Electronic supplementary information for

A problem-based learning activity for enhancing inquiry skills and facilitating conceptual change in a biological chemistry course

Wanda M. Valsecchi, José M. Delfino, Javier Santos, Santiago E. Faraj

This file contains the following materials:

- ***Supplementary figures*** with additional result analysis.
- The ***Assessment instrument*** used for pre- and post-questionnaire.
- The ***Laboratory handout*** used by students to carry out experiments and with instructions on how to produce their lab reports, and the assessment ***Rubric*** used to grade lab reports.

Data sets used to illustrate results throughout this paper, which can be used in a data-analysis exercise, can be found in the public repository https://github.com/wanvalsecchi/Data_Set_SEC_LS.git

SUPPLEMENTARY FIGURES

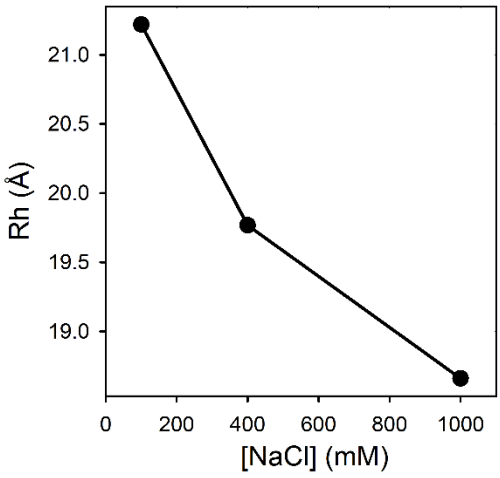


Figure S1. Apparent hydrodynamic behavior of FXN. The Rh of FXN was calculated from SEC data obtained at different NaCl concentrations. Experiments were carried out at room temperature using a Superose 12 prepacked column equilibrated with buffer 20 mM Tris·HCl, pH 7.0, 1 mM EDTA, and varying concentrations of NaCl.

