Optimization of two high-performance size-exclusion chromatography methods with diode-array detection for protein quantification and purity assessment

Camilla Lane

Degree Project in Analytical Chemistry, 2023
Department of Chemistry
Lund University
Sweden

MSc, 30 hp

LUND
UNIVERSITY

Optimization of two high-performance size-exclusion chromatography methods with diode-array detection for protein quantification and purity assessment

Camilla Lane

LUND
UNIVERSITY

Degree Project in Analytical Chemistry
2023
MSc, 30 hp

Supervisor:
Sr Scientist Marika Slivacka
Repligen AB

Ass. Prof. Peter Spégel
Center for Analysis and
Synthesis, Lund university

Examiner:
Sr. Lecturer Margareta Sandahl
Center for Analysis and
Synthesis, Lund university

Lund University
Department of Chemistry
Centre for Analysis and Synthesis
P.O. Box 124
SE-221 00 Lund, Sweden

Development of two methods to measure concentration and purity of proteins used in pharmaceutical production

An important part of pharmaceutical production is the purification of the active ingredient which can be done, in part, using proteins called affinity ligands which selectively bind specific targets. Just like the active ingredients in drugs, companies must be sure that these affinity ligands don't contain harmful impurities and fulfil product specifications in order to protect patient safety.

A common efficient method for determining protein concentration and purity is called high performance size exclusion chromatography, HPSEC. HPSEC is a method where a solution flows through a tube filled with porous particles called a column, which separates compounds according to how fast they travel through the column. How fast something travels depends on the size of the compound. Different affinity ligands have different properties which may impact how they interact with the column which may distort results. Therefore suitable parameter conditions must be determined for each affinity ligand to be measured. The aim of this study is to develop to methods by finding suitable HPSEC parameter conditions for determining purity and concentration of two affinity ligands, here called protein T and protein H. Both methods should be separate the affinity ligands well from impurities and be linear and repeatable.

To this end, four columns were tried; two different sizes for the porous particles, and two different tube lengths. Solutions with different salt concentrations were used to pass through the column as well as different times and temperatures of preparing the samples were tried so that the ligands were in their active states. Once suitable parameters had been selected, different protein amounts were used to see for what sample concentrations the method would be linear. The concentration and purity of a sample were determined on three different days using the final selected parameters to test if the method was repeatable.

The results showed the column with smaller particles were better at separating out the affinity ligand, and worked best in combination with the solution with a low salt concentration. The longer columns could separate more compounds, but took long time for each analysis (40 min vs 20 min) so the shorter column with small particles was selected. High temperature, 85°C, seemed to degrade both affinity ligands the longer they were reduced, while little difference were seen at lower temperatures, 30 and 60°C, no matter the time. Both methods were linear and gave repeatable results when using 2.43 – 16.98µg, 4.85 – 16.98µg and 2.56 – 15.37µg of

protein for purity determination of protein T, concentration determination of protein T and both purity and concentration determination protein H, respectively.

All in all, two methods, one for protein T and one for protein H, were developed that could be used for controlling the quality for later use in the pharmaceutical industry.

Abstract

Background: There is a need to develop methods of determining concentration and purity for two new affinity ligands, protein T and protein H, in order to ensure the quality of an industrial production of said proteins.

Aims: To develop a High Performance Size Exclusion Chromatography, HPSEC, method to determine concentration and purity of protein T and to improve an existing HPSEC method for concentration and purity determination of protein H that show selectivity, display linearity between peak area and injection amount and give repeatable results.

Methods: For protein T, Xbridge Premier Protein and Superdex Increase columns were used of the lengths 15 and 30 cm together with 50mM and 0.3M Sodium phosphate buffers and PBS and mobile phases. For protein H, only a 15 cm Superdex Increase column was used with 0.3M Sodium phosphate buffer. Different reducing conditions during sample preparation were tested. The linearity and measuring ranges were determined by injecting different volumes of reference standard and evaluating . Repeatability was tested with three triplicates of reference standards across three occasions with the same equipment and running conditions.

Results: For protein T the 15 cm Xbridge column with 50mM Sodium Phosphate buffer was selected. Little difference between reduction conditions with exception of reduction at high temperature which resulted in degradation or aggregation for both proteins. Acceptable selectivity and separation between main peak and impurities was demonstrated for both methods. Both methods were linear and gave repeatable results in the measuring ranges 2.43 – 16.98µg, 4.85 – 16.98µg and 2.56 – 15.37µg for purity determination of protein T, concentration determination of protein T and both purity and concentration determination protein H, respectively.

Conclusion: Two methods, one for protein T and one for protein H, were developed which met qualification criteria use in quality control.

Keywords: Method qualification, Protein concentration, Protein Purity, High Performance Size exclusion chromatography, Quality control

Acknowledgement

This thesis would not have been possible without the help of people around me and would like to express my gratitude to all of you now.

Thank you Peter Spégel for your support and guidance, especially through certain academic aspects of data analysis and report writing.

Thank you to the QC and PS department at Repligen Lund for helping me around the lab, letting me borrow computers you were using and giving good feedback on my presentation.

Thank you to all of Repligen Lund for letting me do my thesis with you and being lovely colleagues to be around for the last couple of months.

Finally, a special thank you Marika Slivacka for helping with everything from experiment setups and report writing to how to use a pipette according GMP standard but mostly for being an all-round supportive and kind supervisor.

Table of contents

1 List of abbreviations

BEH – Bridged ethylene hybrid

DAD – Diode array detection

DTT – Dithiothreitol

GMP – Good manufacturing practice

GLP – Good laboratory practice

HMW – High molecular weight

HPLC – High performance liquid chromatography

HPSEC – High performance size exclusion chromatography

LMW – Low molecular weight

MS – Mass spectrometry

PBS – Phosphate buffer saline

RSD – Relative standard deviation

SD – Standard deviation

TRIS – Tris(hydroxymethyl)aminomethane

2 Introduction

Repligen Sweden AB produces different recombinant proteins for use as affinity ligands in various therapeutic applications. Affinity ligands are coupled to gel packed columns and used for preparative purification of other proteins through affinity chromatography, for example monoclonal antibodies. Monoclonal antibodies are widely used as active pharmaceutical ingredients, in a range of treatments, for example vaccines and oral preparatives. Affinity chromatography is highly selective, but not very flexible since one affinity ligand can only bind one affinity binding site. Therefore new affinity ligands are continuously developed for research and development of new pharmaceuticals. To meet the need of flexibility from the pharmaceutical industry, Repligen AB must be able to produce a variety of affinity ligands and introduce new products at the request of customers. Two newer proteins, in this report referred to as protein T and protein H, are currently being produced at the request of two of Repligen AB's customers. These proteins are produced by separate bacterial strains in separate batches then recovered using micro filtration and purified using multiple chromatographic steps.

The affinity ligands are produced in a controlled environment providing the highest quality as the products are used by pharmaceutical companies which are working according to ICH guidelines, including good manufacturing and good laboratory practices (GMP and GLP). The quality is ensured using methods that are validated and approved for analysis of affinity ligands such as high-performance size exclusion chromatography, HPSEC [1].

Protein purity and concentration are very important parameters for proper protein function. Protein H has been produced at the Repligen factory and an HPSEC method for determining the concentration and the purity thereby exists. However the sample preparation is tedious due to requiring multiple dilution steps during sample preparation, and the analysis time is long due to the long column. A method to determine concentration and purity of Protein T has yet to be qualified for use in quality control.

For fast and cheap quantification, a direct spectrophotometric method may be used. Aromatic amino acid residues absorb light at around 280 or 260 nm. Peptide bonds found in the backbone of proteins, absorb light at around 214 nm. Peptide bonds are higher in abundance than the aromatic side chains and detection at 214 nm may therefore be used for a more sensitive assay or for proteins with few or no aromatic residues. However, any other analytes

or matrix components which absorb light at the chosen wavelength are indistinguishable from the analyte of interest. As a result, protein concentration cannot be measured reproducibly in most cases. [2] For higher selectivity, proteins may be separated using liquid chromatography in various modes. Reversed phase, ion exchange chromatography, hydrophobic interaction chromatography and hydrophilic interaction chromatography may be used to isolate peptides and proteins with differences in surface residue composition and with different post-translational modifications, for example aspartate isomerization or sugar addition [3] [4]. Multiple types of modifications will elute both before and after the main peak, therefore modification peaks are identified individually, commonly using mass spectrometry, MS. [4]

The most common mode of HPLC quantifying protein aggregation and fragmentation is HPSEC. In HPSEC, analytes are separated by differences in diffusion rates through porous particles due to differences in hydrodynamic radii. Ideally, there are no interactions between analyte and stationary phase or column walls [5] [3]. Conventional HPSEC columns are packed with polymer based particles, usually with a particle diameter in the range of 3-20μm, normally encased in polymer or glass columns with an inner diameter of 4.6-8mm. Such columns are sensitive to high pressures with pressure maxima of around 100 bar or less, which limits the flowrate and throughput, which are important for analysis costs. Silica based packings with hydrophobic functional groups or with non-aqueous organic phases have also been used for HPSEC, but produce risks of secondary interactions with potential bare silanol groups. Bridged ethylene hybrid organic/inorganic particles, BEH developed during the last decade are more mechanically durable than polymer-based particles, allowing for smaller particle diameters, 1.7-3 μm, that can withstand the resulting pressure. BEH columns are also less susceptible to silanol interactions with analytes than silica particles and are more chemically stable than silica particles as well. [5] [6]

Secondary interactions with the stationary phase and the column are especially prominent for steel columns. Ionic secondary interactions may be mitigated by increasing the ionic strength of the mobile phase, with increased risk of fouling which columns of smaller inner diameter and particles sizes are especially sensitive to. [5]

Once a method is developed, selectivity, linear range, measuring range, and accuracy must be validated for use in quality control. Selectivity is the ability of a method to distinguish analytes in matrices of impurities which give rise to similar responses [7]. In the case of HPSEC with Diode Array Detection, DAD, an important selectivity factor is the ability of the

column to resolve peaks. When peaks are unresolved, it is impossible to determine precisely how much of the light is absorbed by analyte and how much is absorbed by impurities. In the case of protein production, the most significant impurities to distinguish the proteins from are other biomolecules. These biomolecules may be of high molecular weight, HMW, or low molecular weight, LMW, where HMW species may arise from analyte aggregation and LMW species might be degraded analytes.

Linear range and measuring range is determined by injecting different amounts of analyte, either by using different sample concentrations or by using different injection volume, resulting in different mass loads. Current day chromatography systems use autosamplers which can inject different volumes with less variation compared to variation in dilutions performed by humans. The linearity can then be assessed via linear regression and residual analysis, where residuals are the difference between measured values used to make the regression and values predicted by the regression. The residuals should appear to be random with a constant variance across the measured interval, i.e. be homoscedastic, if an appropriate model has been applied.

Accuracy is the trueness and precision of a method where trueness is the closeness of many measurements to an accepted reference value and precision is the closeness of values from different test results, and may be investigated at different levels, repeatability, intermediary precision and reproducibility. For repeatability, the precision is determined from results obtained close in time from the same equipment and analyst, generally within the same assay. For intermediate precision, the analysis is repeated in the same laboratory, typically on different occasions by different analysts and using different equipment. Exactly which parameters are varied should be suitable for the intended end use of the method, for example, if different columns will be used, the intermediary precision should be tested using different columns of the same type. Reproducibility is the precision obtained using the same method in different laboratories. [7] No accepted reference values exist for neither protein T nor protein H samples, therefore the trueness of the methods developed here will not be evaluated.

Precision may be measured as the relative standard deviation, %RSD, which is calculated according equation (1) below,

(1)

$$\%RSD = \frac{100}{\bar{y}} \sqrt{\frac{(y_i - \bar{y})^2}{n-1}}$$

where y_i are individual values, $\bar{y}$ is the mean of individual values, and n is the number of individual values [9]. To ensure that the precision and repeatability does not deteriorate over time or is dependent on the operator of the experiment, one may test the intermediate precision, where the test is repeated at different times and by different operators [8].

The aim of the following study was to develop and validate a HPSEC method to determine concentration and purity of protein T and to improve and test an existing HPSEC method for concentration and purity determination of protein H in terms of time efficiency by testing different column types, mobile phases, sample preparation methods and injection amounts. The specific research questions of this thesis are:

- Find methods with a response linear to the concentration with a $R^2 \geq 0.99$ and homoscedastic residuals with standardized values of <2
- Determine the measuring range for purity and concentration
- The repeatability should give %RSD ≤ 1 for purity determination
- The repeatability should give an %RSD ≤ 3 for concentration determination of pure samples and %RSD ≤ 10 for concentration determination of impure samples.
- Determine selectivity of developed protein T and H methods: retention time variation between reference standard and in process controls (± 10%) and evaluate main peak separation from low molecular weight species and high molecular weight species.

