

NPL REPORT AS 11

Protein Quantitation Assays for Biopharmaceutical Applications

**J.E. Noble, A.E. Knight, *A.J.
Reason, *A. Di Matola and
M.J. Bailey**

NOT RESTRICTED

September 2007

Protein Quantitation Assays for Biopharmaceutical Applications

J.E. Noble, A.E. Knight, *A.J. Reason, *A. Di Matola and M.J. Bailey

*M-Scan Ltd, 3 Millars Business Centre, Fishponds Close, Wokingham,
Berkshire, RG41 2TZ, United Kingdom

Quality of Life Division

ABSTRACT

This report presents a review of current methods for estimation of protein amount in biological solutions, and a comparison of dye-based protein assays for biopharmaceutical analysis. The review details the various methods to determine the amount of protein in an aqueous sample, and highlights the limitations of some of these methods. The comparison of dye-based protein quantitation assays looks at various aspects of these spectroscopic methods in relation to the analysis of biopharmaceuticals including; accuracy, repeatability, dynamic range etc. The report contains recommendations for the use of specific dye-based assays for certain applications.

© Crown copyright 2007
Reproduced with the permission of the Controller of HMSO
and Queen's Printer for Scotland

ISSN 1754-2928

National Physical Laboratory
Hampton Road, Teddington, Middlesex, TW11 0LW

Extracts from this report may be reproduced provided the source is acknowledged and the extract is not taken out of context.

Approved on behalf of the Managing Director, NPL
by Martin Sené, Director, Quality of Life Division

Table of Contents

Contents

1	Introduction and Context	1
1.1	Protein Quantitation Assays.....	1
1.2	Protein Composition can affect Dye-based Quantitation Assays Performance	2
2	Methods.....	3
2.1	Materials	3
2.2	Protein Preparation and Characterisation.....	4
2.3	Amino Acid Analysis.....	4
2.4	Protein PEGylation	5
2.4.1	PEGylation of Insulin and Lysozyme	5
2.4.2	Purification and Characterisation of the PEGylation Reactions	6
2.5	Protein Concentration Assays	6
2.5.1	BCA Assay.....	7
2.5.2	Bradford Assay	7
2.5.3	CBQCA™ Assay	7
2.5.4	DC Assay	8
2.5.5	Fluorescamine Assay	8
2.5.6	Quant-iT™	8
2.5.7	Data Analysis	9
3	Results.....	9
3.1	PEGylation of Lysozyme and Insulin	9
3.1.1	Insulin PEGylation.....	9
3.1.2	Lysozyme PEGylation	10
3.2	Protein-to-Protein Variability	12
3.2.1	Amino Acid Analysis to Determine Protein Amount	12
3.2.1.1	Dye-Based Protein Concentration Assays	13
3.2.1.2	BCA Assay.....	13
3.2.1.3	Bradford Assay	14
3.2.1.4	CBQCA™ and Fluorescamine Assays	14
3.2.1.5	DC Assay	15
3.2.1.6	Quant-iT™ Assay	15
3.3	The use of a BSA Standard to Estimate the Concentration of an Unknown	17
3.4	The effect of PEG-modification.....	18

3.5	The effect of Glycosylation	18
3.6	Quantitation range and the Limits of Detection.....	19
3.7	Variation Associated with the Protein Concentration Assays	20
3.8	Assay Robustness.....	21
3.9	Sensitivity to Sample Additives.....	22
4	Conclusions.....	22
	Acknowledgements.....	24
	Bibliography	25
5	Abbreviations.....	27

1 Introduction and Context

This report was produced under the Product Characterisation (PC4) project, part of the Measurements for Biotechnology Programme 2¹ for 2004-2007. The aims of the PC4 project were:

- Promotion of best practice in circular dichroism in a quality control environment/laboratory
- Identification of best practice in protein concentration measurement
- Validation of a complementary approach to higher order structure characterisation for QC

The determination of protein quantity (concentration) is an important test procedure for biopharmaceutical products as detailed in the ICH Q6B specifications [1]. Various methods for determining protein quantity exist with the choice of assay depending on protein preparation, quantity available and required throughput. Accurate protein concentration measurements are also required for other biopharmaceutical analytical techniques where the amount of sample is important for the performance of the technique. This study supports our concurrent protein structural studies that include circular dichroism (CD) and Fourier transformed infrared (FTIR) spectroscopy, where an accurate determination of the analyte concentration is critical for spectral analysis¹. ICH Q6B guidelines for biotechnological/biological products recommend spectroscopic profiling, highlighting CD as an appropriate technique [1]. The CD signal is proportional to the concentration of the protein; therefore errors in protein concentration estimation will effect spectroscopic structural interpretation [2] and could ultimately lead to biopharmaceutical batch product failure.