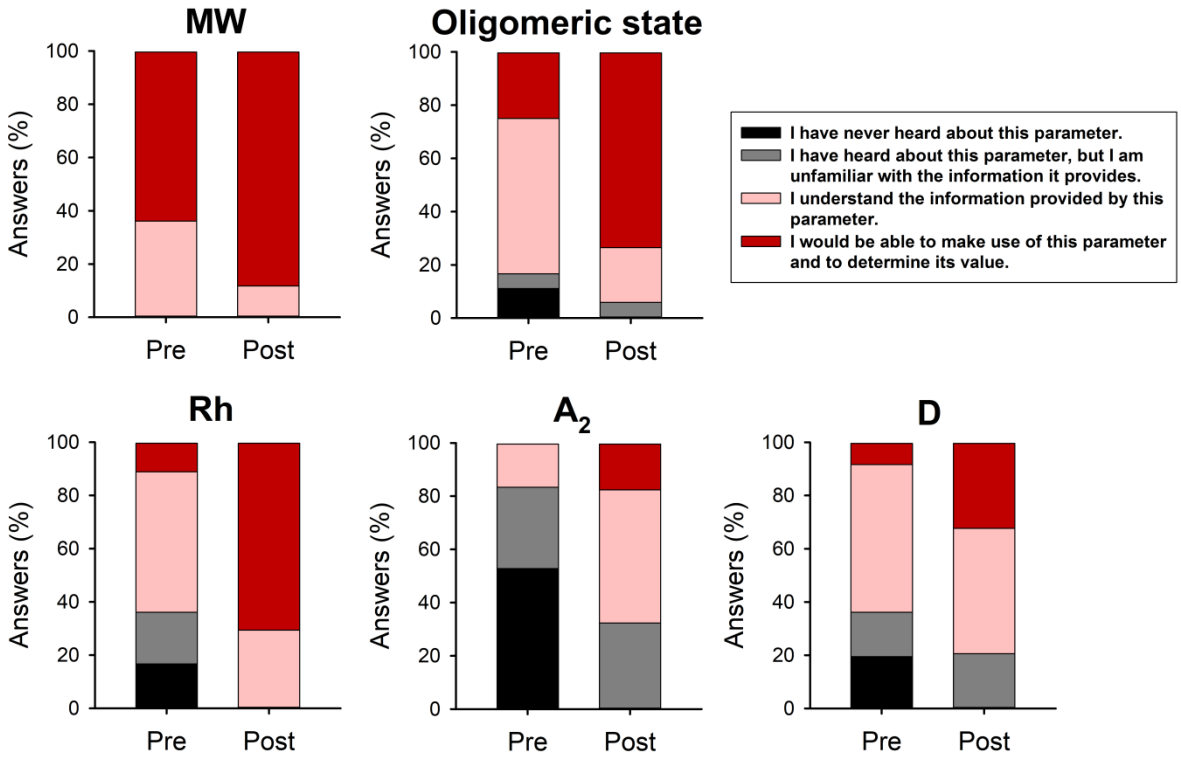


Figure S2. Self-perception. Students were asked to report on their perceived understanding of different parameters discussed throughout the module. This questionnaire was presented as a multiple-choice survey and was answered by 36 respondents in pre-test and 34 respondents post-tests.

Assessment instrument

(Please note that the actual questionnaire was administered in Spanish)

DETERMINATION OF PROTEIN MOLECULAR WEIGHT AND SIZE

A1) Choose the option that better describes your understanding of each of the following techniques: analytical size exclusion chromatography (SEC), static light scattering (SLS) and dynamic light scattering (DLS).

	SEC	SLS	DLS
I understand the theory behind the technique and would be able to make experimental use of it.	☐	☐	☐
I understand the theory behind the technique as well as how it is experimentally applied.	☐	☐	☐
I understand the theory behind the technique, but I am uncertain about how it is experimentally applied.	☐	☐	☐
I have heard about the technique, but I am uncertain about the theory behind the methodology or how it is experimentally applied.	☐	☐	☐
I have never heard about this technique.	☐	☐	☐

A2) Choose the option (1-4) that you feel best describes your situation with respect to each of the following parameters of a protein.

Molecular weight (MW) _____

Hydrodynamic radius (Rh) _____

Second Virial Coefficient (A_2) _____

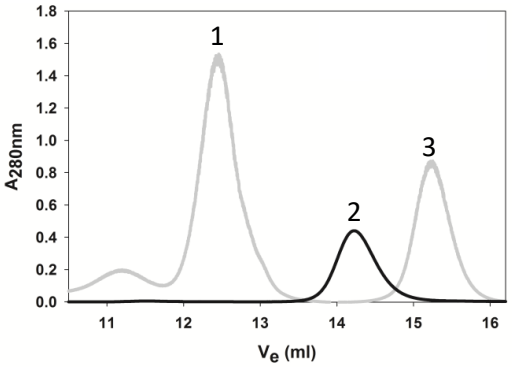
Avece factor (Fa) _____

Diffusional coefficient (D) _____

Oligomeric state _____

1. I would be able to make use of this parameter and to determine its value
2. I understand the information provided by this parameter
3. I have heard about this parameter, but I am unfamiliar with the information it provides
4. I have never heard about this parameter

Q1) The following figure shows the SEC profiles of proteins A (100kDa), B (66kDa) and C (14kDa).



Peak assignment would be:

☐ 1-A, 2-B, 3-C

☐ 1-C, 2-B, 3-A

☐ 1-A, 2-C, 3-B

☐ The provided information is not sufficient to make a decision.

Q2) The molecular weight of a protein was determined by SEC and SLS. Results were the following:

SEC: 25.8 kDa

SLS: 105.3 kDa

The molecular weight of a monomer calculated from the amino acid sequence is 26.6 kDa.

Which is the oligomeric state of the protein?

…….