3 Materials and Methods

3.1 Chemicals

The chemicals used and their purity and manufacturers are listed in **Table 1** below. All water used in buffer, sample and stock solution preparation was filtered through 0.22µm filter from Millipak (MPGP002AI) using a Milli-Q Water system from Merck (ZiQ7000T0).

Table 1. Chemicals used.

Chemical name	Chemical formula	Purity	Grade	Manufacturer	Catalog/reference number
Disodium phosphate	Na_2HPO_4	98-102%	-	Thermo Scientific	448140010
Sodium dihydrogen phosphate	NaH_2PO_4	99%	-	Acros Organics	389870010
Potassium Chloride	KCl	-	-	Merck	1.04935
Acetonitrile	CH_3CH	99.9%	HPLC	Fischer Chemical	A/0627/17X
Methanol	CH_3OH	≥99.9%	HPLC	Fischer Chemical	M/4058/PB17
Ethanol	CH_3CH_2OH	≥99.9%	HPLC, ACS	Scharlau	ET00151000
DTT	$C_4H_{10}O_2S_2$	≥99.9%	-	Fischer Bioreagents	BP172-25
Tris-(hydroxymethyl)-aminomethane	$NH_2C(CH_2OH)_3 \cdot HCl$	-	ACS reagent	Merck	1.08382
Tris-(hydroxymethyl)-aminomethane	$NH_2C(CH_2OH)_3$	≥99.0%	ACS reagent	Merck	1.08219
Hydrochloric acid	HCl	≥99.8%	ACS reagent	Merck	1.00063
Sodium hydroxide	NaOH	99-100%	ACS reagent	Merck	1.06469

3.1.1 Buffers and Solutions

The buffers and solutions used are listed in **Table 2** below.

Table 2. Buffers and solutions used.

Description	Use	Notes
0.3M Sodium phosphate buffer, pH 7	Mobile phase	pH adjusted to pH 7±0.15 using 1M HCl or 5M NaOH if needed
50mM Sodium Phosphate buffer, pH 7	Mobile phase	pH adjusted to pH 7±0.15 using 1M HCl or 5M NaOH if needed, filtered using 0.22µm filter
10% ACN 90% 25mM Sodium phosphate buffer, pH 7, with 100mM KCl	Storage and washing buffer for Xbridge columns	pH adjusted to pH 7±0.15 using 1M HCl or 5M NaOH if needed
20% EtOH in MQ-H_2O	Storage and washing buffer for Superindex Increase columns	-
PBS buffer with HCl, pH 3	Washing solution for Xbridge column	-
100mM DTT	Reducing Agent	-
100mM DTT with 20mM TRIS	Reducing agent	pH adjusted to pH 8±0.15 using 1M HCl or 5M NaOH if needed
15% Isopropanol	Needle wash	-

3.1.2 Protein samples

Protein T were produced at the process development department at Repligen AB and protein H were produced in the production department at Repligen AB according to the scheme in **Figure 1.** Samples are written next to the step which they were taken after. H-proc1 was taken after a third chromatography step not shown in the figure.

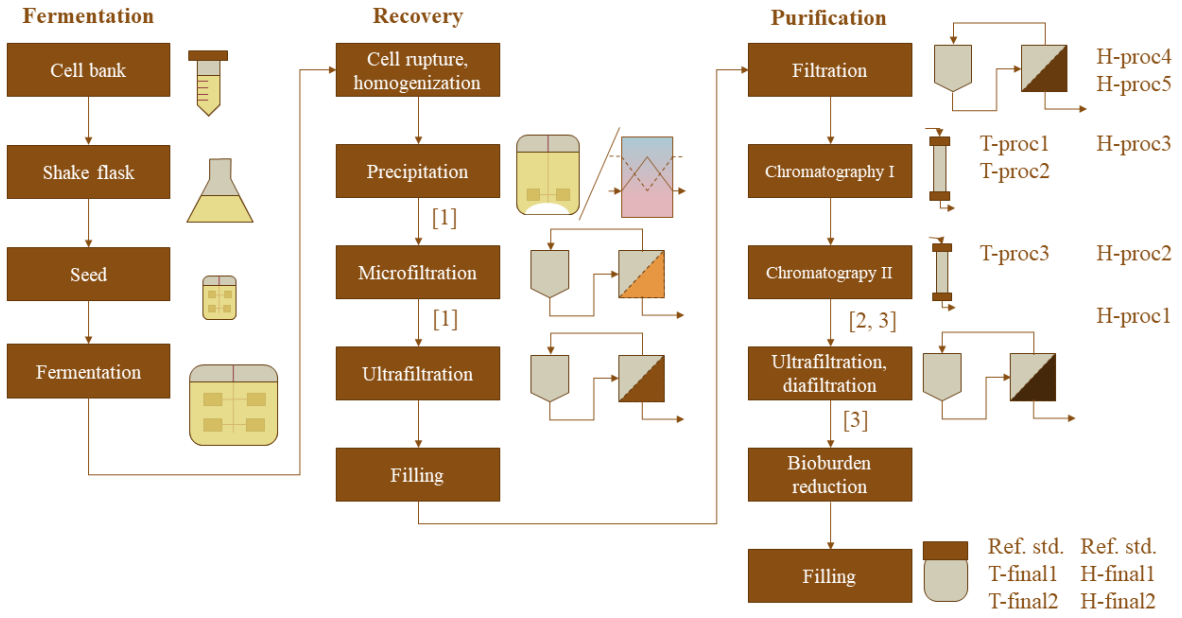


Figure 1. Overview of the protein production process and origin of samples used. More steps are needed for protein H production, marked by numbers [1-3] in figure, [1] Bioburden reduction steps [2] Chromatography III, anion exchange [3] Endotoxin reduction steps. H-proc1 is the eluate from the Chromatography III step.

Protein T reference standard was provided by the customer of Repligen AB, with a concentration of 9.7 mg/mL. T-final1, T-final2, T-proc1, T-proc2 and T-proc3 are all taken from separate batches. Samples of end product protein H produced by Repligen AB was used as protein H reference standard. From this reference standard control samples were prepared by diluting the reference standard x70 in 100mM DTT and aliquoted in 500µL aliquots and stored at -20°C. H-final1 and H-final2 were end product of protein H from a separate batches. H-proc1 and H-proc5 were from the same production batch, and H-proc2, H-proc3 and H-proc4 were from the same production batch but different from that of H-proc1 and H-proc5. All protein samples were stored at -20°C dissolved in water (except protein H control samples), and subjugated to a maximum of four freeze/thaw cycles.

3.2 Equipment

Two HPLC systems were used, which will be referred to as system I and system II. Both systems were from Agilent Technologies from the 1260 Infinity II series with a G7167 multisampler and a G4212-6008 DAD detection flow cell. System I used a G7116A MCT, a G7117C DAD HS mainboard and a G7112B binary pump. System II used a G7116B MCT, a G4212B DAD HS mainboard and a G71312B binary pump. System II was also equipped with a G4225 degasser. The software used was ChemStation.

The columns used were Superindex® Increase 75 5/150 GL by Cytiva (cat. GE29-1487-22), XBridge Premier Protein SEC Column, 250Å, 2.5 µm, 7.8 x 150 mm by Waters (cat. 186009961), and XBridge Premier Protein SEC Column, 250Å, 2.5 µm, 7.8 x 300 mm by Waters (cat. 186009962). For the repeatability tests a guard was used, MaxPeak Premier Protein SEC Guard, 250Å, 2.5 µm, 4.6 x 30 mm (cat. 186009969).

3.3 Methods

3.3.1 Peak Identification

At least one injection of reducing agent and one injection of non-reduced protein for both proteins were made. For protein T it was done using the 30cm Xbridge column with mobile phase 0.3M Sodium phosphate buffer, pH 7, on system I. For protein H it was done with the 15 cm Superindex Increase column with mobile phase 0.3M Sodium phosphate buffer, pH 7, on system I. The monomer peaks were identified as the peak present in both protein samples but not in the reducing agent sample and which increased in intensity for the reduced sample compared to the non-reduced samples.

3.3.2 Protein T

3.3.2.1 Column and mobile phase selection

While selecting column and mobile phase, Protein T reference standard, T-proc1 and T-proc3 samples were diluted to ca 1.5 mg/mL in 100mM DTT and 20mM TRIS and reduced at 85℃ for 15 min.

Table **3** shows the amount of replicates used for each condition along sample dilution factors, injection volumes and column temperatures. 214 nm was used as detection wavelength. The peaks corresponding to the monomer of the target protein were identified by comparison to a

negative control containing only reducing agent and comparison to a non-reduced samples of Protein T reference standard diluted in MQ-H_2O.

Table 3. Parameters used when selecting column and mobile phase.

Column	Mobile phase	Sample	Replicate preparation	Dilution factor	Injection volume	Column temperature
Superdex® increase, 15 cm	0.3M Sodium phosphate buffer, pH 7	Protein T reference standard	Three sample dilutions, single injections, two different days	x6	10µL	25°C
		T-proc1	Duplicate dilutions, single injections, same day	x8	10µL	25°C
Xbridge 15 cm	0.3 M Sodium phosphate buffer, pH 7	Protein T Reference standard	Single sample dilution, three injections, same day	x8	6µL	35°C
	10% ACN 90% 25mM Sodium phosphate buffer, pH 7 with 100mM KCl	Protein T Reference standard	Single sample dilution, single injection	x8	10µL	35°C
		T-proc1	Single sample dilution, single injection	x8	10µL	35°C
	50mM Sodium phosphate buffer, pH 7	Protein T Reference standard	Single sample dilution, single injection	x8	6µL	35°C
Xbridge 30 cm	0.3 M Sodium phosphate buffer	Protein T Reference standard	Single sample dilution, four injections, same day	x8	12µL	35°C
		T-proc1	Single sample dilution, four injections, same day	x12	12µL	35°C
	PBS	Protein T Reference standard	Single sample dilution, four injections, same day	x8	12µL	35°C
		T-proc3	Single sample dilution, four injections, same day	x10	12µL	35°C

3.3.2.2 Reduction

Suitable reduction times and temperatures were screened using single samples of Protein T reference standard. Two tests were performed. For the first test, the reference standard samples were diluted x8 in 100mM DTT with 20mM TRIS then reduced for 5, 10, 15, 25, and 30 minutes at 30, 60 and 85°C. Six µL of each sample was then injected onto the 15cm Xbridge Premier Protein column. For the second test, the reference standard samples were diluted x8 in 100mM DTT with 20mM TRIS then reduced for 5, 10, 15, 20, 25, and 30 minutes at 60°C. Twelve µL of each sample was then injected onto the 30cm Xbridge Premier Protein column. For both tests 50mM sodium phosphate buffer, pH 7, was used as the mobile phase and a flow rate of 0.5mL/min. The column temperature in both tests was 35°C.

3.3.2.3 Linearity and measuring range

Two 7-point calibration curves were made from two tests, where for both tests a single sample of reference standard was diluted x8 in 100mM DTT with 20mM TRIS and reduced at 60°C for 15 min. For both tests, the 15 cm Xbridge Premier Protein column was used on HPLC system I with 50mM sodium phosphate buffer as mobile phase at a flow rate of 0.5mL/min and a column temperature of 35°C. For the first test the injection volumes were 1, 2, 4, 6, 8, 10, 12µL. As a control, a second sample of reference standard was prepared for the first test by diluting x16 in 100mM DTT with 20mM TRIS and reducing at 60°C for 15 min, and then injected in volumes of 1, 2, 12, and 14µL. For the second test, only the x8 dilution was used and the injection volumes were 2, 4, 6, 8, 10, 12, and 14µL. A linear regression was made between the peak areas obtained and the mass loads and the residuals between measured values and values predicted by the regression were calculated using Microsoft Excel data analysis -> Regression function. The residuals were standardized by the same tool by dividing by the square root of the predicted values.