This best practice guide lists the methods commonly used to make protein concentration estimates. It then looks in detail at the relative merits of the different dye-based concentration assays, concentrating on different methodologies rather than manufactures of the same assay. From the conclusions in Table 6 a reader should be able to judge which assay suits their specific application based on certain criteria such as cost, repeatability and quantitation range and availability of a suitable standard protein.

1.1 Protein Quantitation Assays

Many of the physical properties of a protein can be exploited to determine its concentration in solution. Non-destructive protein quantitation techniques include; gravimetric analysis and UV absorbance spectroscopy, while total amino acid, or nitrogen content (Kjeldahl) analysis are destructive. Dyes can be used to estimate concentration by either covalently modifying the protein, or non-covalent interactions through specific amino acids, or hydrophobic residues of the protein.

Currently there are multiple dye binding assays to determine protein concentration, these offer many advantages including limited reagent usage and the ability to be performed in parallel microplate formats when compared to other methods. The dye

¹ Further information about this project, and others under the MfB programme, can be obtained at: <http://www.mfbprog.org.uk> or <http://www.npl.co.uk/biotech>.

binding assays are limited by concerns about variability of the dye binding to different protein structures, reproducibility and sensitivity to contaminants. To help address these concerns we have compared the performance of six protein concentration assays with seven model proteins representing various sizes and covalent modifications that are encountered in the biopharmaceutical industry. The methods selected represent traditional colorimetric assays and the recently introduced fluorescence based assays with reportedly enhanced detection limits. We have attempted to compare protein concentration assays that detect proteins through different mechanisms, rather than compare similar formulations of common assays from different suppliers. The criteria for the selection of dye-based assay were that the method was designed, or could be adapted to a 96-well plate format and was compatible with a top-reading spectrophotometer or fluorimeter as appropriate.

Amino acid analysis (AAA) was used as the “gold standard” technique to determine the absolute protein amount, to which the microtitre plate assays were compared. The analysis of proteins using AAA methods generally involves the hydrolysis of the protein and subsequent separation of the resultant amino acid mixture using chromatography. To improve detection, amino acids are typically labelled with a fluorescent dye either pre or post-separation, before analysis. AAA is considered a ‘gold standard’ method due to the reproducibility and robustness of the technique and the availability of commercial reagents and standards. AAA was chosen over the Kjeldahl and gravimetric analysis, as these methods are limited by the complexity of the techniques and the large amounts of protein required, while the latter is also limited by the extensive dialysis and drying required to remove water and salts [3]. Covalent modification of the protein, for example glycosylation and PEGylation can cause additional problems with the drying process associated with gravimetric analysis, limiting the use of this technique.

AAA was used to determine the amount of protein in the stock solution from which the extinction coefficient at 280 nm (ϵ_{280}) was subsequently derived, using the molecular weight and absorbance of the protein [4]. The extinction co-efficient of a protein can also be estimated using UV spectroscopic models (provided the molecular weight of the protein is known) derived by Gill and von Hippel [5], and later refined by Pace and colleagues [6]. The model uses the absorption coefficients for specific amino acids (Trp, Tyr and Cys) and generates a good estimation of the extinction coefficient where these amino acids are in abundance in the sample protein [6]. However, where there are a low abundance of these amino acids (for example insulin) the model can display deviations of up to 15% from that predicted by physical methods (AAA, gravimetric, Kjeldahl) [6]. The use of AAA to calculate ϵ allows the determination to be performed in the protein’s native buffer and does not depend on the frequency of Trp, Tyr, or Cys in the protein sequence.