Q3) The hydrodynamic radius (Rh) of a protein was determined by SEC and DLS. Results were the following:

SEC: 23.3 Å

DLS: 40.1 Å

The Rh was estimated at 39.9 Å using an empirical correlation between the Rh and the molecular weight.

How does the actual degree of compaction of the protein compare with the expected one?

☐ It is more compact than a typical globular protein

☐ It is as compact as a typical globular protein

☐ It is less compact than a typical globular protein

Q4) State whether the following statements are true or false:

_____ *Given a mixture of proteins of different sizes, a SEC column attached to a UV detector would always allow obtaining the MW of each of them.*

_____ *A light scattering device (685 nm laser) with a single 90° detector may be adequate to obtain the MW of a monodisperse solution of a polymer with a mean radius of 15 nm.*

_____ *In order to obtain the MW of a solution of a small, pure protein, a light scattering device (685 nm laser) with a single 90° detector (online with a SEC column) would be a convenient choice.*

Determination of protein molecular weight and size

Laboratory handout

In this practical we will apply size exclusion chromatography, static light scattering (SLS) and dynamic light scattering (DLS) to obtain the molecular weight (MW) and the hydrodynamic or Stokes radius (Rh) of a set of proteins.

Laboratory guide

MW and Rh determination will be performed by size exclusion chromatography (SEC) and light scattering (LS) in a continuous flow mode. Our system consists of a set of devices arranged in tandem as outlined in Fig. 1, which allows the simultaneous determination of the MW and the value of the diffusional coefficient (D). The samples will be thoroughly centrifuged, loaded into a 100 µl injector loop and filtered through a Sepharose 12 SEC column. The flow rate of the buffer (20 mM Tris-HCl, 100 mM NaCl, pH 7.4) will be set constant at 0.3 ml min^{-1}. As shown in Figure 1, the eluate passes through the UV detector (*Jasco UV 2075 plus*), which allows determining the elution volume (Ve) of each peak, and through the multi-angle LS detector (*miniDAWN Tristar, Wyatt Technology*), which measures the amount of light scattered to calculate the absolute molar mass of the sample by SLS. An optical fiber receiver is mounted in the 90° detector and coupled to the DLS module for the determination of D and posterior calculation of Rh (*WyattQELS, Wyatt Technology*). Collected data is transferred to a computer where it is processed with ASTRA software (*Wyatt Technology*).

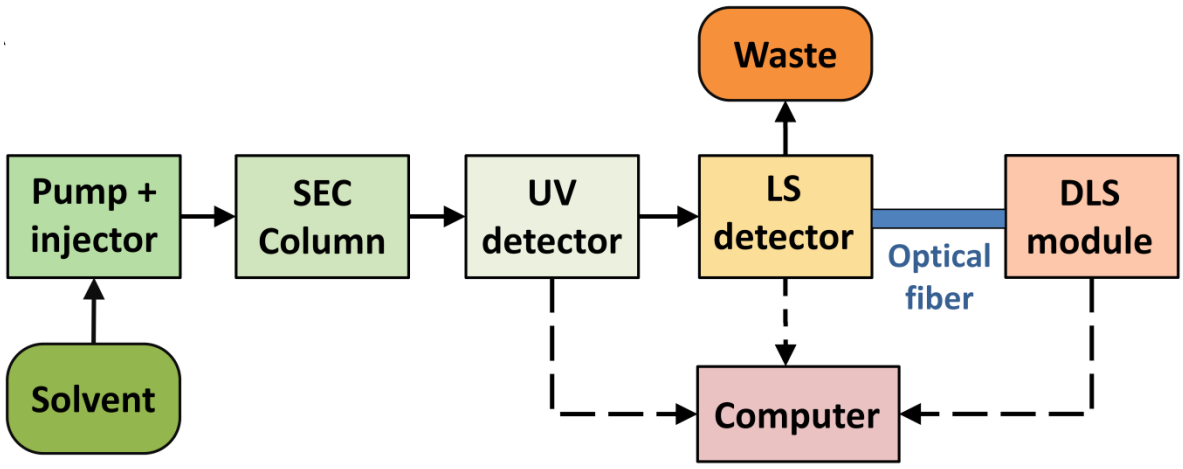


Fig. 1

We will perform three chromatographic runs during this practical:

Sample A: Molecular mass standards that will be used to calibrate the SEC column:

- Thyroglobulin (bovine) 670 kDa
- γ-globulin (bovine) 158 kDa
- Ovalbumin (chicken) 44 kDa
- Myoglobin (horse) 17 kDa
- Vitamin B12 1.35 kDa

Sample B: A mixture of bovine serum Albumin (BSA) and human Frataxin (FXN):

- BSA 66 kDa *
- FXN 14 Da *

Sample C: A solution of Hypoxanthine phosphoribosyl transferase (HPRT) from *T. cruzi*:

- HPRT 27 kDa *

*The molecular weights shown are those of a monomer of the protein, calculated using the amino acid sequence.

Data from experiments will be exported as spreadsheet files.

Data analysis

Open the spreadsheet file. You will find three sheets, one for each run. Identify the columns within each spreadsheet:

- Column A is the independent variable **time**, expressed in minutes.
- Column B is the **absorbance at 280 nm** of the SEC column eluate collected by the UV detector.
- Column C is the **MW** obtained using data collected by the LS detector, expressed in Daltons.
- Column D is the **Rh**, calculated using the diffusional coefficient value (D) obtained by the DLS module, expressed in nanometers.

Part 1: MW and Rh determination by SEC

First, we will calculate the MW and Rh using only elution volumes of the different species.

Use data obtained for **Sample A** (MW standards).

1. Using the buffer flow rate, calculate the Ve for each time. Plot the chromatogram using Ve as the independent variable: absorbance (in ordinate y axis) vs. Ve (in abscissa x axis).
2. Identify the peaks on the chromatogram and assign them to the different species present in the mixture of MW standards. Note that the components of the mixture elute according to their MW.
3. Make note of the Ve of the maximum absorbance signal for each of the peaks.
4. Plot Log(MW) vs. Ve for each MW standard and find the best-fitting line to the data. This will be the characteristic equation of the system in our working conditions.

Use UV absorption data (column B) obtained for **Samples B and C**.

5. Using the buffer flow rate, calculate the Ve for each time. Plot the chromatograms using Ve as the independent variable (absorbance vs. Ve).
6. Identify the peaks on the chromatograms and find the Ve that yielded the maximal absorbance signal for each of them.
7. Using the equation of the best-fitting line calculate the MW of each of the peaks.
8. In the work by Uversky (1993)[*], identify the equation that correlates Rh with MW for proteins in native conformation, and use it to estimate the Rh of each species from the value of MW obtained for each peak (assume that each peak corresponds to a single species). Use MW values in Da (not kDa).

Part 2: MW and Rh determination by light scattering

Using data collected by LS and UV detectors, we can obtain the MW of polymers that elute from the size exclusion column (SLS). Additionally, the DLS module allows obtaining the value of the diffusion coefficient (D), which is used to calculate the Rh value. Since we performed an on-line experiment (flow mode), components present in the sample are fractionated in the column and reach the light scattering detector cell at different times. In this fashion, many light scattering data are obtained for each peak. We will use the MW and Rh data from **Samples B and C** (columns C and D in the spreadsheets).

9. Calculate the mean value and standard deviation of the MW and Rh for each peak.

[*] Uversky VN, Use of fast protein size-exclusion liquid chromatography to study the unfolding of proteins which denature through the molten globule (1993), *Biochemistry* 32(48):13288-98.

Laboratory report *[structured questionnaire – new format only]*

Follow the instructions in the *Data analysis* section. Then answer the following questions.