For routine use, the method developed in this thesis will be used with a fixed injection volume, but on samples of varying concentration. To estimate the size of potential errors in the concentration determination, theoretical and predicted concentrations were calculated for the injections made. Theoretical concentrations based on the mass load and the future intended injection volume for the method was calculated, according to equation (2)

(2)

$$Theoretical\ concentration = \frac{Injected\ volume}{Method\ injection\ volume} \cdot Concentration\ of\ injected\ sample\ .$$

Predicted concentrations were calculated using the slope and intercept from the 7-point regression as shown in equation (3),

(3)

$$Predicted\ concentration = \left(\frac{Main\ Peak\ Area - Intercept\ of\ regression}{Slope\ of\ regression}\right) Method\ injection\ volume.$$

The difference between theoretical and predicted concentration were then calculated as a percentage as shown in equation (4),

(4)

$$Difference\ theoretical\ vs\ predicted = \frac{Predicted\ concentration}{Theoretical\ concentration} \cdot 100 - 100.$$

Potential errors in the concentration determination from using one point calibration were also estimated by predicting concentrations using the main peak area obtained for the method injection volume as shown in equation (5)

(5)

$$Predicted\ concentration\ 1-point = \frac{Area}{\left(\frac{\text{Main peak area for Method injection volume}}{\text{Theoretical concentration for Method injection volume}}\right)}$$

The error was estimated by calculating the difference between the theoretical and concentration predicted by the one point calibration as shown in equation (6)

(6)

$$Difference\ theoretical\ vs\ predicted = \frac{Predicted\ concentration\ 1-point}{Theoretical\ concentration} \cdot 100 - 100$$

3.3.2.4 Repeatability

Triplicate injections of protein T reference standard, T-final1, T-final2 and T-proc2 were made to determine the repeatability of the method. All samples were diluted in 100mM DTT with 20mM TRIS and reduced at 60°C for 15 min, with a x8 dilution for protein T reference standard, x30 dilution for T-final1, x30 dilution for T-final2 and x12 dilution for T-proc2. Two tests were performed using the 15 cm Xbridge Premier Protein column on HPLC system I with 50mM sodium phosphate buffer as mobile phase at a flow rate of 0.5mL/min and a column temperature of 35°C. The injection volume was 8µL and the detection wavelength was 214 nm for both tests. The protein T reference standard results were used to calibrate the concentration measurements by linear regression between the average of the main peak areas and a forced point through origin. The repeatability test was also repeated twice with triplicate dilutions of T-proc1 on separate days. Triplicate dilutions of protein T reference standard were used as 1-point calibration with the regression line forced through origin on the separate days as well.

3.3.3 Protein H

3.3.3.1 Reduction

Two tests were performed: one to identify a suitable DTT concentration agent for reduction, one to identify a suitable reduction temperature and time. For the first test, reducing agent was diluted in series from 100mM DTT to 50, 25 and 12.5mM DTT. Protein H reference standard was diluted x40 in triplicates for each concentration of DTT, 100, 50, 25 and 12.5mM as well

as one triplicate using 100mM DTT supplemented with 20mM TRIS. All samples were reduced for 30 min at 40℃. For the second test protein H reference standard samples were diluted x40 in 100mM DTT then reduced for 5, 10, 15, 20, 25, and 30 minutes at 30, 60 and 85℃. 8µL of each replicate or sample was then injected onto a 15cm Superdex Increase column. 0.3M Sodium phosphate buffer was used as mobile phase with a flow rate of 0.5mL/min. The column temperature was 25℃ and the detection wavelength was 214nm. The injection volume for the control sample was also 8µL. For the first test, system I was used, for the second test, system II was used.

3.3.3.2 Linearity and measuring range

Two triplicates of protein H reference standard sample were diluted x40 in 100mM DTT then reduced for 10 min at 60℃ and 30℃. Samples injected onto a 15cm Superdex Increase column connected to system I in volumes of 1, 2, 4, 6, 8, 10 and 12µL. Triplicates were injected for 1, 6, 8, and 12µL, single samples were injected for 2, 4 and 10µL. 0.3M Sodium phosphate buffer was used as mobile phase with a flow rate of 0.5mL/min. The column temperature was 25℃ and the detection wavelengths were 214 and 280nm. Linear regression, residual analysis and measuring range determination was done in the same way as for protein T, see section *3.3.2.3 Linearity and measuring range*.

3.3.3.3 Repeatability

Triplicate injections of protein H reference standard, H-final1, H-final2, H-proc2, H-proc3 and H-proc4 were made to determine the repeatability of the method. All samples were diluted in 100mM DTT and reduced at 60℃ for 10 min, see dilution factors in **Table 4** below.

Table 4. Dilution factors used for repeatability test for protein H method.

	Protein H reference standard	H-final1	H-final2	H-proc1	H-proc2	H-proc3	H-proc4
Dilution Factor	40	40	25	20	2	2	10

The 15 cm Superdex Increase column was used on HPLC system I with 0.3M sodium phosphate buffer as mobile phase at a flow rate of 0.5mL/min and a column temperature of 25℃. The injection volume was 6µL and the detection wavelength was 214 nm. The repeatability test was also repeated twice with triplicate dilutions of H-final1 and reference standard on separate days. Protein H reference standard results were used to calibrate the concentration measurements by linear regression between the average of the main peak areas and a forced point through origin.

4 Results and discussions

4.1 Peak identification and selectivity

Monomer peaks were identified as described in materials and method in section *3.3.1 Peak Identification*. Representative chromatograms for protein T can be seen in **Figure 2**. Two peaks are present around 24 to 26 min in the DTT blank and the reduced sample, but not in the non-reduced sample where DTT is not present. These were therefore determined to correspond to DTT. Apart from the DTT peaks, two prominent peaks were present in the non-reduced sample, and one prominent peak was present in the reduced samples. The prominent peak in the reduced sample eluted at the same retention time as one of the non-reduced peak, but was higher in intensity, while the other prominent peak in the non-reduced sample was not present in the reduced sample. As such, the monomer protein peak was determined to be the prominent peak in reduced samples which did not correspond to DTT. The same pattern could be seen for protein H but dimer/trimers were not resolved from the monomer, see figure below.

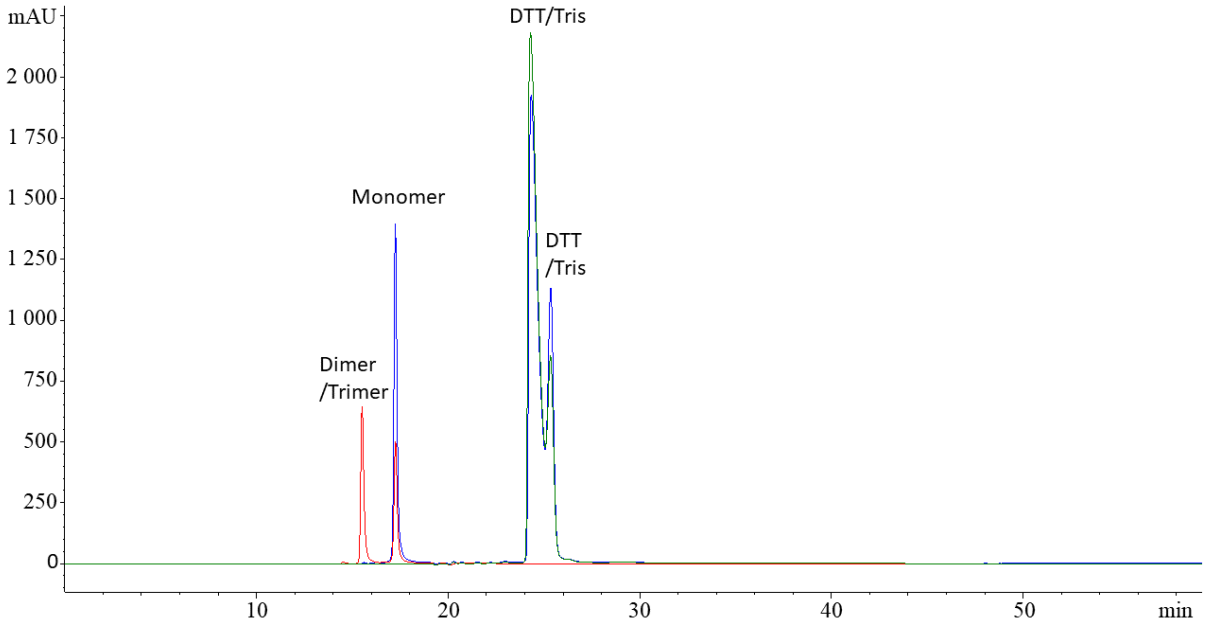


Figure 2. Representative example chromatograms of reduced (blue) and non-reduced (red) protein T reference standard as well as only reducing agent (green). The column used was the Xbridge Premier Protein 30cm column with 0.3M Sodium phosphate buffer, pH 7, as mobile phase. 14.55µg of protein was injected for both samples. The reduced sample was reduced at 85℃ for 15 min in 100mM DTT with 20mM Tris.

Low molecular weight, LMW, impurities can be clearly seen to the right of the main peak in the impure samples of both proteins, see **Figure 3**, and it is therefore assumed that LMW species may be present in the pure sample as well.

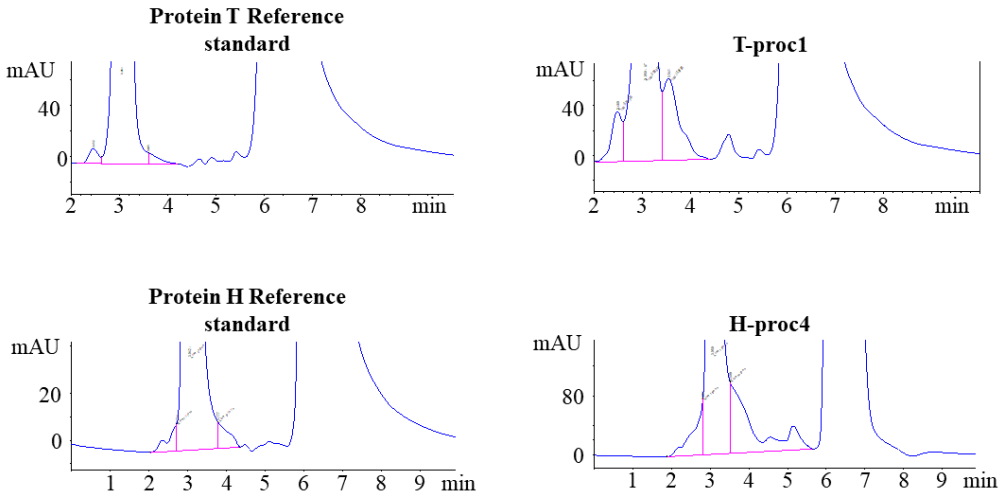


Figure 3. Chromatograms of reference standard and impure protein samples, T-proc1 and H-proc4. 7.686µg of protein H reference standard and 5.19µg of H-proc4 was used. 16.17µg of protein T reference standard was used and 21.25µg of T-proc1 was used. Both tests were performed using a 15 cm Superdex Increase column with 0.3M sodium phosphate buffer, pH 7, flow rate of 0.5mL/min, column temperature of 25°C and detection wavelength of 214nm

Peaks were integrated following the baseline, often at a slight angle aiming at the baseline between peaks from the DTT as DTT appears to raise the baseline slightly, see the horizontal pink line in **Figure 4**. The integrated peaks were then split to exclude unresolved high molecular weight, HMW, and LMW impurities see vertical pink line in **Figure 4**. The peak is split on the LMW side at around the same retention time as where the bump is clear in the impure sample **Figure 3**. To unify LMW specie estimation, automatic integration should be encouraged.

The protein H dimer/trimers were much less resolved from the monomer than for protein T, see **Figure 4** below. The dimers or trimers however formed distinguishable bumps on the front of the monomer peak, enabling sufficiently consistent splitting of the protein peak. The chromatographic profiles for both proteins are comparable to other proteins produced at Repligen AB, where robust and reliable methods for concentration and purity determination are in place.

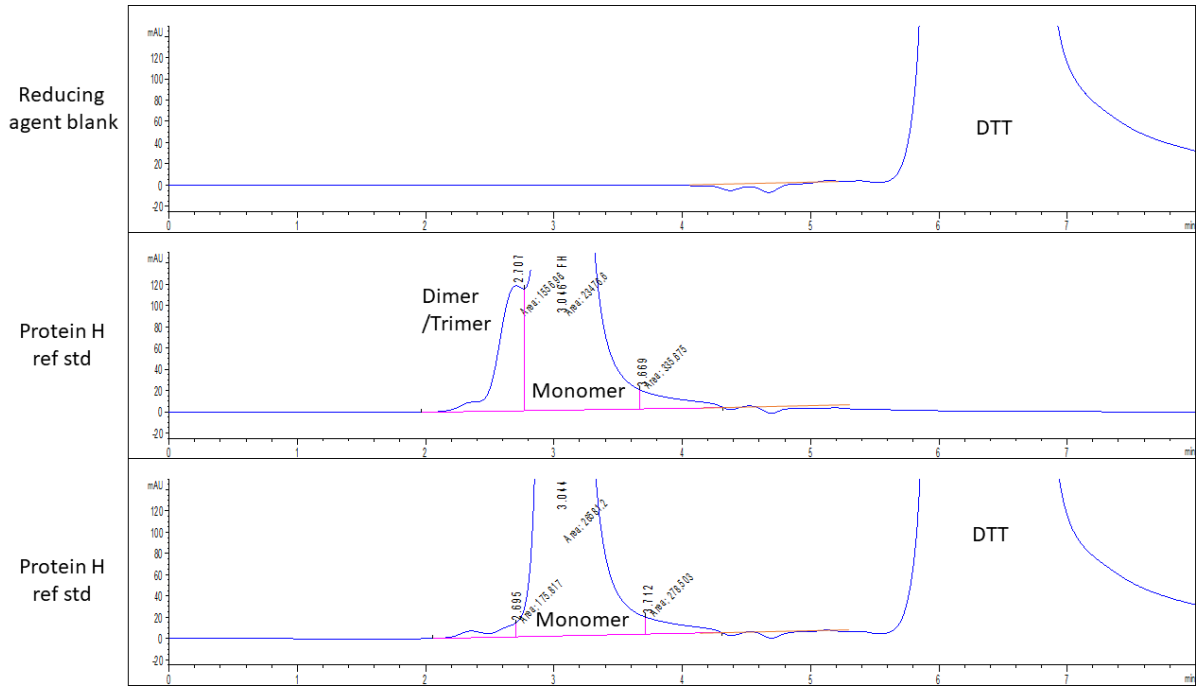


Figure 4. Representative example chromatograms of reduced (bottom) and non-reduced (middle) protein H reference standard as well as only reducing agent (top) showing how chromatograms were integrated. Integration baselines and peak splits are shown in pink. The orange line shows how the baseline is interpreted after the main peak for the integration. The column used was the Superdex Increase 15 cm column with 0.3M Sodium phosphate buffer, pH 7, as mobile phase. 10.24µg of protein was injected for both samples.

