$$D_i(q, t_k) = \gamma_i(q)[I(q, t_{k-1}) - B_i(q, t_{k-1})]$$

$$B_{i+1}(q, t) = \sum_{t'=1}^{t} D_i(q, t')$$

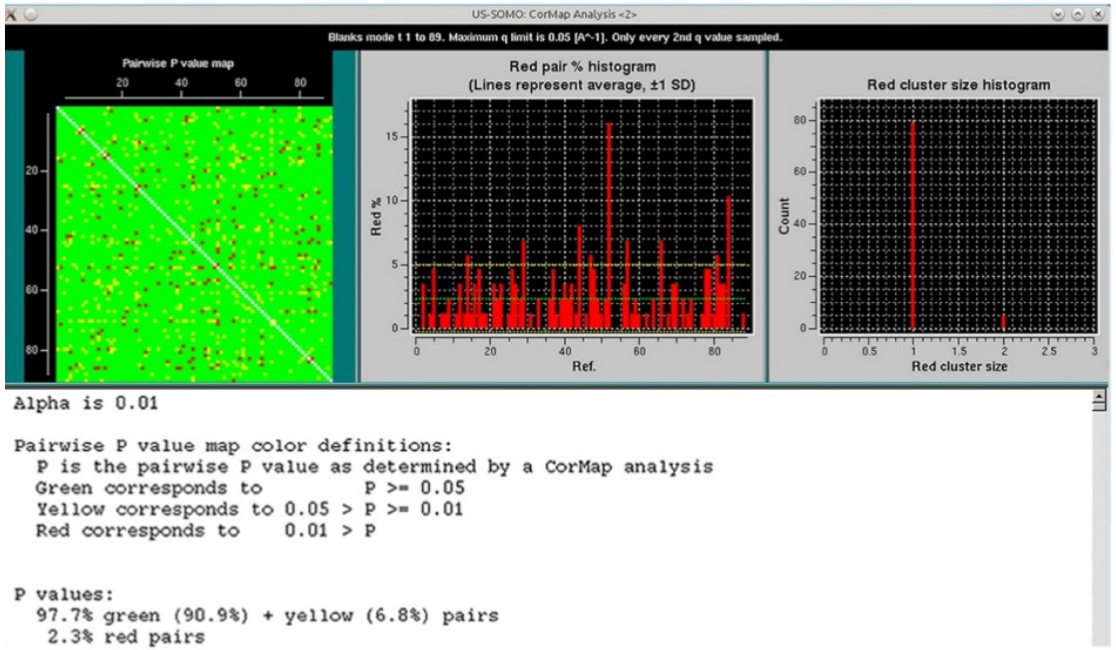
Lysozyme SEC-SAXS Experiment Analysis

University of Lethbridge

- Initially, buffer frames were shown to be significantly different in a control.