1.2 Protein Composition can affect Dye-based Quantitation Assays Performance

Previous reviews in the literature have highlighted the variability of signal response with protein composition and other limitations of many of the colorimetric protein determination assays [7, 8]. Direct comparisons between glycosylated and non-glycosylated proteins using Coomassie, BCA (Bicinchoninic acid) and Lowry quantitation assays suggested that glycosylation can affect protein concentration determination with these methods when compared to AAA [9]. Another study has

compared AAA with the BCA, Bradford and Lowry assays using a range of proteins and concluded that the BCA assay generated results that were consistent with AAA, whereas those determined using Lowry and Bradford assays were more variable [10]. In this report, we have sought to compare a wider range of proteins, using both the traditional colorimetric and the more recently developed fluorescence assays. We have sought to make the study applicable to the biopharmaceutical industry by analysing not only the protein-to-protein variability and accuracy compared to a commercial bovine serum albumin (BSA) standard, but also by concentrating on repeatability, quantitation range, limit of detection (LOD) and the sensitivity of the assay to excipients typically found in biopharmaceutical formulations. We have also evaluated the effect of covalent modifications on protein concentration assays. Both the host-cell related modification ‘glycosylation’ and PEGylation, a chemical modification of the protein by adding a synthetic polymer were analysed to see if they changed protein concentration estimates performed using the dye-based assays. The addition of polyethylene glycol (PEG) polymers to a biopharmaceutical can improve the pharmacokinetic properties of the protein and reduce unwanted immunogenic responses to the drug.

By comparing the performance of multiple protein quantitation assays simultaneously with a large number of proteins, the data generated will add to and support other studies in the literature. The choice of protein assays was based on the European Pharmacopoeia approved methods 2.5.33 [11], except that the fluorimetric derivatisation method using o-phthalaldehyde was replaced with fluorescamine. When both assays were compared in a microtitre plate format the fluorescamine assay generated a superior signal noise (S/N) ratio (results not shown). The Biuret assay was not assessed due to the high quantity of protein required to generate a signal [7].

2 Methods

2.1 Materials²

The model proteins (lysozyme from chicken egg white, RNase A Type XII-A from bovine pancreas, RNase B from bovine pancreas, recombinant human insulin expressed in yeast and bovine IgG) were purchased at the highest purity available from Sigma (Poole, UK). The BSA (Fraction V) protein standard used as a reference in all the protein quantitation assays was purchased from Pierce (Rockford, Illinois). The protein quantitation kits/reagents were obtained from the following suppliers: 3-(4-carboxybenzoyl)quinoline-2-carboxaldehyde (CBQCATM), Quant-iTTM and fluorescamine (FluoroPureTM Grade) were from Invitrogen (Paisley, UK); Bradford reagent was acquired from both Sigma and Pierce, the DC Protein Assay kit was from Bio-Rad (Hemel Hempstead, UK); and the BCA Protein Assay Kit was from Novagen (Beeston, UK)³. The polyethylene glycol (PEG) derivatisation agent O-methyl-O’-

² Certain commercial materials, instruments, and equipment are identified in this manuscript in order to specify the experimental procedure as completely as possible. In no case does such identification imply recommendation or endorsement by the National Physical Laboratory, nor does it imply that the material, instrument or equipment identified is necessarily the best available for the purpose.

³ Protein quantitation kits were selected because they were different examples of common methods, however this is not an endorsement of the kit manufacturers. The CBQCATM and Quant-iTTM assays are unique to one supplier and have been analysed for comparison with other methods, rather than as an endorsement.

succinylpolyethylene glycol 5000 N-succinimidyl ester (PEG-5000 SE), unmodified polyethylene glycol 5,000 (PEG-5000) and other buffer agents were obtained at reagent grade at the highest purity available from Sigma.

2.2 Protein Preparation and Characterisation

Proteins were typically prepared at 5 mg/ml in assay buffer: sodium phosphate buffer (10 mM phosphate buffer, 2.7 mM potassium chloride and 145 mM sodium chloride, pH 7.4) except insulin and insulin-PEG, which were prepared in 10 mM sodium acetate buffer, pH 5.0. The protein samples were then extensively dialysed into the relevant assay buffer to produce the 'stock solution' and stored at 4°C. All proteins were > 95 % pure (protein content) as determined by SDS-PAGE and mass spectrometry. Throughout the course of the study protein stability was periodically checked using SDS-PAGE and UV absorbance; no degradation, adsorption, or precipitation (light-scattering) of the proteins was observed.