An appropriate detection wavelength for protein T was screened from 190 nm to 400, an isoplot around the monomer peak of protein T reference could be seen in **Figure 5** below where different absorbances in terms of mAU for different wavelengths and elution times are shown in a colour scale from blue, low absorbance, to red, high absorbance. As expected of proteins, two maxima are seen around 214 and 280 nm. The maximum around 214nm appears significantly more intense than that at 280nm, as expected since peptide bonds are more abundant than aromatic amino acid in proteins in general. The presence of the 280nm maximum shows that there are aromatic amino acids present in protein T. No significant impurities were detected at higher wavelengths, indicating non-peptide impurities are not undetected around this retention time with single wavelength settings. This implies good selectivity in the method, but it cannot be guaranteed that no other species are eluting at the same time with absorbances around 214nm. Going forward 214nm was used for its sensitivity. The same trend were seen for T-proc1 and T-proc3.

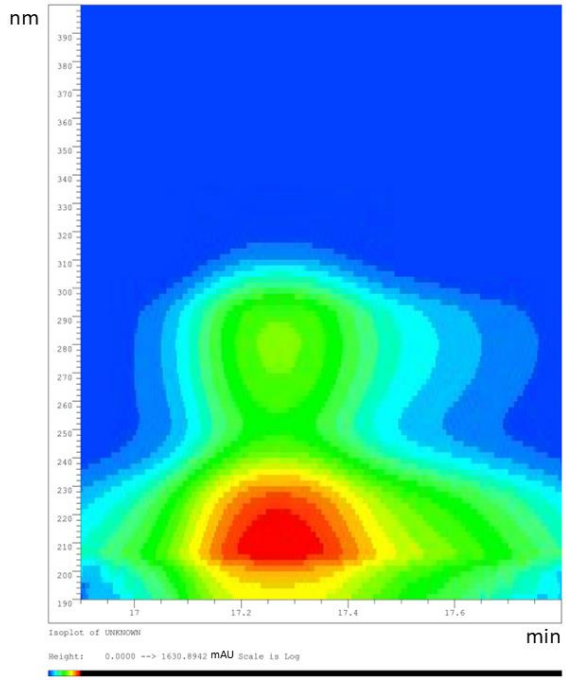


Figure 5. Isoplot protein T ref std. around main peak. The recorded wavelengths were 190-400nm. The mobile phase was 0.3M Sodium phosphate buffer with a flow rate of 0.5mL/min. The 30cm Xbridge column was used with a column temperature of 35°C and the injection amount was 14.55µg. The sample was reduced at 85°C for 15 min red (100mM DTT 20 mM TRIS).

No spectrum was measured for protein H but some test were made with detection at 280 nm along with 214 nm. The response signal in terms of mAU for protein H was much lower at 280nm than 214nm, same as for protein T.

4.2 Column and mobile phase selection

Due to time constraints, different columns and mobile phases were only investigated for protein T. Columns and mobile phases were evaluated for their selectivity in terms of resolution between main peaks and HMW and LMW species since high resolution implies little overlap between main peaks and impurity peaks. Due to the low resolving power of size exclusion chromatography, the resolution factor, which is normally used to quantify resolution, was not possible to calculate. Instead plate numbers were used to give an indication of resolving power along with visual evaluation of the chromatographic profile. A clear increase in plate number can be seen when using Xbridge Premier columns rather than the previously used 30cm Superdex Increase columns, see **Table 5**. This is expected as the Xbridge Premier columns have a smaller particle diameter than Superdex Increase, 2.5µm compared to 9µm, which decreases zone broadening from multiple flow paths. Less zone broadening and overlap between the main peak and impurities corresponds to higher selectivity.

Lower ionic strength appears to be preferrable for the Xbridge columns compared to the Superdex columns, as pH 7 and 50mM Sodium phosphate buffer, pH 7 gave lower plate numbers than both 0.3M Sodium phosphate buffer, pH 7 and PBS, pH 7.4. This implies other affects are more significant than interactions with stationary phase or column hardware or that 50mM sodium phosphate is sufficient to limit these interactions for the Xbridge column. Salt precipitation or salt interactions with the stationary phase might be issues for the Xbridge columns 0.3M sodium because pressure rose to the maximum recommended by the column manufacturer during the 0.3M sodium phosphate buffer tests and backflushing of the column was required to restore function.

Table 5. Analysis time, average plate number and average tailing factor for different columns and mobile phase combinations. Results are given based on tests with protein T reference standard reduced in 100mM DTT with 20mM TRIS at 85℃ for 15 min, see **Table 3** for further details. Plate numbers and tailing factor were calculated in Chemstation using the extended statistics of the report functions. The tangent method was used to calculate plate numbers. Averages were calculated from all protein T reference standard injections made and the error given corresponds to two standard deviations.

Column	Mobile phase	Analysis time (min)	Average plate number	Tailing Factor
Superdex Increase 30 cm	0.3M Sodium phosphate buffer, pH 7	40	4 450±32	1.088±0.0044
Superdex Increase 15 cm	0.3M Sodium phosphate buffer, pH 7	10	1 160±3	1.04±0.0040
Xbridge Premier Protein 30 cm	0.3 M Sodium phosphate buffer	40	55 000±12 000	1.36±0.1160
	PBS	40	28 000±5 400	1.58±0.0435
Xbridge Premier Protein 15 cm	0.3M Sodium phosphate buffer, pH 7	20	25 700±290	1.34±0.0059
	50mM Sodium phosphate buffer, pH 7	20	28 000*	1.12*

*No standard deviation shown due to only one data point being collected.

The purity appeared to be the similar across the columns for the reference standard (Superdex Increase 30cm: 98.9% Superdex Increase 15cm: 99.0% Xbridge 30cm: 98.2% Xbridge 15 cm: 97.9%), while purity decreased with longer columns and smaller particle size for the impure sample Tproc1 (Superdex Increase 30cm: 91.5% Superdex Increase 15cm: T-proc1 93.5% Xbridge 30cm: T-proc1 87.7%). This suggests that the selectivity of the Xbridge 30cm column is better than that of the Superindex Increase columns which may not resolve impurity peaks as well as the Xbridge 30 cm column and thereby overestimate the purity of the sample.

The consistent purity of the reference standard shows that it contains no significant undetected impurities and is a reliable reference, as expected.

The behavior and selectivity of the different columns can be seen more clearly in the chromatograms presented in **Figure 6** below, most clearly illustrated by the chromatograms for the impure sample T-proc1. Three HMW peaks which can be distinguished in the 30cm Superdex Increase column are all indistinguishable for the 15 cm Superdex increase column, a clear illustration of the drop in resolution and plate number when reducing the column length. Meanwhile, there are four distinguishable HMW peaks in the chromatogram from the 30 cm Xbridge column. The same trend can be seen on the LMW side, with more distinguishable peaks for the longer columns compared to the shorter, and more distinguishable peaks for the Xbridge columns compared to the Superdex Increase columns. The differences in separation are not as evident for the reference standard, but the a peak on the HMW side for the Superdex Increase columns appear as two peaks in the Xbridge columns.

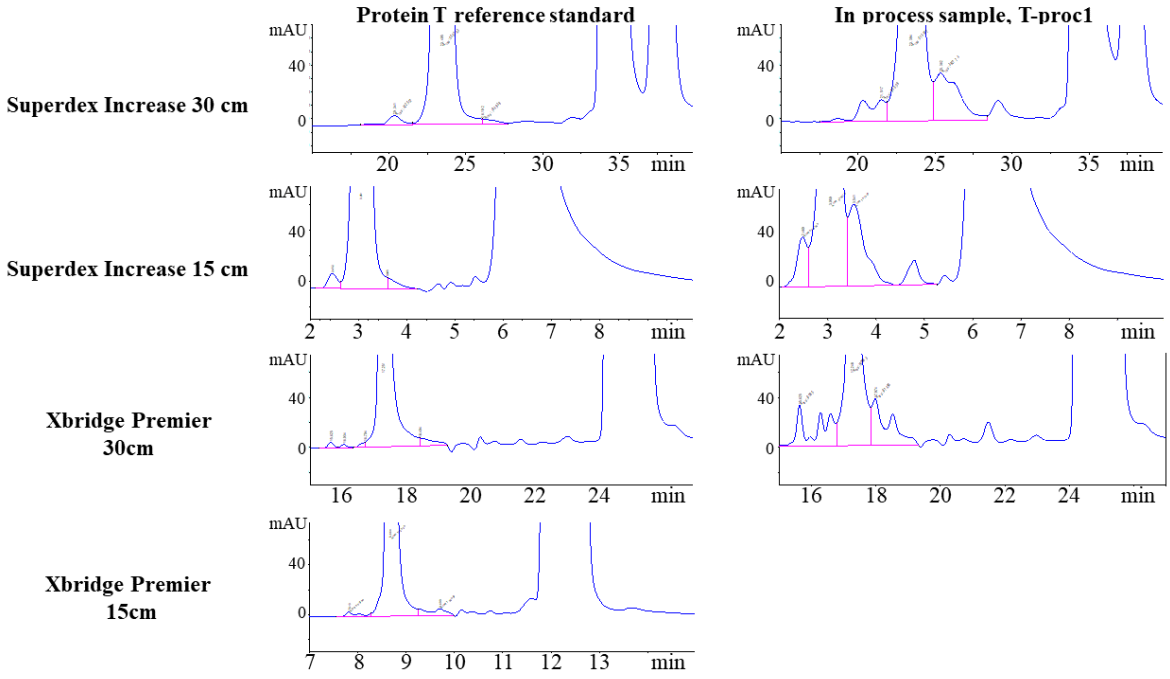


Figure 6. Representative chromatograms from each column tested with protein T. 0.3M sodium phosphate buffer was used as mobile phase with a flow rate in all cases. The column temperature was 25°C for Superindex columns, while 35°C was used for Xbridge columns. The injection amounts were as follows: Superdex Increase 30cm: 48.50 µg reference standard and 63.75µg T-proc1. Superdex Increase 15cm: 16.17µg reference standard and 21.25µg T-proc1 sample. Xbridge 30cm: 14.55µg of reference standard and 17.00µg of T-proc1. Xbridge 15 cm: 7.28µg reference standard.

4.3 Selectivity

The selectivity is evaluated based on a visual evaluation of the separation between monomer and impurities, as well as the difference in retention time of the monomer between pure

samples and impure samples. The separation between monomers and impurities are discussed at length in section 4.1 and 4.2. Little difference was seen in retention time between pure and impure samples, see table below. Up to 10% variation in retention time would be acceptable, but only -0.00733% and 0.208% was seen for protein T and protein H respectively. The retention of protein H may have been more affected by impurities since a column with lower resolution was used, meaning the monomer and impurities are less separated from each other. Such an affect would affect the retention time of impure sample more since more impurities are present.

Table 6. Selectivity of both methods in terms of difference in retention time between pure and impure sample. Data is taken from first repeatability tests for both methods.

Replicate	Reference standard T (min)	T-proc2 (min)	Replicate	Reference standard H (min)	H-proc4 (min)
1	8.78226	8.78228	1	3.044	3.049
2	8.78185	8.78199	2	3.042	3.05
3	8.7826	8.78051	3	3.043	3.049
Average (min)	8.782237	8.781593	**Average (min)**	3.043	3.049333
SD (min)	0.000307	0.000775	**SD (min)**	0.000816	0.000471
%RSD*	0.003491	0.008827	**%RSD***	0.026832	0.015459
Difference (min)**	-	-0.00064	**Difference (min)****	-	0.006333
Difference (%)***	-	-0.00733	**Difference (%)*****	-	0.208128

* $\%RSD = \frac{\text{SD}}{\text{Average}} \cdot 100$

** $Difference = Retention\ time_{Reference\ Standard} - Retention\ time_{impure\ sample}$

*** $Difference = \frac{Retention\ time_{impure\ sample}}{Retention\ time_{Reference\ Standard}} \cdot 100 - 100$

4.4 Reduction

The purified protein in solution is not the final product, but the protein is coupled to a column in reducing conditions. To emulate this environment and to measure concentration of active protein, in the monomer form, the protein samples are reduced.