The chromatograms obtained for **Samples B and C** are shown in Fig. 2. *Note that absorbance values might not exactly correspond to those you obtained.*

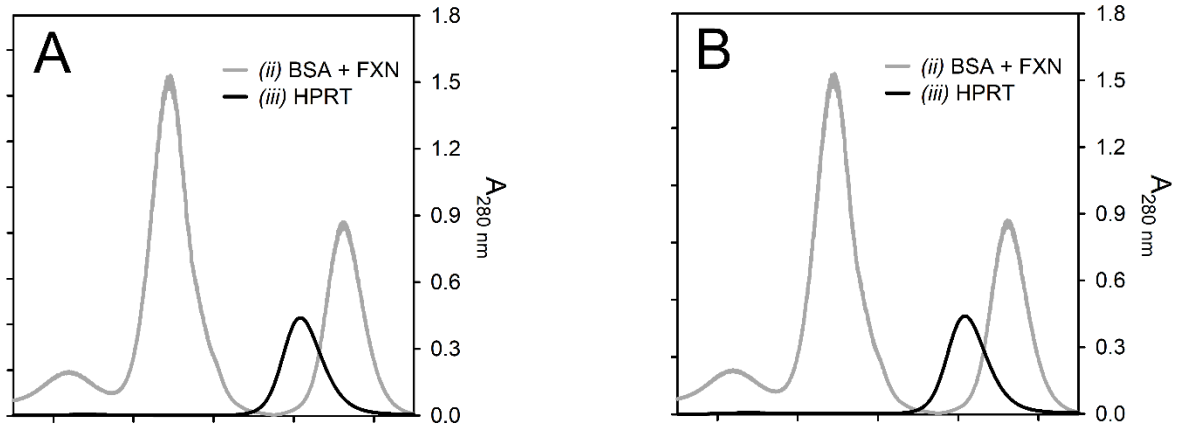


Fig. 2

1 Using the left (primary) axes, plot **MW values obtained by SLS in Fig. 2A** and **Rh values obtained by DLS in Fig. 2B.** Suggest suitable labels and units for the graph axes when appropriate.

2 Identify which peaks correspond to each protein, **BSA, HPRT and FXN**, and clearly label them on Fig. 2.

3 Complete the missing spaces in Table 1 with the MW values obtained by SEC and Rh values calculated using the equation that describes the dependence of Rh with the MW for proteins in their native conformation proposed by Uversky (1993):

$$\log(Rh) = -(0.254 \pm 0.002) + (0.369 \pm 0.001)\log(MW) \qquad \textbf{Eq.1}$$

Bear in mind that you should use MW values in Da (not kDa) when using this equation.

Table 1. MW values obtained by SEC and Rh values calculated using an empirical correlation for each peak obtained for Samples B and C.

Sample	Peak (left to right)	MW obtained by SEC (kDa)	Calculated Rh (Å)
B	1	……	……
B	2	……	……
B	3	……	……
C	1	……	……

Table 2 shows sequence-based MW values for several possible species of the proteins under study.

Table 2. Calculated MW and Rh values obtained for several species of BSA, FXN and HPRT.

Protein		Sequence-based MW (Da)	Calculated Rh (Å)
BSA	Monomer	66 300	……
	Dimer	132 600	……
	Trimer	198 900	……
FXN	Monomer	13 700	……
	Dimer	27 400	……
HPRT	Monomer	26 600	……
	Dimer	53 200	……
	Tetramer	106 400	……

4 Complete the missing spaces in Table 2 with the Rh values calculated using **Eq. 1**.

5 Using information in Tables 1 and 2, and considering that FXN does not interact with BSA, suggest the most probable scenario for species (protein and oligomeric state) present in each peak.

Sample B, peak 1:……

Sample B, peak 2:……

Sample B, peak 3:……

Sample C, peak 1:……

6 Complete the missing spaces in Table 3 with MW values obtained by SLS and Rh values obtained by DLS.

Table 3. MW values obtained by SEC and Rh values calculated using an empirical correlation for each peak obtained for Samples B and C.