The concentration of protein in these stock solutions was determined by AAA, from which the extinction coefficient at a wavelength of 280 nm (ϵ_{280}) for the native protein was determined (see Amino Acid Analysis section). Using the ϵ_{280} values determined further protein absorbance measurements were performed on a Lambda 850 spectrometer from PerkinElmer (Seer Green, UK), using the European Pharmacopoeia method 2.5.33 [11], in which the extent of light scattering at 280 nm was calculated and subsequently corrected. Briefly the logarithm of the absorbance from 320-350 nm was plotted against the logarithm of the wavelength, from which a standard curve was generated by linear regression. The curve was extrapolated to determine the logarithm of the absorbance at 280 nm that was attributed to light scattering from the protein solution. For all protein samples tested the absorbance attributed to light scattering represented less than 5 % of the total protein signal. Absorbance scans from 280-350 nm (0.5 nm interval) were performed using matched pairs of 1 cm cuvettes from Hellma (Müllheim, Germany), using the appropriate assay buffer as the blank.

2.3 Amino Acid Analysis

AAA was used to determine the amount of protein in the stock solution from which the extinction coefficient could be calculated using the additional data derived from absorbance and mass spectrometry measurements. Prior to AAA, 20 µg of the protein sample was analysed by on-line LC/ES-MS (LCT Premier MS from Waters (Elstree, UK)) and the raw data were processed using MassLynx Transformation to generate the average molecular mass spectra. Typically a 20 µg sample of intact protein was eluted using an aqueous formic acid (0.05% v/v) and acetonitrile gradient on a reverse phase Phenomenex C₄ peptide trap cartridge at a flow rate of 200 µl/minute. MS detection was performed using the LCT Premier; data acquisition was started 3 minutes following injection. Sodium format was used to calibrate the ES-MS instrument over the appropriate mass range. Absorbance measurements at 280 nm were performed on aliquots of the protein stock using a Jasco UV spectrophotometer (Great Dunmow, UK).

AAA was based on the method described by Barrett [12]; briefly, six approximate 100 µg aliquots of each protein were mixed with 50 nanomoles of the internal standard, norleucine. Each mixture was then lyophilised and re-dissolved in 6 M HCl; the optimum hydrolysis conditions were then determined for each protein (lysozyme and RNase A: 24 hour hydrolysis at 110°C; insulin and RNase B: 4 hour hydrolysis at

145°C). The optimum conditions were defined as the point when the highest hydrolysis rate was achieved with the least decomposition of the released amino acid, when comparing the two-hydrolysis conditions detailed. The requirement to analyse different hydrolysis conditions arises from the difficulty in cleaving highly hydrophobic areas of the protein, while trying to retain the structure of the less robust amino acids. The hydrolysed protein samples were then derivatised with phenylisothiocyanate (PITC) and analysed in duplicate by RP-HPLC (Supelcosil column: LC-18-DB, 5 μ m, 25 cm $\times$ 4.6 mm (Sigma) run on a Hewlett Packard 1050 series modular HPLC system) using a gradient of water to acetonitrile buffered to pH 6.4 using sodium acetate. The amino acids detected by absorbance at 254 nm were then quantified by comparison to the data produced from a standard mixture of known amounts of PITC-derivatised amino acids. For each hydrolysis condition, three aliquots of hydrolysed protein were prepared and each aliquot was analysed in duplicate by HPLC. The average amount of each amino acid present was calculated from these 6 HPLC analyses using samples from the optimum hydrolysis conditions. The data generated for the most robust amino acids (typically Pro, Leu, Phe, Tyr, Met and Ala) were used to calculate the amount of protein present; the ϵ_{280} was then calculated using the Beer-Lambert Law also using the absorbance and mass spectrometry data detailed above.

2.4 Protein PEGylation

Modification of biopharmaceuticals by PEGylation changes many of the physical (hydrodynamic radius) and therapeutic (immunogenicity, plasma clearance) properties of the modified protein. Physical changes caused by PEGylation may affect the performance of the dye-based assays described; this was therefore assessed using PEGylated insulin and lysozyme.