The first reduction method test for the protein T method, showed that while the highest purity could be obtained at 60℃, the differences in purity between different reduction times and temperatures were small, as can be seen in **Figure 7A**. Only 25 min and 30 min at 85℃ were noticeably lower in purity, at 94.7% and 94.2% respectively. For the 85℃ samples, the additional impurities appeared to be HMW species as these were 3.9% and 4.4% for 25 min and 30 min respectively, compared to the 1.8±0.05% HMW species at 30℃ and 1.2±0.04% at 60℃ (average and standard deviation calculated from all reduction times for each temperature), see **Figure 7B**. Raised amounts HMW species may be due proteins unfolding

from the high temperature, resulting in a larger hydrodynamic radius without a change in mass.

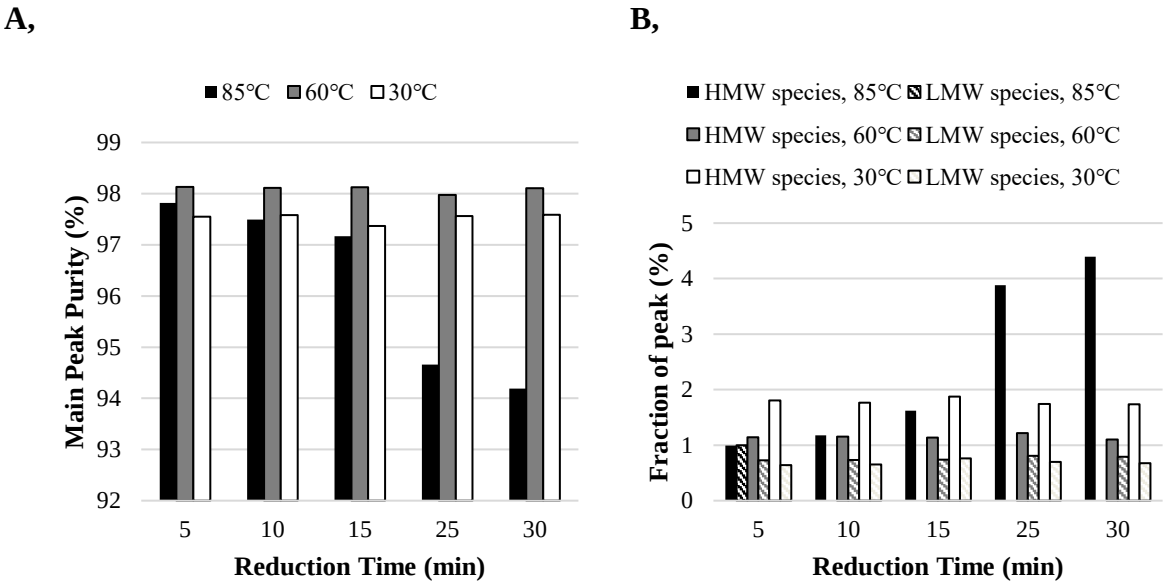


Figure 7. **A,** Monomer purities of protein T reference standard in different reduction conditions. **B,** Impurities of protein T reference standard in different reduction conditions. 100mM DTT with 20mM TRIS was used as reducing agent, results were obtained using Xbridge 15cm column with 50mM Sodium phosphate buffer, pH 7, as the mobile phase. Each bar corresponds to a single injection from a single protein sample.

A similar trend could be seen for protein H, where the reduction time and temperature did not appear to affect the purity when reducing at 40 and 60°C, but the purity decreased as reduction time increased when reducing at 85°C, see **Figure 8**. However, for protein H, the LMW species increased as monomer purity decreased rather than the HMW species as for protein T. This suggests that there might be some degradation at high temperatures.

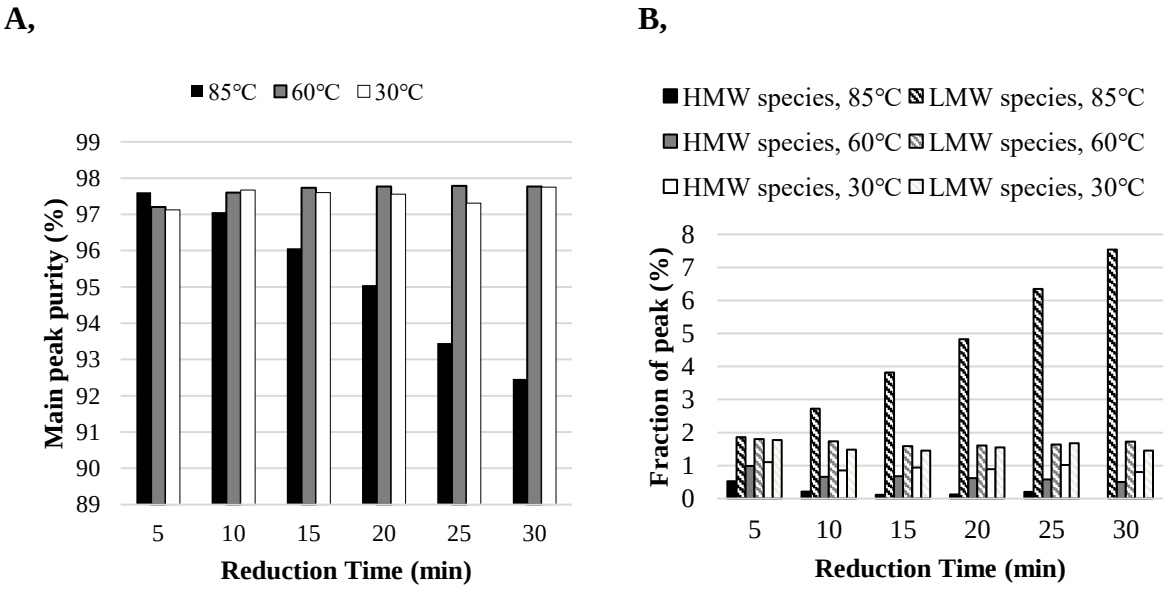
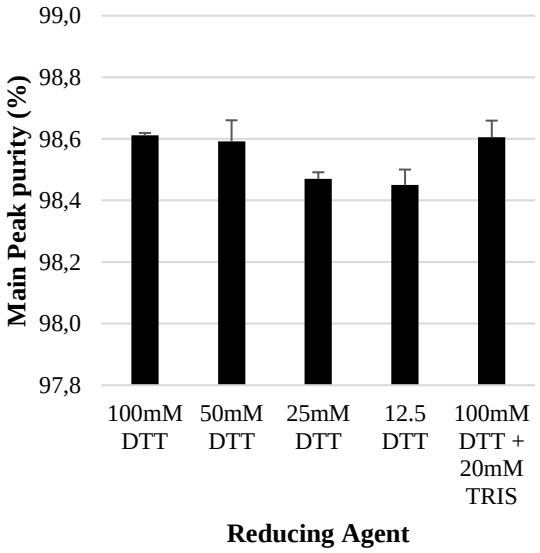


Figure 8. **A,** Monomer purities of protein T reference standard in different reduction conditions. **B,** Impurities of protein T reference standard in different reduction conditions. All samples were reduced in 100mM DTT, results were obtained using Superindex Increase 15cm column with 0.3M Sodium phosphate buffer, pH 7, as the mobile phase. Each bar corresponds to a single injection.

Different concentrations of reducing agent were also tried for protein H, also with little to no difference in the resulting purity, see **Figure 9** below. A slight decrease in purity and increase in HMW species could be seen for the lower concentrations of DTT, perhaps suggesting incomplete reduction of the sample. These samples were reduced at a lower temperature, 40℃ for 30 min, for future studies one may investigate if sufficient reduction can be achieved with lower concentrations of DTT if the temperature is raised. The addition of TRIS did not appear to affect the result either.

A,



B,

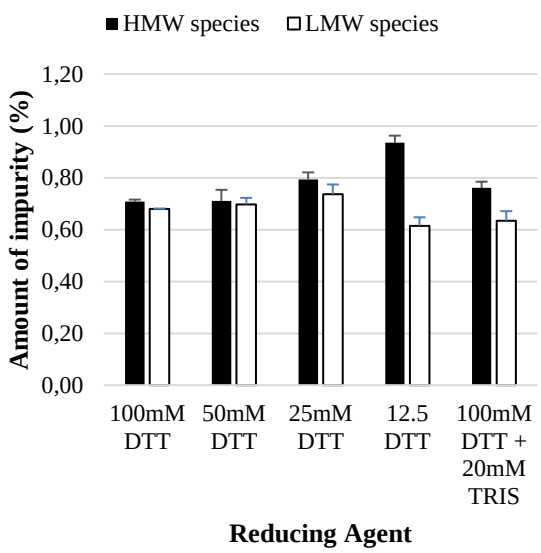


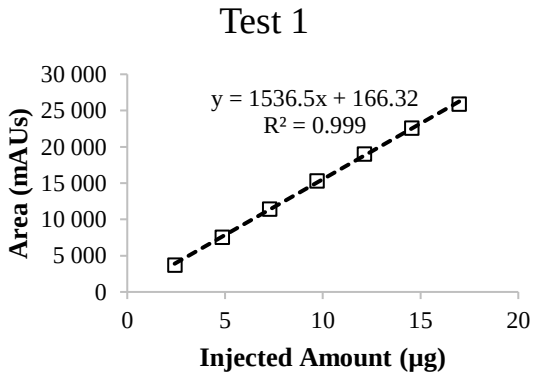
Figure 9. **A,** Monomer purities of protein H reference standard in different reduction conditions. **B,** Impurities of protein H reference standard in different reduction conditions. All samples were reduced for 30 min at 40℃, results were obtained using Superindex Increase 15cm column with 0.3M Sodium phosphate buffer, pH 7, as the mobile phase. Each bar corresponds to triplicates, and the error bar corresponds to the standard deviation for the given triplicate.

While both proteins appear to favor similar reduction conditions, 60℃, 10 or 15 min, such is not the case for all proteins. Other methods at Repligen both use higher and lower temperatures and longer and shorter times for reduction. Protein T and protein H are both relatively small, <15 kDa, so cysteine residues are likely to be relatively exposed and easily reduced compared to larger proteins.

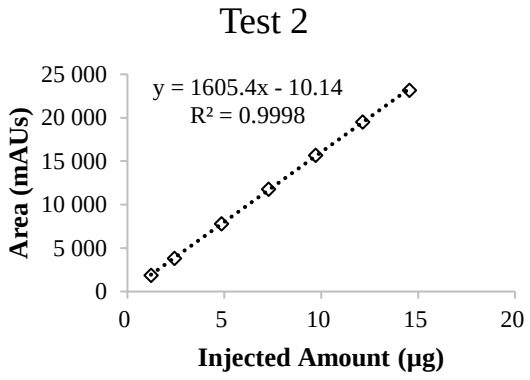
4.5 Linearity

One of the aims of the study was to establish that the developed methods showed a linear response in relation to the amount of injected protein with an R^2 value greater than 0.99. Both linearity tests, shown in **Figure 10**, showed the response to mass load relationship to be linear with a R^2>0.99 as per the requirements for validation (R^2=0.999 and 0.9998 for test 1 and 2, respectively). The linear regressions appear to underestimate mass load at low and very high mass loads, as well as overestimating mass load in the middle range, as can be seen by the resulting standardized residuals, see **Figure 10**C, perhaps suggesting a non-linear relationship between mass load and main peak area. However, the residuals are all relatively small, <±2, and comparable to residuals for models for other proteins. The inaccuracy model is therefore considered negligible for the intended purpose. Homoscedasticity of the protein T data was not determined since no triplicates were made.

A,



B,



C,

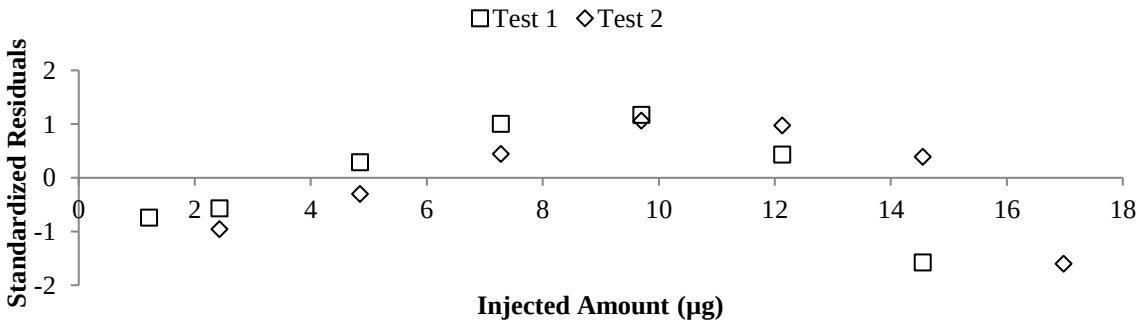


Figure 10. Results from linearity tests for protein T. Each data point corresponds to a single injection of protein T standard reference sample. **A.** Raw data from linearity test 1 with fitted linear regression**. B.** Raw data from linearity test 2 with fitted linear regression. **C.**Standardized residuals obtained from linear regression.