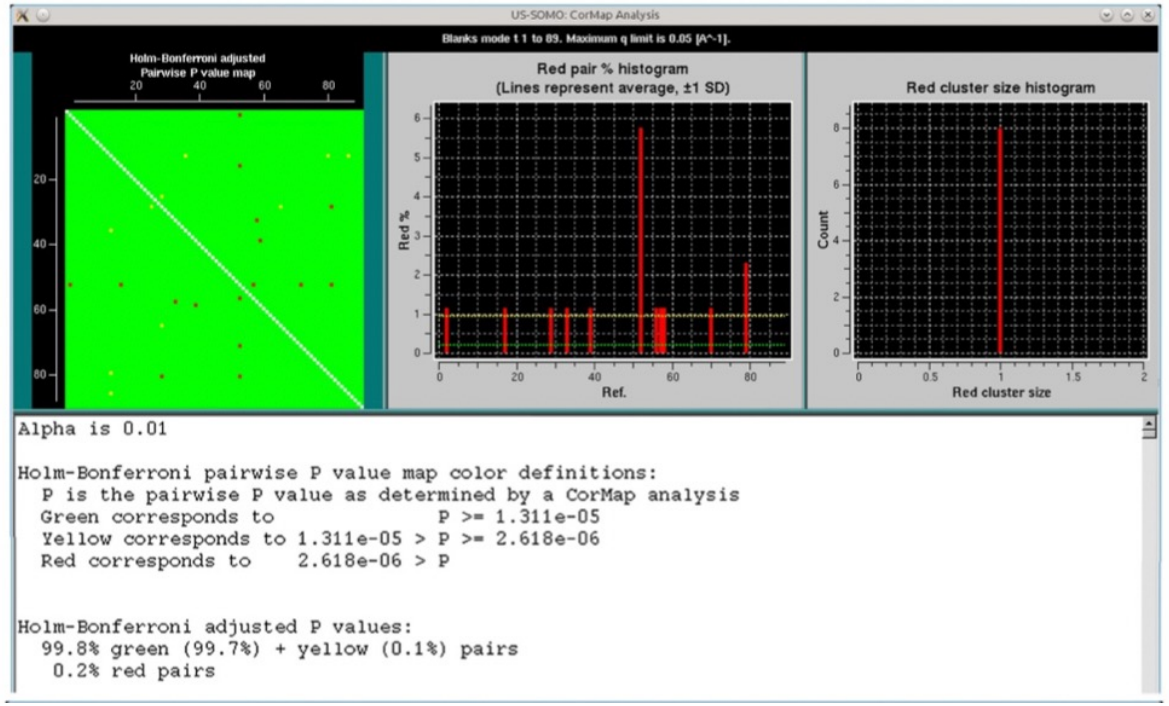
Brookes et al., 2016



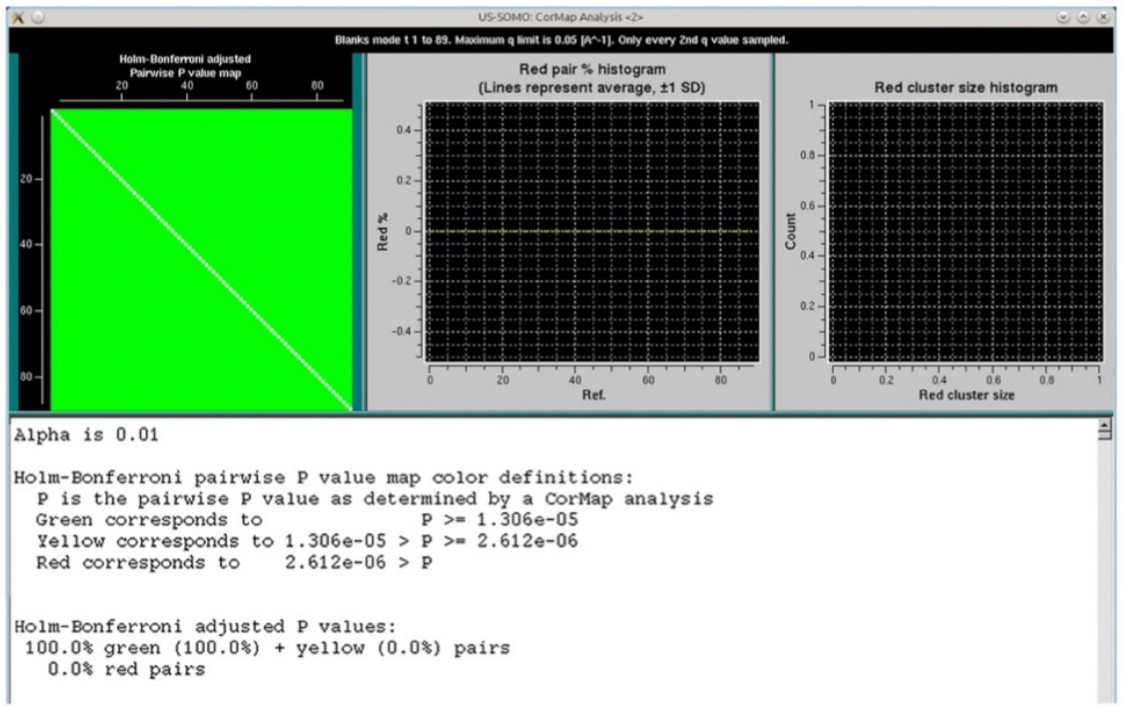
Brookes et al., 2016

Holm-Bonferroni Adjustment

University of Lethbridge



Brookes et al., 2016

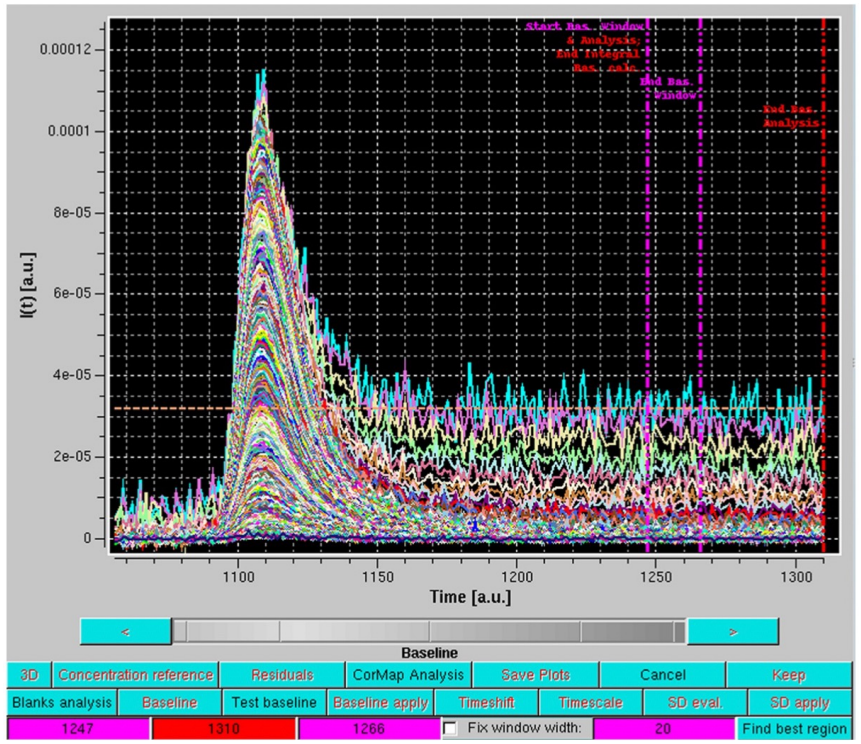


Brookes et al., 2016

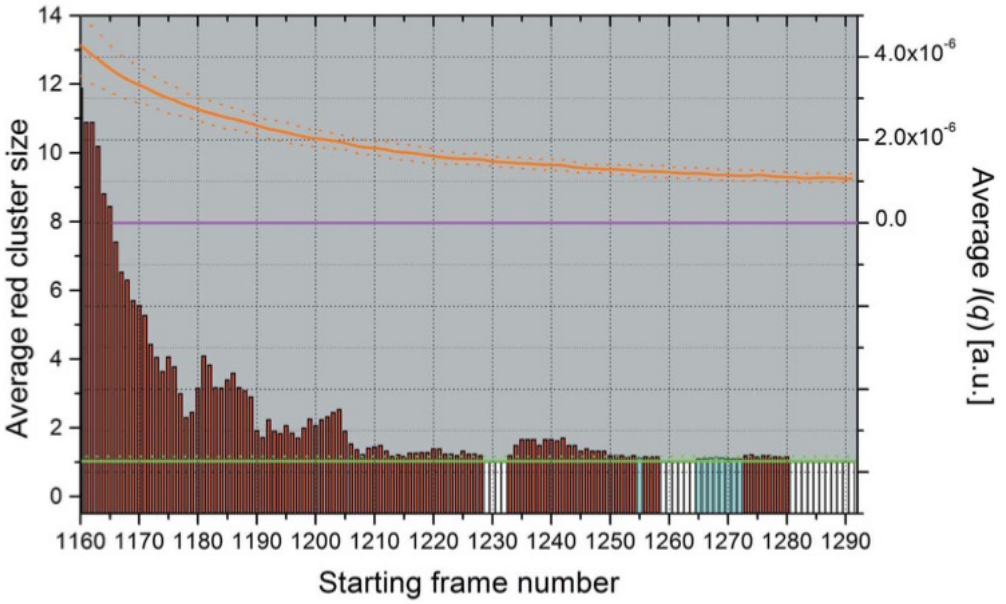
Identification of the Baseline Region

University of Lethbridge

- Identification of a stable final baseline region is identified using correlation map.