2.4.1 PEGylation of Insulin and Lysozyme

Lysozyme and insulin were PEGylated through covalent attachment of PEG-5000 SE to lysine residues and/or the N-terminus of the proteins. Lysozyme and insulin were dialyzed extensively into 'PEGylation labelling buffer' (100 mM sodium bicarbonate pH 8.5) at 5 mg/ml and 1 mg/ml respectively (insulin was prepared at 1 mg/ml due to its limited solubility in alkaline solutions). Dialysis was performed using Slide-A-Lyzer dialysis cassettes from Pierce using Mw cut-offs of 3500 and 5000 Da for insulin and lysozyme respectively. Dialysis was performed overnight at 4°C in 2000 ml of PEGylation labelling buffer using constant stirring. The PEGylation reaction was started by adding PEG-5000 SE in powder form to the protein solution while stirring, and the reaction was allowed to proceed at room temperature (21°C) for 30 minutes. The final concentration of PEG-5000 SE in the labelling solutions was at a 10-fold excess (molarity) for the lysozyme and 5-fold for insulin labelling reaction. The reaction was quenched by the addition of 100 μ l of 1.5 M N-hydroxylamine prepared at pH 8.5, the reaction was allowed to proceed for 30 minutes at room temperature. The N-hydroxylamine is a low molecular weight amine and will covalently modify unreacted PEG-5000 SE. The large excess of N-hydroxylamine effectively stops 'quenches' the PEG labelling of the proteins. Side reaction with other amino acids can occur with SE labelling chemistry, quenching with N-hydroxylamine is reported to also remove any coupling between PEGylation reagents and tyrosine residues [13], limiting the amount of non-specific labelling.

2.4.2 Purification and Characterisation of the PEGylation Reactions

Quenched PEGylated protein reactions were purified using IEX chromatography on an AKTA FPLC system from GE Healthcare (Little Chalfont, UK). PEGylated insulin was purified on a Q-Resource column equilibrated with 20 mM Tris-HCL pH 8.0, while PEGylated lysozyme was purified on a S-Resource column equilibrated with 20 mM HEPES pH 7.0, both at a flow rate of 1 ml/min. The bound protein was eluted with a NaCl gradient (0–600 mM for PEG-insulin and 0–400 mM for PEG-lysozyme). All the fractions from the purification were analysed for the presence of PEG using a colorimetric assay based on the partitioning of ammonium ferrothiocyanate from the aqueous to a chloroform phase in the presence of PEG, as described by Nag and colleagues [14]. Briefly ammonium ferrothiocyanate was prepared by dissolving 16.2 g of anhydrous ferric chloride and 30.4 g of ammonium thiocyanate in distilled water to a final volume of 1 litre. All steps of the assay were performed in a fume hood using appropriate techniques for handling hazardous materials. Estimations of the amount of PEG in column fractions were derived from a PEG-5000 standard curve (amount of PEG-5000 determined by mass). Typically PEG standards and unknown column aliquots of 50 μ l were prepared in 1.5 ml microfuge tubes to which was added 0.5 ml of both ammonium ferrothiocyanate and chloroform. The tubes were mixed using vigorous vortexing for 30 minutes to facilitate partitioning of the PEG, these tubes were then centrifuged at 3000 g for 2 minutes. The lower chloroform layer was removed and the absorbance recorded at 510 nm using preparations containing no PEG as the blank. Absorbance measurements were performed as detailed in Section 2.2 and the chloroform solution analysed was disposed of using appropriate methods. The ferrothiocyanate PEG assay generate a linear response from 5–100 μ g of PEG and was used to estimate which fractions from the IEX purification contained either unreacted PEG, or PEGylated proteins.

Purified PEG conjugated proteins were then dialysed into the appropriate assay buffer and concentrated using a centrifugal concentrator from Sartorius (Epsom, UK). The concentration of the PEG-modified proteins was determined using ϵ_{280} , with a correction for the PEG-SE moiety (absorbance at 280 nm of a 1 mg/ml solution of PEG-5000 SE was 0.008), using the method detailed in Section 2.2. The number of PEG units per protein was estimated using the proteins' migration on SDS-PAGE and by determination of the number of free amines using the fluorescamine assay (Section 2.5.1).