Similar results could be seen for protein H, peak areas showed a linear relationship to injected amount R^2=0.9993 for the sample reduced at 30℃ and R^2=0.9995 for the sample reduced at 60℃. The residuals appear more random than for protein T, but with less variation for the lowest injection amount, 1.28µg. Nevertheless, residuals are still within the acceptable range of <±2.

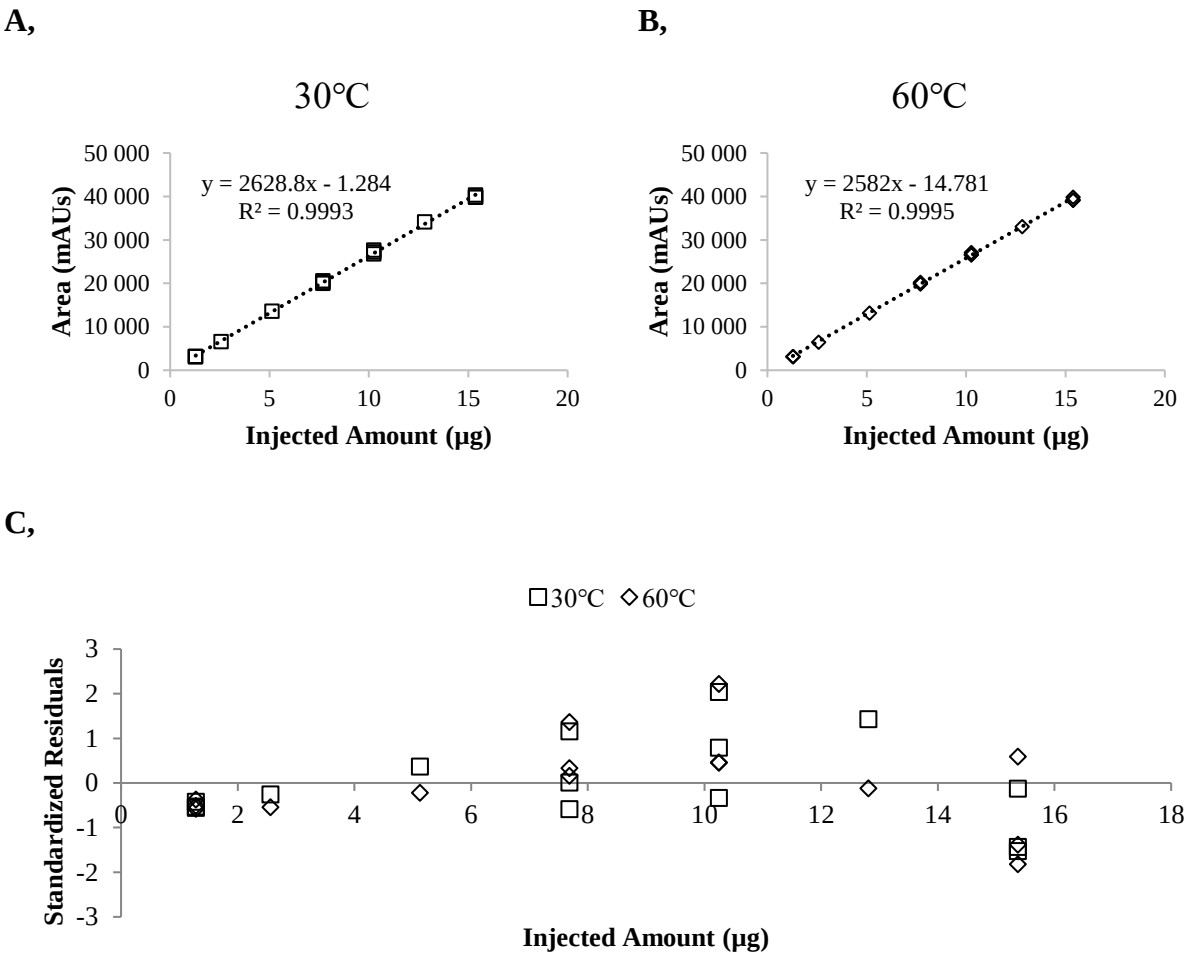


Figure 11. Results from linearity tests for protein H. Each data point corresponds to single samples. **A**. Raw data from sample reduced at 30℃ with fitted linear regression. **B.** . Raw data from sample reduced at 60℃ with fitted linear regression. **C.**Standardized residuals obtained from linear regression.

Triplicates where made of 1.28, 7.69, 10.25 and 15.37µg protein H injections and F-tests were used to determine whether the data was homoscedastic or not. The data was homoscedastic for the 7.69, 10.25 and 15.37µg injections with F-test results ≤2.3, below the value of 39.00 corresponding to two degrees of freedom for two data sets in the F-distribution. For 1.28µg the F-test against the other three injection amounts all showed heteroscedasticity with 144,

260 and 113 for 7.69, 10.25 and 15.37µg respectively. Linearity can therefore not be guaranteed for injection of less than 7.69µg of protein H.

4.6 Measuring range

The tests of the measuring range for purity showed that differences in mass load resulted in little variation in measured purity for both proteins, see **Figure 12**. Lower mass loads might result in underestimation in purity as the result for ≤5µg showed a purity which deviated more than method criteria of one percent unit from the mean, for both proteins. For purity measurements the lower limit of the measuring range should therefore be 1.21µg and 1.28µg for protein T and protein H, respectively. No upper limit for consistent purity results could be seen in these tests, and therefore the upper limit for the methods are simply set to the highest injection amount used here, 16.98µg and 15.37µg for protein T and protein H, respectively. Interestingly, no difference could be seen in purities between 30°C and 60°C reduction across different injection amounts for protein H.

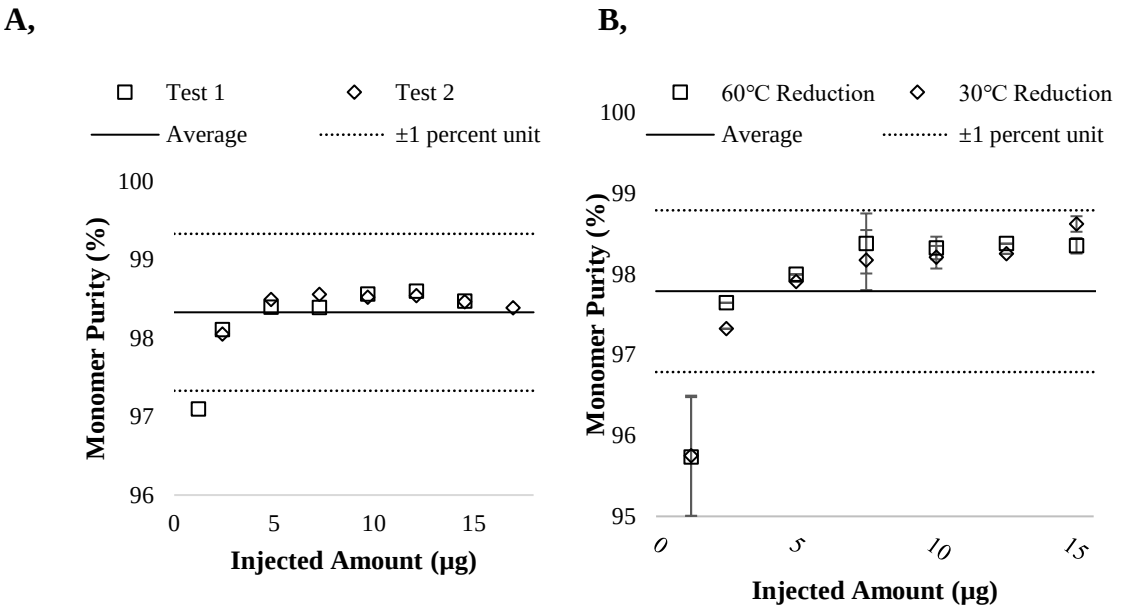


Figure 12. Purity of protein reference standards for different injected amounts. **A.** Purities for protein T. Seven injections, each of different volumes were made from single samples from two separate tests. The average and standard deviations were calculated from all 14 injections. **B**. Purities for protein H.

In order to evaluate the measuring range for concentration determination, the difference between theoretical concentrations injected on the column and concentrations predicted by the model was calculated as a percentage. The method should not over- or underestimate the concentration by more than 5%. For the protein T method, the percentage difference fell

within the interval -5% to 5% for all injection amounts except 2.43µg where the percentage difference was -5.5%.

For repeatability, concentration determinations were done using one point calibration forced through the origin since one point calibrations allow for an easier daily application in a quality control department.

This was done after differences between theoretical concentrations and concentrations predicted by one point calibration from the linearity test had been analysed, shown in **Table 7**. The limits for over and underestimation were set to ±5% for one point calibration. All measured points for protein T fell within these limits, indicating that one point calibrations forced through give acceptable results within the determined measuring range. Indeed, for injection amounts 7.28µg to 12.13µg the deviation of the predicted concentrations from the theoretical concentrations are smaller for the one point calibration than for the seven point calibration. However, one point calibration is very susceptible to errors, therefore control samples and system test runs should be included during analysis to capture any deviations.

Based on both the one point and seven point calibrations, the measuring range for the concentration determination of protein T was set to 4.85µg to 16.98 µg.

Table 7. Theoretical and predicted concentration protein T and percentage difference between theoretical and predicted for both 7-point and 1-point calibration based on Test 2. Concentration of the injected sample was 1.2125mg/mL.

Injection volume (µL)	Injection Amount (µg)	Theoretical concentration (mg/mL)*	Main Peak Area (mAUs)	Predicted concentration (mg/mL)**	Difference theoretical vs. predicted (%)***	Predicted concentration 1-point (mg/mL)****	Difference, 1-point (%)*****
2.00	2.43	0.303	3687.40	0.286	-5.5	0.292	-3.6
4.00	4.85	0.606	7553.27	0.601	-0.9	0.599	-1.2
6.00	7.28	0.909	11438.60	0.917	0.8	0.907	-0.3
8.00	9.70	1.213	15297.70	1.231	1.5	1.213	0.0
10.00	12.13	1.516	19004.80	1.533	1.1	1.506	-0.6
12.00	14.55	1.819	22606.10	1.826	0.4	1.792	-1.5
14.00	16.98	2.122	25907.00	2.094	-1.3	2.053	-3.2

*$Theoretical\ concentration = \frac{Injected\ volume}{Method\ injection\ volume} \cdot Concentration\ of\ injected\ sample$

$Predicted\ concentration = (\frac{Main\ Peak\ Area - 67.288 mAUs}{2\,585.4 \mu g})/6 \mu L$, intercept and slope from **Figure 10B.

***$Difference\ theoretical\ vs\ predicted = \frac{Predicted\ concentration}{Theoretical\ concentration} \cdot 100 - 100$

**** $Predicted\ concentration\ 1-point = \frac{Area}{(\frac{15\,297.70 mAUs}{1.213 mg/mL})}$

***** $Difference\ theoretical\ vs\ predicted = \frac{Predicted\ concentration\ 1-point}{Theoretical\ concentration} \cdot 100 - 100$

The difference for concentration determination of protein H fell within the interval of -5% to 5% for all injection amounts. Differences between concentrations predicted 1-point calibration and theoretical concentrations were also <±5% for all measured points except

1.28µg which had a difference of -5.7%. Thereby the measuring range for quantification of protein H is 2.56µg to 15.37µg. As mentioned in section 4.5,it is not guaranteed that the residuals are homoscedastic for 2.56 to 5.12µg, so measurements in this range should be viewed with caution, but are still included in the range since they fulfill the other criteria for linearity.

Table 8. Theoretical and predicted concentration protein H and percentage difference between theoretical and predicted for both 7-point calibration and 1-point calibration. Error interval given as two standard deviation based on triplicate dilutions. Where no error is given only a single injection was made. Concentration of the injected sample was 1.2810mg/mL.

Injection volume (µL)	Injection Amount (µg)	Theoretical concentration (mg/mL)*	Average Main Peak Area (mAUs)	Predicted Concentration (mg/mL)**	Difference theoretical vs. predicted (%)***	Predicted concentration 1-point (mg/mL)****	Difference, 1-point (%)*****
1.00	1.28	0.214	3146.58 ±50.56	0.207	-3.0	0.201	-5.7
2.00	2.56	0.427	6438.66	0.419	-1.8	0.412	-3.5
4.00	5.12	0.854	13148.60	0.852	-0.2	0.842	-1.5
6.00	7.69	1.281	20013.90 ±316.62	1.295	1.1	1.281	0.0
8.00	10.25	1.708	26756.13 ±495.63	1.729	1.2	1.713	0.3
10.00	12.81	2.135	33024.60	2.133	-0.1	2.114	-1.0
12.00	15.37	2.562	39414.17 ±625.13	2.545	-0.7	2.523	-1.5

*$Theoretical\ concentration = \frac{Injected\ volume}{Method\ injection\ volume} \cdot Concentration\ of\ injected\ sample$

$Predicted\ concentration = (\frac{Main\ Peak\ Area - 166.32 mAUs}{1\,536.5\ mAUs/\mu g})/8\mu L$, intercept and slope from **Figure 11B.