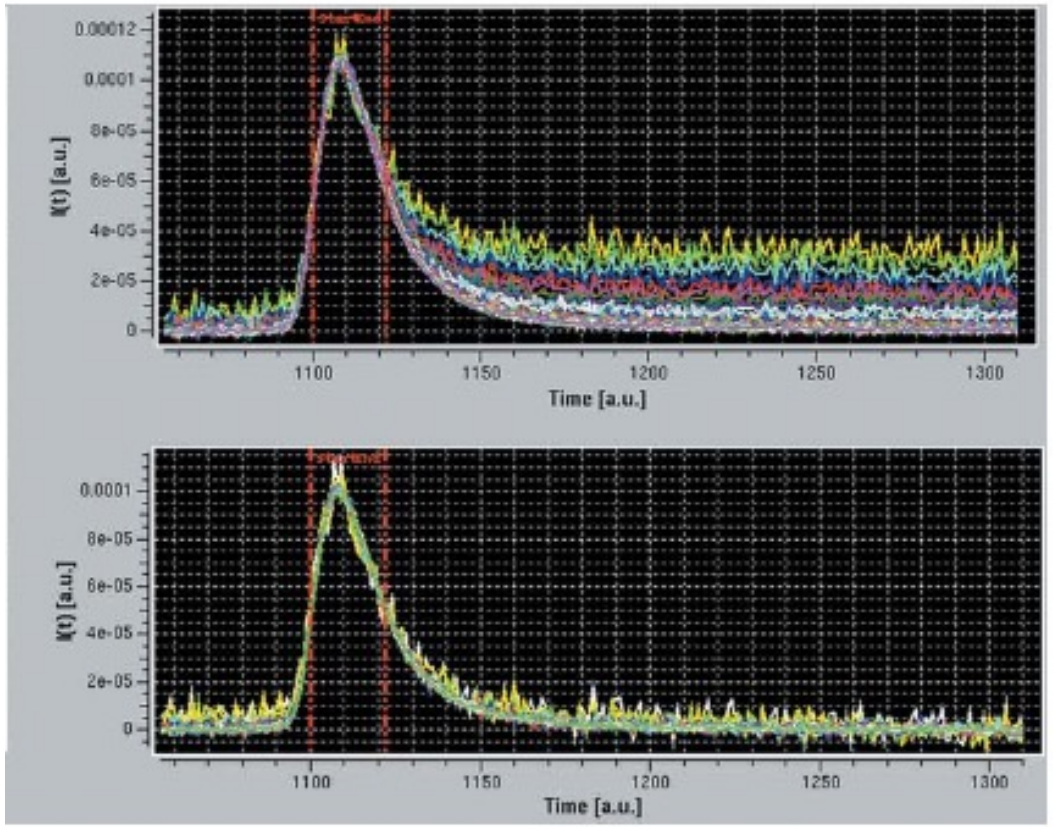
Brookes et al., 2016



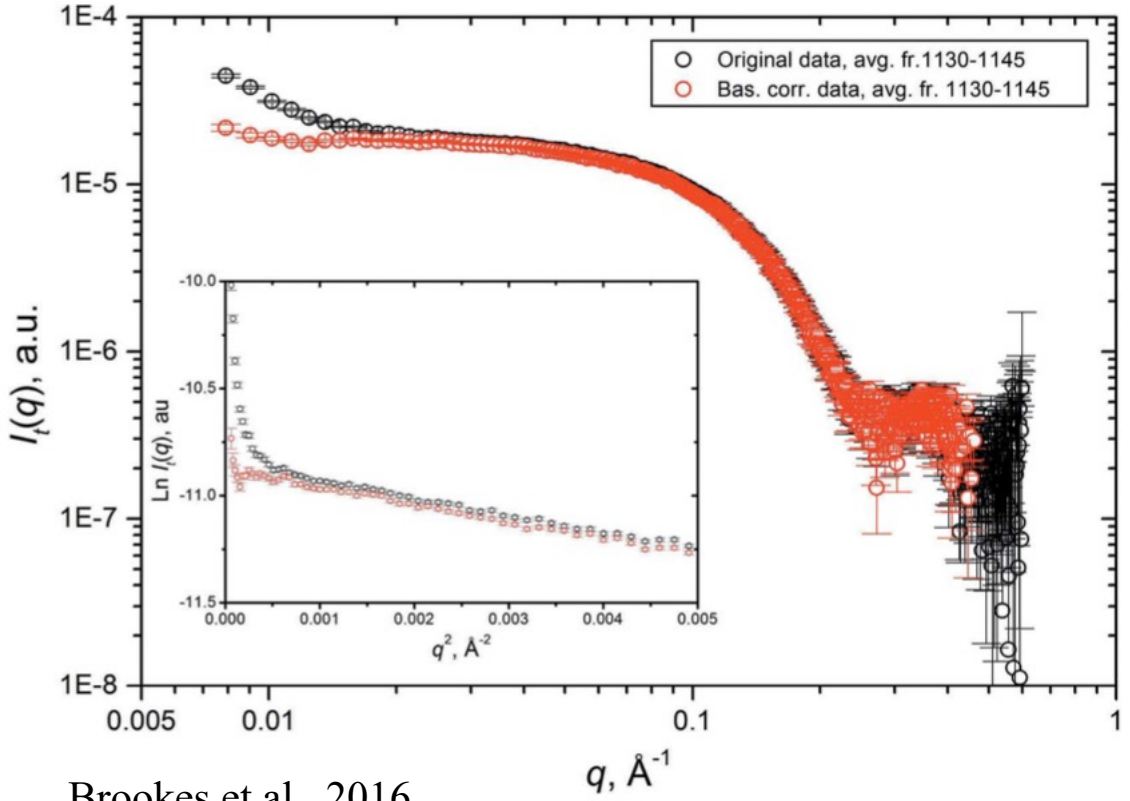
Brookes et al., 2016

Results of Baseline-Correction Procedure

University of Lethbridge