2.5 Protein Concentration Assays

Protein concentration assays were performed according to the kit manufacturers' instructions in a microtitre plate format and detailed briefly below. Parameters for all of the protein concentration assays are detailed in Table 1. Data from concentration series for all the proteins were compiled from four independent experiments, each performed four times per plate, unless stated. Every plate contained a BSA (Pierce standard) concentration series performed in duplicate, allowing direct comparison (normalisation) between experiments performed on separate 96-well plates.

Fluorescence assays were prepared in FluoroNunc flat-bottomed black polystyrene 96-well plates and absorbance assays were prepared in clear flat-bottomed polystyrene (non-treated) 96-well plates, both from Nunc (Roskilde, Denmark). Microplate measurements were performed on a Victor³ 1420 Multilabel Counter from

PerkinElmer (Seer Green, UK) at ambient temperature (range 25-29°C). Protein dilutions were prepared in sodium phosphate buffer (as detailed above), except insulin (10 mM sodium acetate, pH 5.0) and BSA standard (155 mM NaCl and 7.7 mM sodium azide), and stored in low protein binding microfuge tubes from Eppendorf (Histon, UK).

Protein concentration curves were diluted into the appropriate assay buffer into low-protein binding microfuge tubes and stored at 4°C until use. Liquid transfer of protein dilutions to the microplate were performed using an electronic pipette in repeat dispensing mode, whereas assay reagents were transferred using multichannel pipette.

2.5.1 BCA Assay

The bicinchoninic acid reagent detects the reduction of copper ions by proteins in an alkaline solution. BCA working reagent was prepared by diluting the 4% cupric sulphate solution 50 fold into the BCA solution, 200 µl BCA working solution was then added to a clear microplate containing 25 µl of the protein standards or unknown. The plate was sealed and incubated for 30 minutes at 37°C, then 10 minutes at room temperature. The plate seal was removed carefully to avoid transfer of evaporated liquid into other wells and the absorbance measured as indicated in Table 1. Data was fitted to a second order polynomial response using least-squares regression as detailed in Section 2.5.7.

2.5.2 Bradford Assay

Two preparations of Bradford reagent were analysed due to the low signal generated by Bradford reagent 1 with low molecular weight proteins. Bradford assays consist of a formulation of Coomassie dye in an acidic solution; this dye shows a metachromatic shift when it reacts with specific regions of the protein. To a clear microplate 5 µl of standard, or unknown protein solution was pipetted, to which either 250 µl, or 300 µl of Supplier 1 and 2 Bradford reagent respectively was added. The plate was incubated for 40 minutes before reading the absorbance as directed in Table 1. Data was fitted to a 4-parameter quadratic response using least-squares regression as detailed in Section 2.5.7.

2.5.3 CBQCA™ Assay

The CBQCA™ derivitization reagent (ATTO-TAG™) reacts with amines in the presence of cyanide to yield a fluorescent product that can be used to quantitate the amount of protein. Protein dilutions were prepared in 0.1 M sodium borate, pH 9.3 and made up to 135 µl in the same buffer, and then added to a black 96-well plate. To the plate was added 5 µl of 20 mM KCN solution (caution KCN solutions are toxic and should be handled with care, attention should be taken when disposing of solution containing cyanide). The ATTO-TAG™ working reagent was prepared at 2 mM by diluting the 40 mM DMSO stock 20-fold in sodium borate buffer, 10 µl of this working reagent was then added to every well to initiate the reaction. The plate was then incubated at room temperature for 120 minutes, before analysing the fluorescence using the parameters detailed in Table 1. Data was fitted to a linear response as detailed in Section 2.5.7.

2.5.4 DC Assay

The DC protein assay is based on the reaction of the protein with an alkaline copper tartrate solution and Folin reagent. DC working reagent was prepared by diluting reagent S 50-fold into reagent A to produce 'working reagent A'. To a clear microplate was added 5 µl of the protein unknown, or standard, 25 µl of working reagent A and 200 µl of reagent B in sequential order. The plate was left to incubate for 40 minutes before reading the absorbance as directed in Table 1. Data was fitted to a second order polynomial response using least-squares regression as detailed in Section 2.5.7.