***$Difference\ theoretical\ vs\ predicted = \frac{Predicted\ concentration}{Theoretical\ concentration} \cdot 100 - 100$

**** $Predicted\ concentration\ 1-point = \frac{Area}{(\frac{26\,756.13 mAUs}{1.708\ mg/mL})}$

***** $Difference\ theoretical\ vs\ predicted = \frac{Predicted\ concentration\ 1-point}{Theoretical\ concentration} \cdot 100 - 100$

4.7 Repeatability

Once all running parameters were chosen the precision and the repeatability of the methods were established over three different occasions. As can be seen in **Table 9**, the average purity of T-final1 was determined to be 98.21% with an intermediate precision (n=9) standard deviation of 0.22% and an %RSD 0.22. while the average purity of reference standard was determined to be 98.23% with a standard deviation of 0.23% and an %RSD of 0.24. Both samples fall below the repeatability for purity assessment of an %RSD ≤ 1. The overall average concentration of T-final1 (n=9) was determined to be 42.6 mg/mL with a standard deviation of 0.9 mg/mL and an %RSD of 2.09, which meets the criteria of an %RSD ≤ 3 for concentration determination of pure samples as well. In conclusion, the repeatability of the method is sufficient for determination of concentration and purity in quality control of protein production.

Table 9. Intermediary precision for protein T method across 3 occasions. n=9. The same equipment and method used across the three occasions and the tests were performed by the same analyst.

	Purity			Concentration
	T-final1	**Protein T ref std**		**T-final1**
Average (%)	98.21	98.23	**Average (mg/ml)**	42.61262
SD (%)	0.22	0.23	**SD (mg/ml)**	0.890
%RSD*	0.22	0.24	**%RSD***	2.089

$*\%RSD = \frac{\text{SD}}{\text{Average}} \cdot 100$

The protein H method also appears to meet criteria for repeatability for purity, ≤1, as the %RSD was 0.16 for H-final1 and 0.21 for the reference standard (n=9), see **Table 10**. The %RSD for quantification of protein H reference standard was 1.38 (n=9), showing that the repeatability criteria for quantification. %RSD ≤3 for pure samples is also met.

Table 10. Intermediary precision for protein H method across 3 occasions. n=9. The same equipment and method used across the three occasions and the tests were performed by the same analyst.

	Purity			Concentration
	H-final1	**Protein H ref std**		**H-final1**
Average (%)	98.40	98.07	**Average (mg/ml)**	48.69
SD (%)	0.15	0.21	**SD (mg/ml)**	0.67
%RSD*	0.16	0.21	**%RSD (%)**	1.38

$*\%RSD = \frac{\text{SD}}{\text{Average}} \cdot 100$

The repeatability of both methods were also tested using impure samples but with only one triplicate from one occasion. The %RSD of purity assessment was 0.012 for T-proc2 and 0.210. 0.043. 0.987 and 0.489 for H-proc1. H-proc2. H-proc3. and H-proc4 respectively (n=3). The %RSD for concentration determination was 2.30 for T-proc2 and 1.317. 1.196. 1.486 and 1.934 for H-proc1. H-proc2. H-proc3. and H-proc4 respectively (n=3).

5 Conclusions

In conclusion, two methods have been developed, one for protein T and one for protein H, which fulfill the criteria stated in the aims. Both methods displayed acceptable selectivity since main peaks were distinguishable or separated from high molecular weight species and low molecular weight species and the retention time was consistent (± 10%). Both methods also showed good linearity within their determined measuring ranges for quantification, 4.85µg to 16.98µg for protein T and 2.56µg to 16.98µg for protein H. Concentration and purity determinations were repeatable for both proteins. The results were in general in line with known theory, smaller particle size as well as longer column length in the stationary phase leads better resolution. All in all, two accurate methods have been developed with good basis to be fully validated for use in a quality control setting.

6 Future aspects

Before use in quality control settings, the intermediary precision of both methods should be validated for different analysts and different equipment. Analysts and equipment introduce the largest variability in results. Since impurities tend to introduce more variability in the results the repeatability and intermediary precision can also be determined for impure samples, but is not required. As for improving the methods, one may develop a method for protein H using the silica based Xbridge columns for better resolution and operating stability of the column. Since Xbridge columns are more pH tolerant, one may then use mobile phases with pH closer to that of the analyte pI may decrease interactions with stationary phase and improve peak shape. The effects of column temperature can also be further explored. For sample preparation improvements one may also investigate reduction of the protein using lower DTT concentration at higher temperatures in order to decrease DTT consumption.

7 References

[1] ICH, *ICH Topic Q 6 B - Specifications: Test Procedures and Acceptance Criteria for Biotechnological/Biological Products*, European Medicines Agency (EMA), 1999.

[2] B. J. Olson, "Assays for Determination of Protein Concentration," *Current Protocols in Pharmacology*, vol. 73, no. 1, pp. A.3A.1- A.3A.32, 2016.

[3] L. R. Snyder, J. J. Kirkland and J. W. Dolan, "Biochemical and Synthetic Polymer Separations," in *Introduction to Modern Liquid Chromatography, Third Edition*, L. R. Snyder, K. J. J. and D. J.W., Eds., John Wiley & Sons, Inc, 2009, pp. 570-616, 631-641.

[4] K. Sandra, I. Vandenheede and P. Sandra, "Modern chromatographic and mass spectrometric techniques for protein biopharmaceutical characterization," *Journal of Chromatography A*, vol. 1335, pp. 81-103, 2014.

[5] S. Fekete, A. Beck, J.-L. Veuthey and D. Guillarme, "Theory and practice of size exclusion chromatography for the analysis of protein aggregates," *Journal of Pharmaceutical and Biomedical Analysis*, vol. 101, pp. 161-173, 2014.

[6] S. Fekete, L. Kizekai, Y. T. Sarisozen, N. Lawrence, S. Shiner and M. Lauber, "Investigating the secondary interactions of packing materials for size-exclusion chromatography of therapeutic proteins," *Journal of Chromatography A*, vol. 1676, p. 463262, 2022.

[7] J. Vessman, R. I. Stefan, J. F. Van Staden, K. Danzer, W. Lindner, D. Thorburn Burns, A. Fajgelj and H. Müller, "Selectivity in Analytical Chemistry (IUPAC REcommendations 2001)," *Pure Appl. Chem*, vol. 73, no. 8, pp. 1381-1386, 2001.

[8] ICH, *Q2(R2) Validation of analytical procedures - Scientific guideline*, EMA, 1995.

[9] EDQM, "2.2.46. CHROMATOGRAPHIC SEPARATION TECHNIQUES," in *European Pharmacopoeia 11.2*, Strasbourg, Cedex: European Directorate for the Quality of Medicines & HealthCare of the Council of Europe (EDQM), 2023, pp. 86-96.

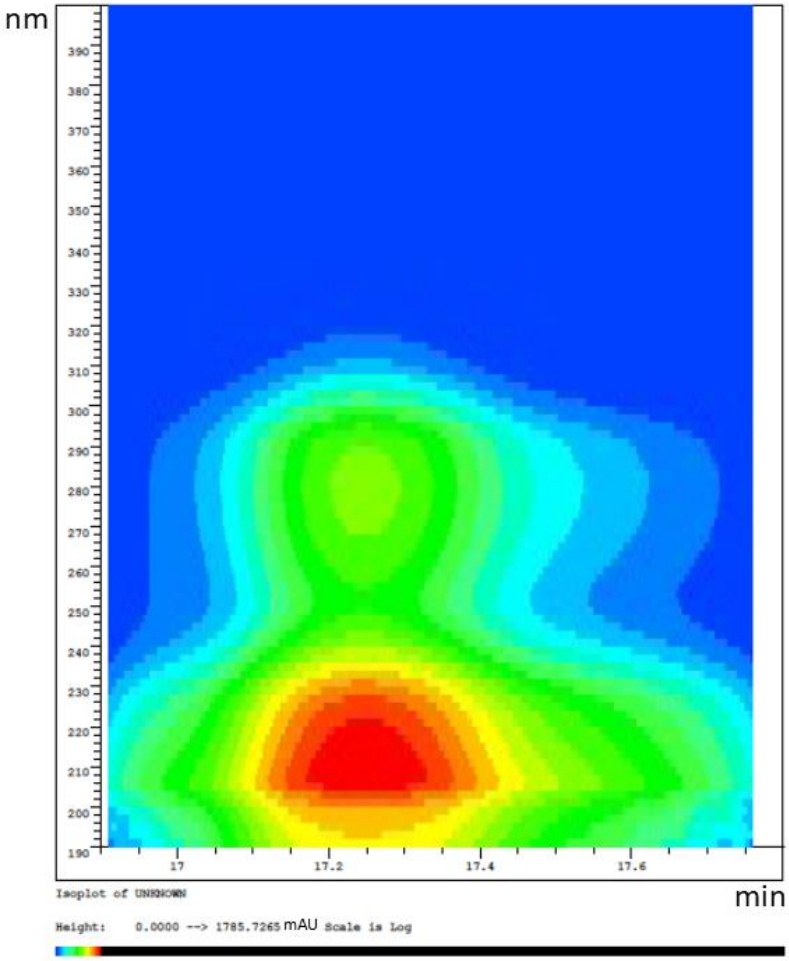
[10] ISO, *5275-1:1994 Accuracy (trueness and precision) of measurement methods and results – Part 1: General principles and definitions*, ISO (the International Organization for Standardization), 2003.

[11] M. Lauber, T. H. Walter, M. Gilar, M. DeLano, C. Boissel, K. Smith, R. Birdsall, P. Rainville, J. Belanger and K. Wyndham, *Low Adsorption LC Technology Based on MaxPeak High Performance Surfaces*, Milford, MA: Waters Corporation, 2021.

8 Appendix

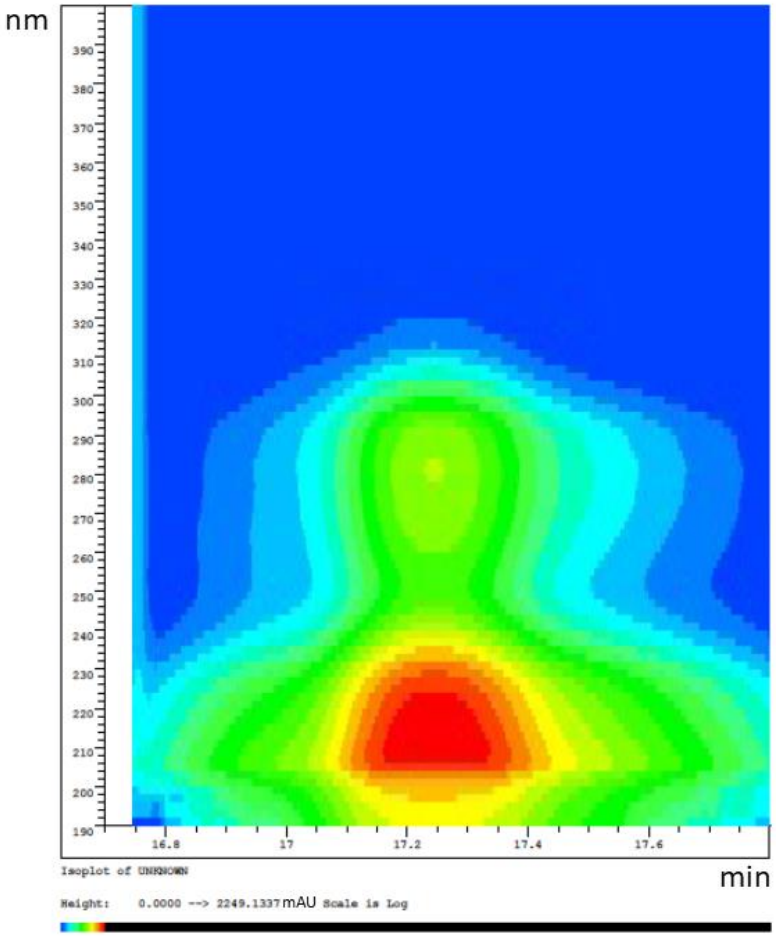
8.1 Isoplot of T-proc1 around main peak.

The recorded wavelengths were 190-400nm. The mobile phase was 0.3M Sodium phosphate buffer with a flow rate of 0.5mL/min. The 30cm Xbridge column was used with a column temperature of 35°C and the injection amount was 17.00µg. The sample was reduced at 85°C for 15 min red (100mM DTT 20 mM TRIS).