Brookes et al., 2016

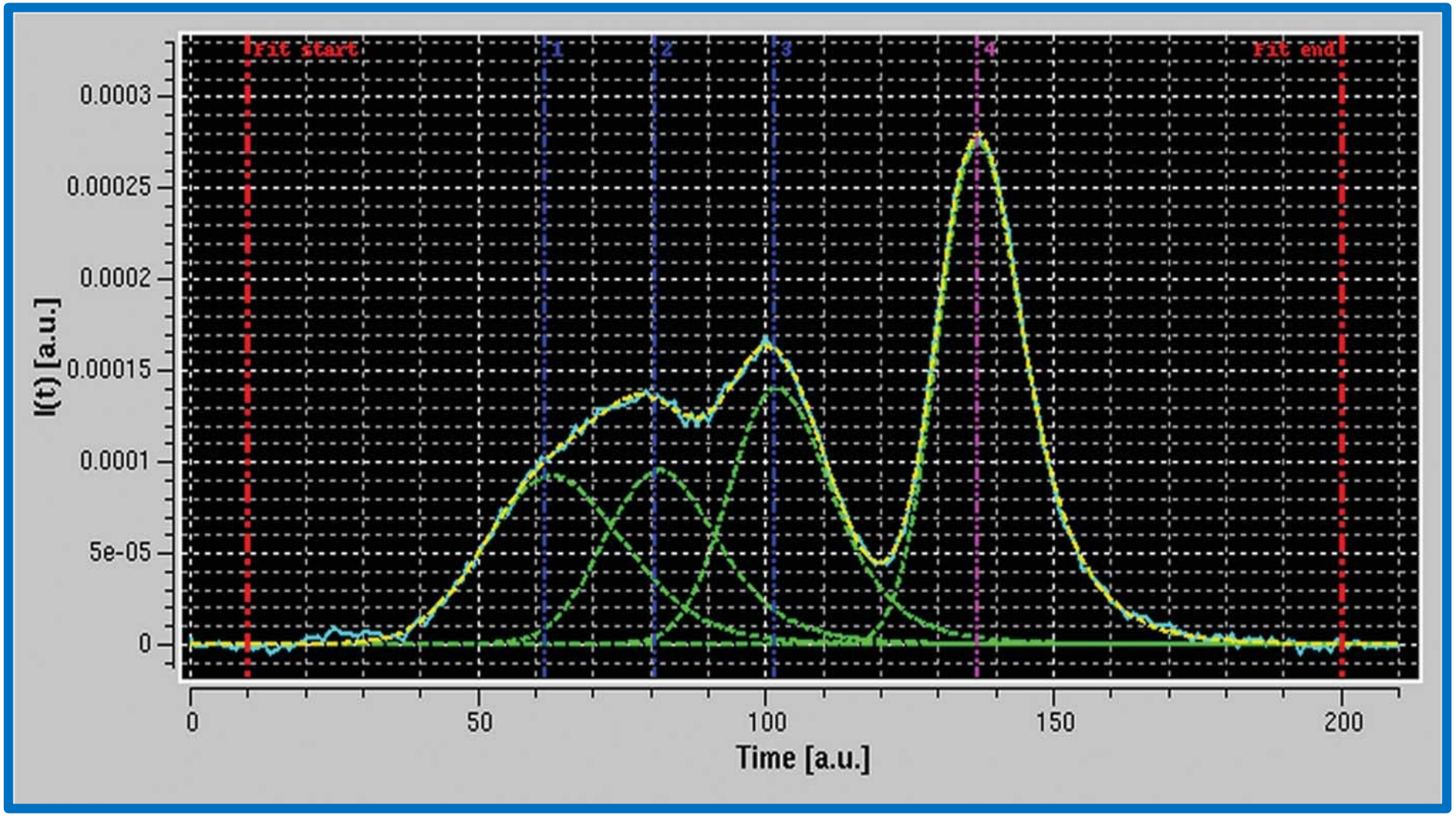


Brookes et al., 2016

Poor HPLC Resolution

University of Lethbridge

- Different species may not be completely resolved by the size exclusion column.