2.5.5 Fluorescamine Assay

The fluorescamine assay was based on the microplate method developed to measure amino acids by Bantan-Polak and colleagues [15], with the parameters optimised to generate the best S/N ratio using BSA as a model protein (data not shown). To a black 96-well plate, 25 µl of protein sample, 155 µl of sodium borate buffer pH 10.4 (final concentration 50 mM) and 20 µl of 0.6% (w/v) fluorescamine prepared in acetonitrile (final amount 0.06%), were sequentially added. The plate was then incubated at room temperature for 120 minutes, before analysing the fluorescence using the parameters detailed in Table 1. Data was fitted to a linear response as detailed in Section 2.5.7.

2.5.6 Quant-iT™

The Quant-iT™ preparation contains a dye formulated in detergent, when the dye binds detergent coated protein, or hydrophobic patches of the protein an enhanced fluorescence signal is obtained that can be used to determine protein concentration. The Quant-iT™ working solution was prepared by diluting the Quant-iT™ dye 200-fold into the buffer solution, 200 µl of the working solution was added to a black microwell plate containing 25 µl of the protein standard, or unknown. The plate was incubated for 32 minutes before reading the absorbance as directed in Table 1. Data was fitted to a 4-parameter logistic (Boltzman) response using least-squares regression as detailed in Section 2.5.7.

Assay	Instrument Filters (nm)	
	Absorbance/Excitation	Emission
Bradford	595/450 ^s	-
BCA	550	-
CBQCA™	485	535
DC	650	-
Fluorescamine	405	460
Quant-iT™	450	570

^s A ratio of A_{595}/A_{450} was taken as this is reported to improve the signal noise ratio [7]

Table 1: Dye-based Protein Concentration Assay Parameters. The absorbance and fluorescent wavelengths used for each protein concentration assay.

2.5.7 Data Analysis

To compare the dye-based assays curve fitting was performed on the protein dilution series using Origin Version 5 from OriginLab (Northampton, USA), allowing the response to be determined at specific protein concentrations. Curve fitting was performed as detailed in the manufacturer's instructions, where each data point on the curve represents the average of 4-independent assays, each performed 4 times. Points that displayed signal saturation, or low precision ($CV > 10\%$) were excluded from the curve fitting procedure.

3 Results

3.1 PEGylation of Lysozyme and Insulin

3.1.1 Insulin PEGylation

Insulin is reported to contain three possible PEGylation sites; the N-termini of chains A and B (Gly-A1 and Phe-B1, respectively) and Lys-B29 from the B-chain [16]. The alkaline pH of the labelling reaction would favour conjugation to Lys-B29 ($pK_a \sim 10.5$), rather than Phe-B1 ($pK_a \sim 7.0$) and Gly-A1 ($pK_a \sim 8.0$) [16, 17]. Elution of the insulin PEGylation reaction mix on the Q-Resource column generated two major peaks that bound to the column, analysis on an SDS-PAGE gel indicated that the first peak (135-275 mM NaCl) contained PEGylated insulin, whereas the second peak (338-478 mM NaCl) contained unreacted insulin (Figure 1A). The first series of peaks (150, 180 and 220 mM NaCl) that could not be resolved using this system appear to contain insulin with varying degrees of PEG modification. The fluorescamine data (Figure 2) suggests that the pooled insulin-PEG peak (135-275 mM NaCl) had 60% less amine reactive sites than the unmodified insulin; assuming all three amine sites could react with the fluorescamine dye, this would generate a labelling ratio of 1.8 PEG units per protein. Unreacted PEG-5000 SE did not bind to the column, and its

presence was detected in the column flow-through by using the ammonium ferrothiocyanate assay [14].

3.1.2 Lysozyme PEGylation

Lysozyme contains 7 possible PEGylation sites (the N-terminus and 6 Lys residues); analysis of the reaction product by SDS-PAGE suggests that all the detectable lysozyme had been modified by at least one PEG moiety (data not shown). Purification of the PEGylated lysozyme on an S-Resource column generated two distinct PEGylated protein populations, one that did not bind to the column, and another that was eluted at 80 mM NaCl. Comparison of the bound and unbound PEG-modified proteins on an SDS-PAGE gel suggested that the population that did not bind to the column displayed reduced mobility when compared to the fraction that was eluted at 80 mM NaCl (Figure 1B). Again unreacted PEG-5000 SE was not retained by the column and was therefore separated from the bound fraction of PEGylated lysozyme (elution from 40-123 mM NaCl). Fluorescamine analysis of the PEGylated lysozyme (80 mM NaCl fraction) indicated that over 80% of the accessible amine groups in the protein had been modified, generating a labelling ratio of ~5.8 PEG molecules per protein.