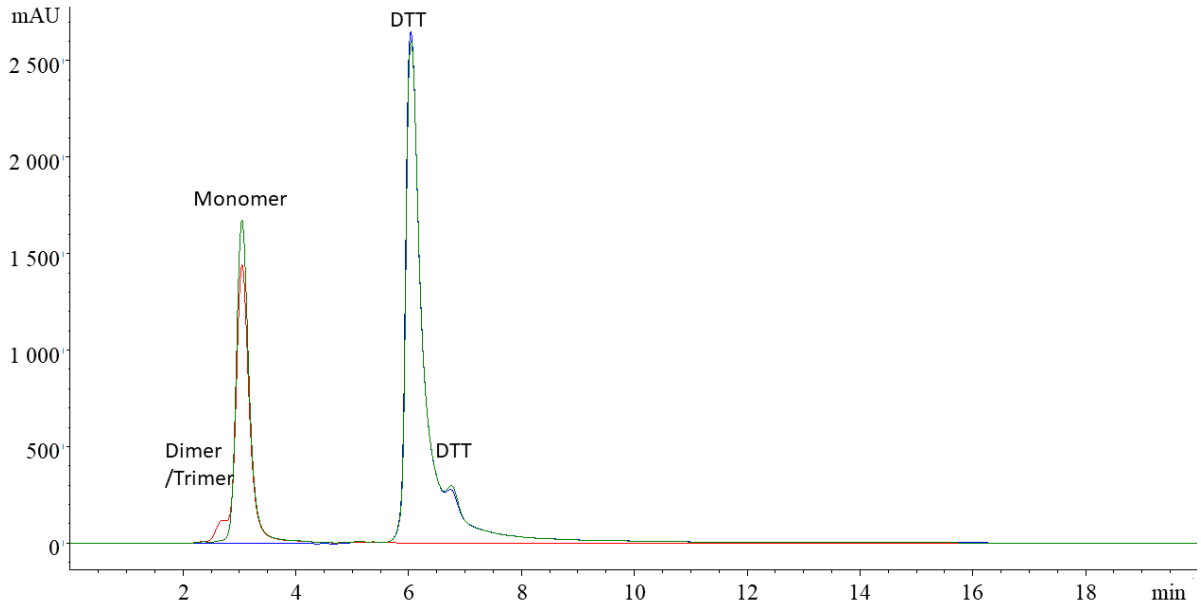
8.2 Isoplot of T-proc3 around main peak.

The recorded wavelengths were 190-400nm. The mobile phase was 0.3M Sodium phosphate buffer with a flow rate of 0.5mL/min. The 30cm Xbridge column was used with a column temperature of 35℃ and the injection amount was 23.52µg. The sample was reduced at 85℃ for 15 min red (100mM DTT 20 mM TRIS).



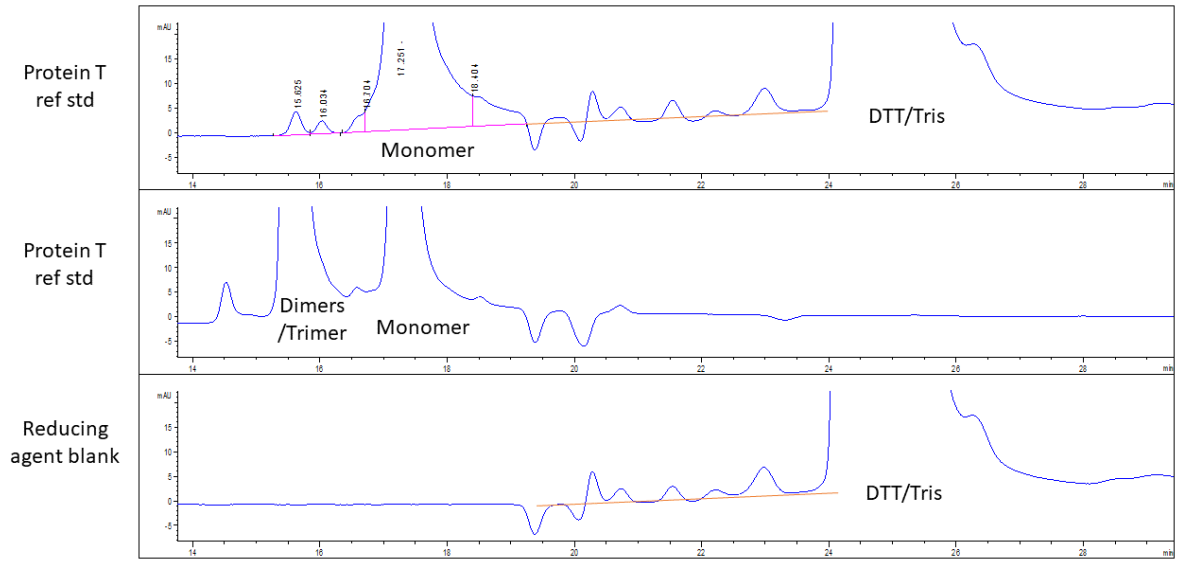
8.3 Protein H comparisons to blank

Representative example chromatograms of reduced (green) and non-reduced (red) protein T reference standard as well as only reducing agent (blue). The column used was the Xbridge Premier Protein 30cm column with 0.3M Sodium phosphate buffer, pH 7, as mobile phase. 10.24µg of protein was injected for both samples. The reduced sample was reduced at 85°C for 15 min in 100mM DTT with 20mM Tris.



8.4 Protein T comparisons to blanks and integration baseline

Representative example chromatograms of reduced (top) and non-reduced (middle) protein T reference standard as well as only reducing agent (bottom) showing how chromatograms were integrated. Integration baselines and peak splits are shown in pink. The orange line shows how the baseline is interpreted after the main peak for the integration. The column used was the Xbridge Premier Protein 30cm column with 0.3M Sodium phosphate buffer, pH 7, as mobile phase. 14.55µg of protein was injected for both samples.



8.5 Calculated F-values for residuals from protein H linearity test, with 60°C reduction.

Values were calculated as $F = \frac{s_A^2}{s_B^2}$, where s_A^2 ands_B^2 variance of the residuals for two injection amounts and s_A^2 is from the injection amount with the larger variance of the two. The degrees of freedom for all injection amounts were two, corresponding to a value of 39.00 on the F-distribution.

		Denominator Injection amount (µg)			
Numerator Injection amount (µg)	**Variance, s^2 (mAU^2s^2)**	**1.28**	**7.69**	**10.25**	**15.37**
1.28	9 900 000	-	40.5	72.3	157
7.69	400 000 000	-	-	1.79	3.88
10.25	716 000 000	-	-	-	2.16
15.37	1 550 000 000	-	-	-	-

8.6 Test of repeatability for determining purity of protein T.

The same equipment and method used across the three occasions and the tests were performed by the same analyst.

	Occasion 1: 2023-02-22				**Occasion 2: 2023-02-23**		**Occasion 3: 2023-02-24**	
Replicate	**T-final1**	**T-final2**	**Ref. std.**	**T-proc2**	**T-final1**	**Ref. std.**	**T-final1**	**Ref. std.**
1	98.15	98.43	98.17	96.03	98.06	97.97	98.53	98.53
2	98.09	98.49	98.17	96.01	97.97	98.04	98.50	98.54
3	98.08	98.46	98.11	96.02	98.04	98.04	98.45	98.51
Average (%)	98.10	98.46	98.17	96.02	98.02	98.02	98.49	98.53
SD (%)	0.038	0.033	0.002	0.011	0.048	0.039	0.043	0.017
%RSD	0.039	0.034	0.002	0.012	0.049	0.040	0.043	0.017
Average Max (%) CI (95%)**	98.20	98.54	98.17	96.05				
Average Min (%) CI (95%)***	98.01	98.38	98.16	95.99				

*$\%RSD = \frac{\text{SD}}{\text{Average}} \cdot 100$

** Average Max (%)CI (95%) = Average $+ \left(\text{T.INV}(0.05.2) \cdot \frac{\text{SD}}{\sqrt{3}}\right)$

*** Average Min (%)CI (95%) = Average $- \left(\text{T.INV}(0.05.2) \cdot \frac{\text{SD}}{\sqrt{3}}\right)$

8.7 Test of repeatability for determining concentration of protein T.

The same equipment and method used across the three occasions and the tests were performed by the same analyst.

	Occasion 1: 2023-02-22				Occasion 2: 2023-02-23	Occasion 3: 2023-02-24
Replicate	**T-final1**	**T-final2**	**Ref. std.**	**T-proc2**	**T-final1**	**T-final1**
1	42.2	52.2	9.74	20.7	43.4	43.0
2	40.9	49.3	9.80	21.1	42.3	43.1
3	41.8	50.5	9.73	20.2	43.6	43.3
Average (mg/mL)	41.6	50.7	9.75	20.7	43.1	43.1
SD (mg/mL)	0.71	1.45	0.037	0.47	0.69	0.16
%RSD*	1.69	2.86	0.38	2.30	1.60	0.36
Average Max (%) CI (95%)**	43.4	54.3	9.84	21.7		
Average Min (%) CI (95%)***	39.9	47.1	9.66	19.5		

*$\%RSD = \frac{\text{SD}}{\text{Average}} \cdot 100$

** Average Max (%)CI (95%) = Average $+ \left(\text{T.INV}(0.05.2) \cdot \frac{\text{SD}}{\sqrt{3}}\right)$

*** Average Min (%)CI (95%) = Average $- \left(\text{T.INV}(0.05.2) \cdot \frac{\text{SD}}{\sqrt{3}}\right)$

8.8 Repeatability for determining purity of protein H.

The same equipment and method used across the three occasions and the tests were performed by the same analyst.

	Occasion 1: 2023-04-04							Occasion 2: 2023-04-05		Occasion 3: 2023-04-06	
Replicate	**Ref. std.**	**H-final1**	**H-final2**	**H-proc1**	**H-proc2**	**H-proc3**	**H-proc4**	**H-final1**	**Ref. std.**	**H-final1**	**Ref. std.**
1	98.27	98.48	95.37	97.80	98.55	68.78	72.39	97.75	98.29	98.16	98.46
2	98.14	98.53	95.37	97.91	98.47	71.07	71.71	97.89	98.17	98.14	98.43
3	98.24	98.64	95.60	97.51	98.48	69.70	72.22	97.77	98.19	98.28	98.40
Average (mg/mL)	98.22	98.55	95.44	97.74	98.50	69.85	72.11	97.80	98.22	98.19	98.43
SD (mg/mL)	0.066	0.082	0.134	0.206	0.042	1.153	0.352	0.074	0.068	0.076	0.029
%RSD*	0.067	0.083	0.140	0.210	0.043	1.650	0.489	0.076	0.069	0.077	0.029
Average Max (%) CI (95%)**	98.38	98.75	95.78	98.25	98.60	72.72	72.98	97.99	98.39	98.38	98.50
Average Min (%) CI (95%)***	98.05	98.34	95.11	97.23	98.39	66.99	71.23	97.62	98.05	98.38	98.50

*$\%RSD = \frac{\text{SD}}{\text{Average}} \cdot 100$

** Average Max (%)CI (95%) = Average $+ \left(\text{T.INV}(0.05.2) \cdot \frac{\text{SD}}{\sqrt{3}}\right)$

*** Average Min (%)CI (95%) = Average $- \left(\text{T.INV}(0.05.2) \cdot \frac{\text{SD}}{\sqrt{3}}\right)$

8.9 Repeatability for determining concentration of Protein H.

The same equipment and method used across the three occasions and the tests were performed by the same analyst.

	Occasion 1: 2023-04-04						Occasion 2: 2023-04-05	Occasion 3: 2023-04-06
Replicate	**H-final1**	**H-final2**	**H-proc1**	**H-proc2**	**H-proc3**	**H-proc4**	**H-final1**	**H-final1**
1	48.36	19.07	23.35	12.72	1.90	1.76	49.53	47.95
2	48.48	19.25	22.77	12.51	1.96	1.70	48.96	48.31
3	49.25	19.36	22.91	12.43	1.92	1.75	49.60	47.77
Average (mg/mL)	48.69	19.23	23.01	12.55	1.93	1.73	49.36	48.01
SD (mg/mL)	0.482	0.148	0.303	0.150	0.029	0.034	0.451	0.272
%RSD*	0.989	0.768	1.317	1.196	1.486	1.934	0.913	0.566
Average Max (mg/mL) CI (95%)**	49.89	19.59	23.76	12.93	2.00	1.82	49.89	19.59
Average Min (mg/mL) CI (95%)***	47.50	18.86	22.26	12.18	1.86	1.65	47.50	18.86

*$\%RSD = \frac{\text{SD}}{\text{Average}} \cdot 100$

** Average Max (%)CI (95%) $= \text{Average} + \left(\text{T.INV}(0.05.2) \cdot \frac{\text{SD}}{\sqrt{3}}\right)$

*** Average Min (%)CI (95%) $= \text{Average} - \left(\text{T.INV}(0.05.2) \cdot \frac{\text{SD}}{\sqrt{3}}\right)$