Brookes et al., 2016

Non-Symmetrical Gaussian Functions

University of Lethbridge

Classical Gaussian Function:

$$y = \frac{a_0}{2\pi^{\frac{1}{2}} a_2} e^{\left[-\frac{1}{2}\left(\frac{x-a_1}{a_2}\right)^2\right]}$$

Exponentially Modified Gaussian Function:

$$y = \frac{a_0 e^{\left[\frac{1(x-a_1)^2}{2\ a_3^2+a_2^2}\right]}\left\{1 + erf\left[\frac{a_3(x-a_1)}{2^{\frac{1}{2}} a_2 (a_3^2+a_2^2)^{\frac{1}{2}}}\right]\right\}}{(2\pi)^{\frac{1}{2}}(a_3^2+a_2^2)^{\frac{1}{2}}}$$

Half-Gaussian Modified Gaussian Function:

$$y = \frac{a_0}{2a_3} e^{\left(\frac{a_2^2}{2a_3^2}+\frac{a_1-x}{a_3}\right)}\left[erf\left(\frac{x-a_1}{2^{\frac{1}{2}} a_2} - \frac{a_2}{2^{\frac{1}{2}} a_3}\right) + \frac{a_3}{|a_3|}\right]$$

Exponentially Modified Gaussian + Half-Gaussian Modified Gaussian Function:

$$y = \frac{a_o}{4a_3} e^{\left(\frac{2a_1a_3-2a_3x+a_2^2}{a_3^2}\right)} erfc\left(\frac{a_1a_3 - a_3x + a_2^2}{2^{\frac{1}{2}} a_2 a_3}\right) + \left[\frac{a_0}{2(2\pi)^{\frac{1}{2}}(a_2^2+a_4^2)^{\frac{1}{2}}}\right] e^{\left[\frac{1(a_1-x)^2}{2\ a_2^2+a_4^2}\right]} erfc\left[\frac{a_4(a_1-x)}{2^{\frac{1}{2}} a_2 (a_2^2+a_4^2)^{\frac{1}{2}}}\right]$$

Goodness of Fit Estimation

University of Lethbridge

- χr_σ is used as a noise independent goodness of fit estimator.

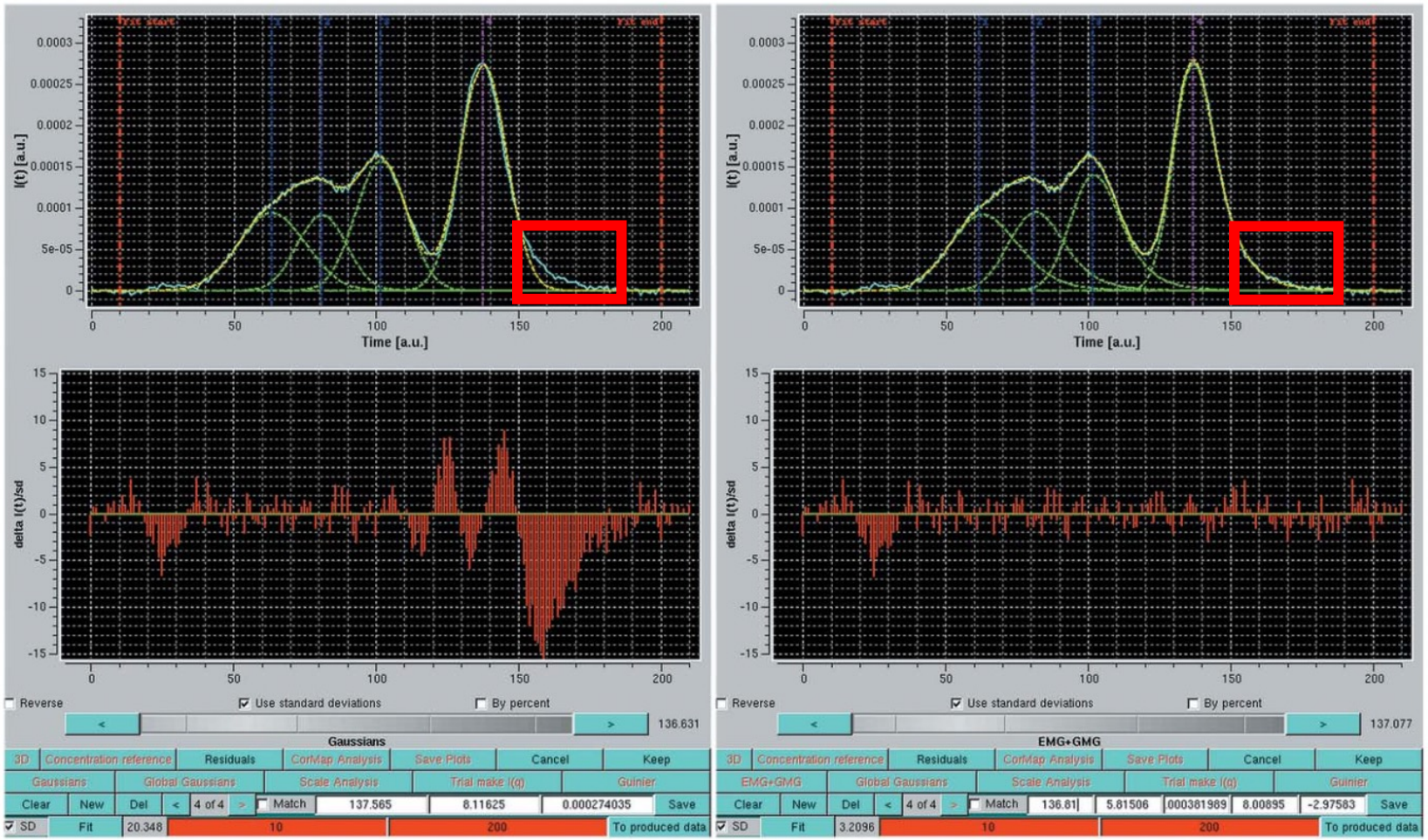
$$r_\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{\sigma_{exp}(q_i)}{I_{exp}(q_i)}\right]^2$$

- Correlation Map statistical analysis was also used.

$$\chi^2 = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{I_{exp}(q_i) - I_{calc}(q_i)}{\sigma_{exp}(q_i)}\right]^2$$