Sample	Peak (left to right)	MW obtained by SLS (kDa)	Rh obtained by DLS (Å)
B	1	……	……
B	2	……	……
B	3	……	……
C	1	……	……

7 Using information in Tables 2 and 3, and considering that FXN does not interact with BSA, suggest the most probable species (protein and oligomeric state) present in each peak.

Sample B, peak 1:…...……..……

Sample B, peak 2:…...……..……

Sample B, peak 3:…...……..……

Sample C, peak 1:…...……..……

8 Compare results obtained by SEC, SLS and DLS (use the mean ± SD as a comparison criterion) and suggest possible reasons for any discrepancy you find. In view of all the information you have analyzed, suggest the most probable species (protein and oligomeric state) in each peak and their MW and Rh value. Explain your answer in full.

……..……

……..……

……..……

……..……

……..……

……..……

Sections	Relevant information	Inadequate = 1	Adequate = 2	Good = 3	Very good = 4	Excellent = 5
Data processing	Figures: chromatograms Abs. vs Ve for MW markers; calibration curve; LS results; DLS results. Tabulated information: MW and Rh values calculated by sequence; SEC; and LS.	Most of the figures and tabulated information is missing. Figures or tables contain unclear information. *Orig. Format: figures have no or inadequate title or legend. Axes are not named or lack labels.	Some of the figures or tabulated information is missing or contains unclear information. *Orig. Format: figures have no or inadequate title or legend. Axes are not named or lack labels.	All figures are presented. Some of the tabulated information is missing or contains unclear information. *Orig. Format: figures have no or inadequate title or legend. Axes are not named or lack labels.	Most the required information is included. There is still room for improvement. *Orig. Format: figures have adequate titles and legends. Axes are named and correctly labeled.	All the required information is included in an effective way to convey results. *Orig. Format: figures have adequate titles and legends. Axes are named and correctly labeled.
Results analysis	Identification of peaks; Description of BSA, HPRT and FXN behaviors; Comparison of SEC with sequence-based values; Comparison of LS with sequence-based values; Comparison between SEC and LS results.	Most components are missing.	Some components are missing. Topics are weakly mentioned: e.g., no comparison is done. Some values are wrongly calculated.	Some of the components are missing. Topics are weakly mentioned: e.g., when contradiction between SEC and LS, the student ignores it or makes no effort for explaining it. Most values are correctly calculated.	Most components are analyzed, and all or most values are well calculated. Good general analysis but there is still room for improvement.	All components are very well analyzed. In some cases, values can be wrongly calculated or analyzed, but this is compensated with a very good general analysis and conclusions.
Conclusions	HPRT and FXN behave anomalously; Compaction hypothesis; Matrix interaction hypothesis.	Poorly organized. Reviewed of analysis but not proper discussion. Most key items are missing.	Conclusions are too brief. Student states MW and Rh by SEC without taking into account LS results. No hypothesis is proposed to explain apparently contradictory data. No discussion exists on Rh results.	Conclusions are brief. Student hypothesizes that FXN and HPRT behave anomalously, but does not propose any hypothesis. MW analysis is performed, but a discussion on Rh results is missing.	Conclusions are brief. Student finds discrepancies between SEC and SLS MW data and proposes some hypotheses. The Rh discussion is vague only about SEC, and does not compare it with LS.	Excellent conclusions. Student finds discrepancies between SEC and SLS MW data and proposes all or most of the expected hypotheses. The Rh discussion is complete, comparing SEC with LS in full.

*Requirements considered in the original format.

Letter grade	Criteria
A	Earned 5 in all sections.
AB	Earned at least 4 in *Data processing* AND *Results analysis*, AND 5 in *Conclusions.*
B	Earned at least 4 in all sections.
BC	Earned at least 3 in *Data processing* AND *Results analysis*, AND at least 4 in *Conclusions.*
C	Earned at least 3 in all sections.
D	Earned at least 2 in all sections.
E	Does not meet the minimum criteria for grade D