Efficacy of SVD with Different Gaussian Functions

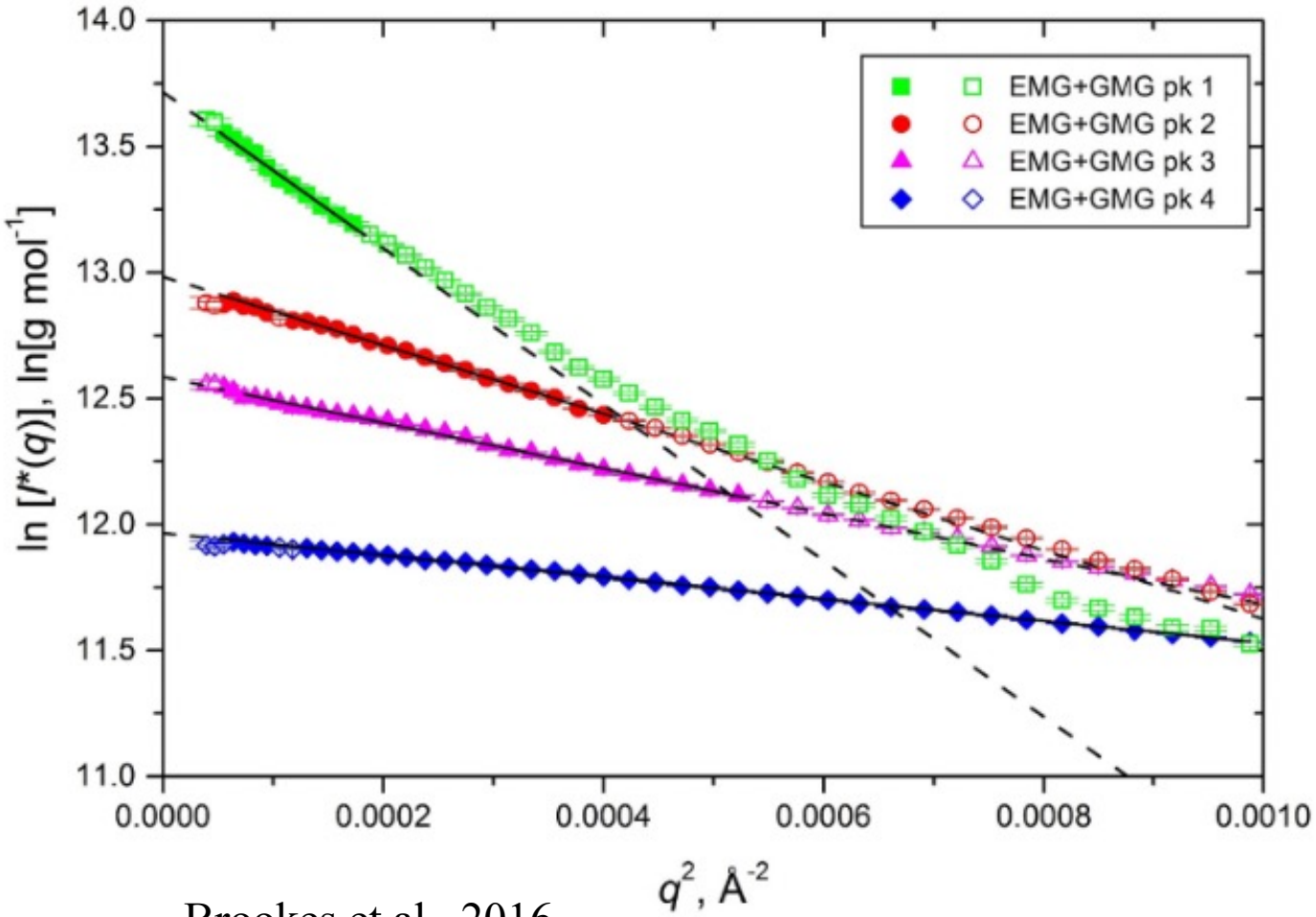
University of Lethbridge



Cyan represents the original curve, yellow is the generated curve, and green represents the individual Gaussian distributions generated. (Brookes et al., 2016)

Efficacy of SVD with Different Gaussian Functions 2

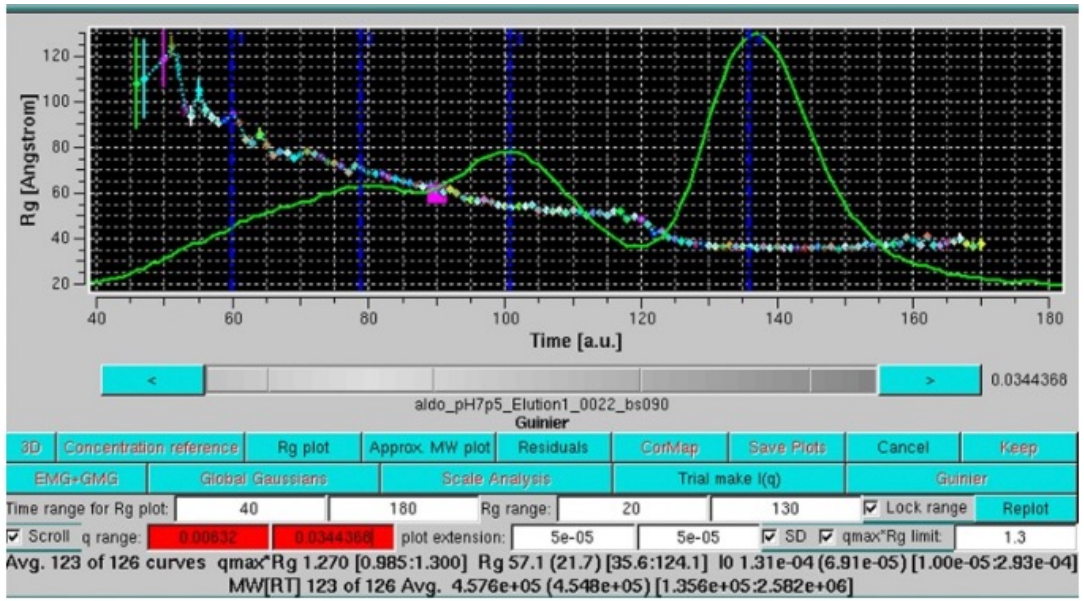
University of Lethbridge



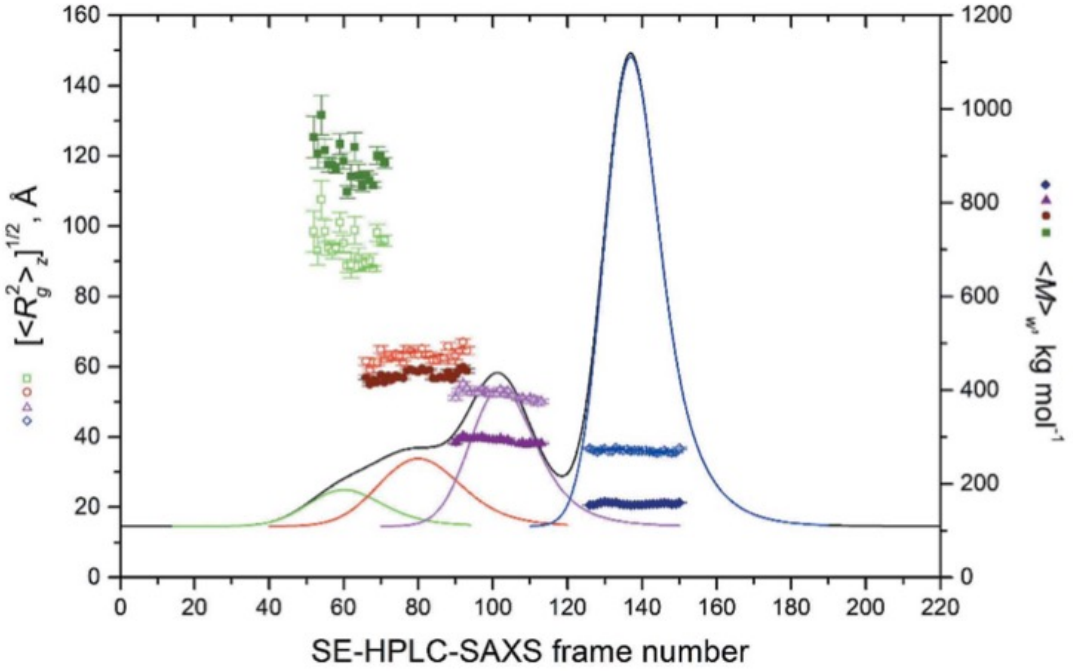
Brookes et al., 2016

Efficacy of SVD with Different Gaussian Functions 3

University of Lethbridge



Brookes et al., 2016



Brookes et al., 2016

Efficacy of SVD with Different Gaussian Functions 4

University of Lethbridge

EMG + GMG peak No.	Frame(s)	$[\langle R_g^2 \rangle_z]^{1/2}$ (Å)	Guinier $\langle M \rangle_w$ (kg mol^{-1})	$q_{min} - q_{max}$ (Å^{-1})	χ^2	*SAXS-MoW* $\langle M \rangle_w$ (kg mol^{-1})
1	60 (top)	101.1 ± 2.8	925 ± 22	0.00743–0.01257	0.0205	n.d.
1	Average of frames 53–72 results	94.4 ± 5.1	886 ± 40	n.a.	n.a.	n.d.
1	Average frame of frames 53–72	94.8 ± 1.4	897 ± 11	0.00743–0.01372	0.0126	1000 ± 3
2	80 (top)	63.6 ± 0.9	441 ± 4	0.00800–0.01886	0.0152	n.d.
2	Average of frames 67–94 results	63.1 ± 1.7	432 ± 10	n.a.	n.a.	n.d.
2	Average frame of frames 67–94	63.4 ± 0.3	434 ± 1	0.00800–0.01886	0.0058	462 ± 1
3	102 (top)	53.5 ± 0.6	297 ± 2	0.00857–0.02229	0.0127	n.d.
3	Average of frames 90–113 results	52.1 ± 1.4	292 ± 5	n.a.	n.a.	n.d.
3	Average frame of frames 90–113	52.0 ± 0.3	292 ± 1	0.00857–0.02286	0.0065	306 ± 4
4	137 (top)	36.4 ± 0.1	156 ± 0	0.00857–0.03143	0.0015	
4	Average of frames 126–150 results	36.1 ± 0.4	157 ± 2	n.a.	n.a.	n.d.
4	Average frame of frames 126–150	36.0 ± 0.1	157 ± 0	0.00857–0.03143	0.0031	154 ± 1

Brookes et al., 2016

Theoretical molecular weight of an aldolase homotetramer: 157kDa

Theoretical molecular weight of dimer of homotetramers: 314kDa

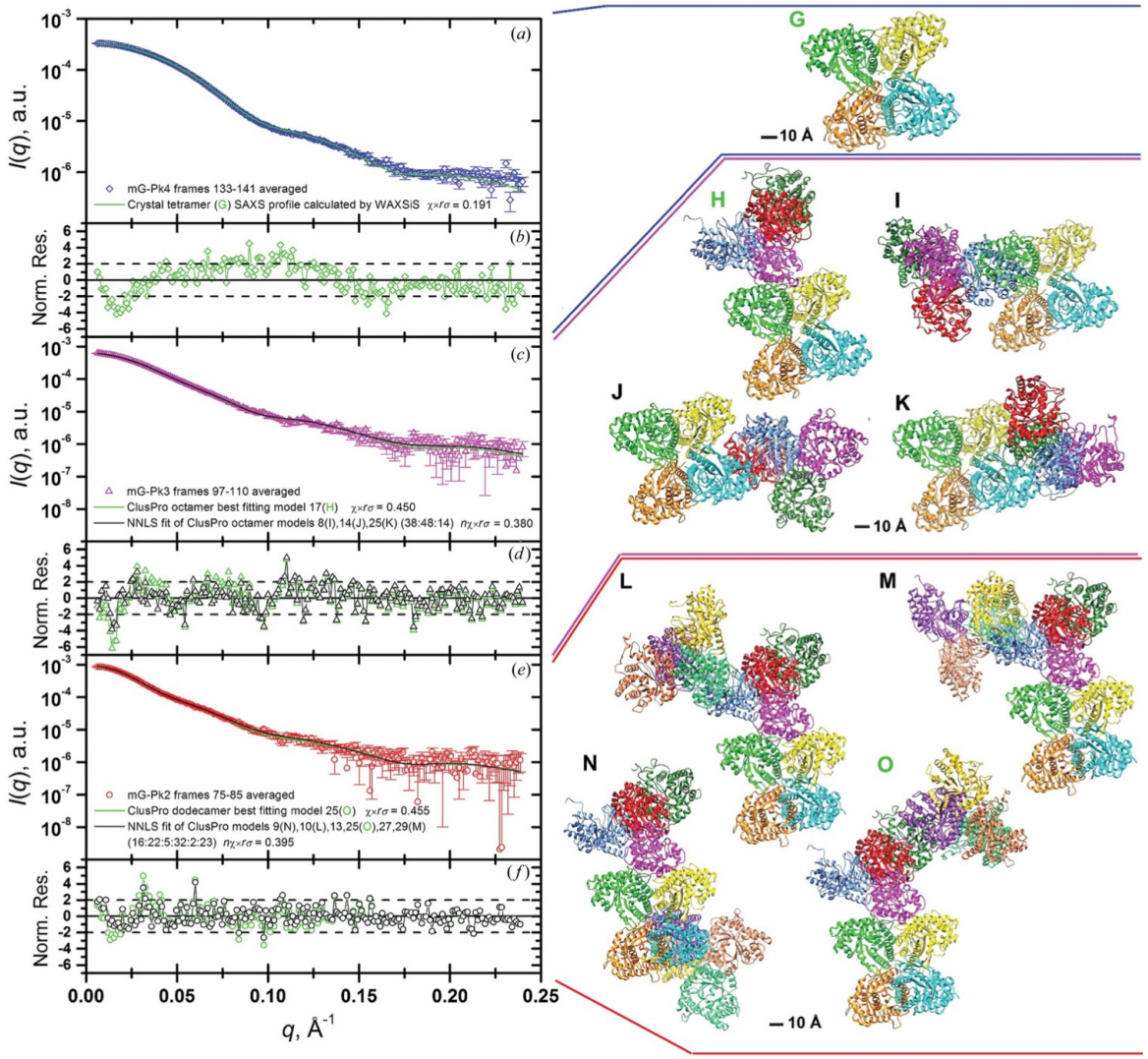
Theoretical molecular weight of trimer of homotetramers: 471kDa

Theoretical molecular weight of hexamer of homotetramers: 942kDa

Modelling of Identified Complexes

University of Lethbridge

- Molecular docking experiments were carried out to try to visualize what the observed complexes would look like structurally.



Brookes et al., 2016

Conclusions/Future Directions

University of Lethbridge

- The new analysis methods are effective in reducing the affect of capillary fouling and poorly resolved SEC peaks on experimental SAXS data.

- The analysis methods presented are more efficacious than most alternatives available at the time.

- Currently, the software cannot be considered a high-throughput tool, due to the number of manual adjustments that must be made. The authors may release a subsequent edition of the software, automating several steps in a future release.

University of Lethbridge

Questions?

Literature Cited

Brookes, E., Pérez, J., Cardinali, B., Profumo, A., Vachette, P., & Rocco, M. (2013). Fibrinogen species as resolved by HPLC-SAXS data processing within the UltraScan Solution Modeler (US-SOMO) enhanced SAS module. *Journal of Applied Crystallography*, *46*(6). https://doi.org/10.1107/S0021889813027751

Brookes, E., Vachette, P., Rocco, M., & Pérez, J. (2016). US-SOMO HPLC-SAXS module: Dealing with capillary fouling and extraction of pure component patterns from poorly resolved SEC-SAXS data. *Journal of Applied Crystallography*, *49*(5). https://doi.org/10.1107/S1600576716011201

Franke, D., Jeffries, C. M., & Svergun, D. I. (2015). Correlation Map, a goodness-of-fit test for one-dimensional X-ray scattering spectra. *Nature Methods*, *12*(5). https://doi.org/10.1038/nmeth.3358

Graewert, M. A., Da Vela, S., Gräwert, T. W., Molodenskiy, D. S., Blanchet, C. E., Svergun, D. I., & Jeffries, C. M. (2020). Adding size exclusion chromatography (SEC) and light scattering (LS) devices to obtain high-quality small angle X-ray scattering (SAXS) data. *Crystals*, *10*(11). https://doi.org/10.3390/cryst10110975

Malaby, A. W., Chakravarthy, S., Irving, T. C., Kathuria, S. V., Bilsel, O., & Lambright, D. G. (2015). Methods for analysis of size-exclusion chromatography-small-angle X-ray scattering and reconstruction of protein scattering. *Journal of Applied Crystallography*, *48*. https://doi.org/10.1107/S1600576715010420