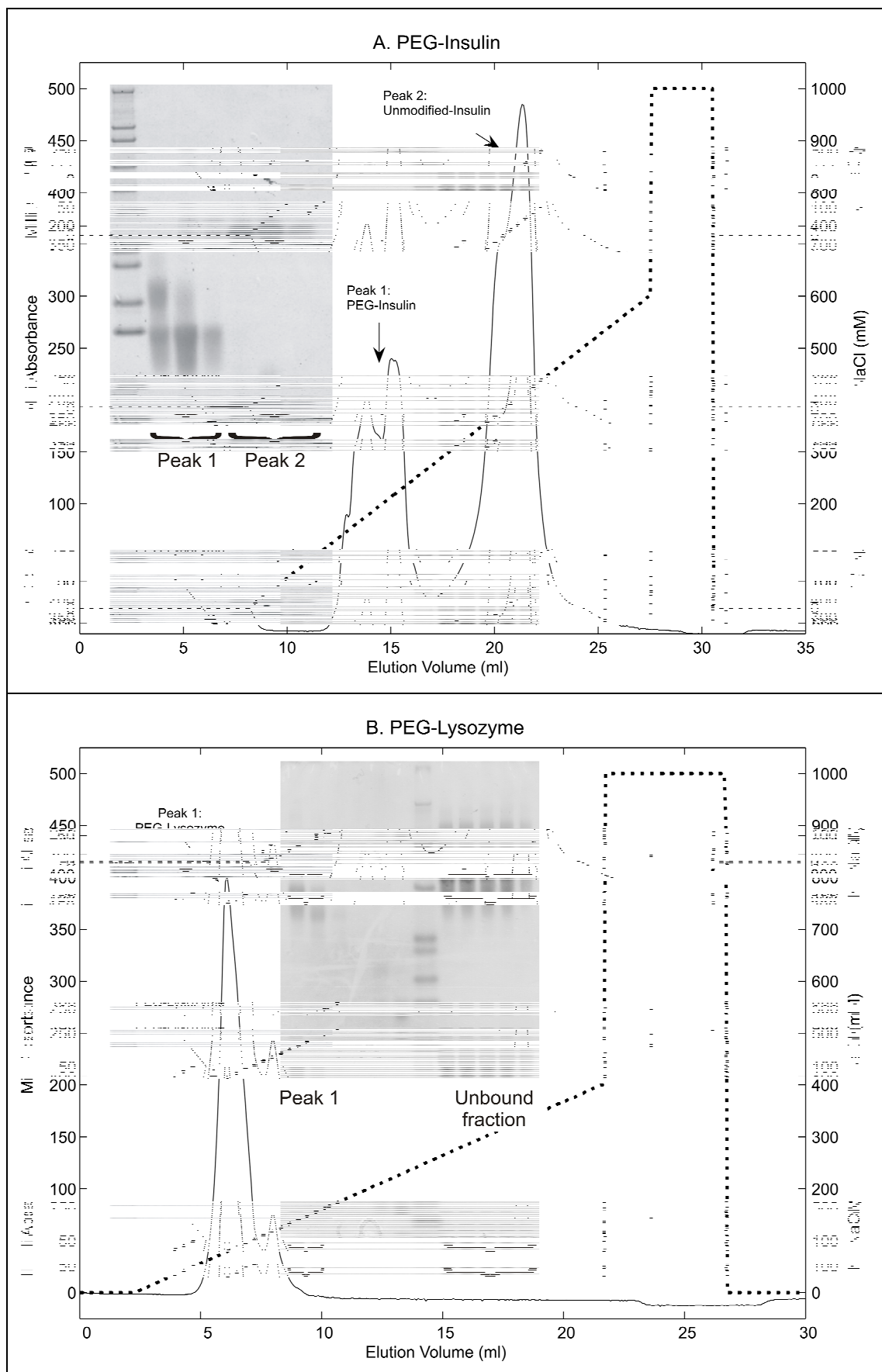


Figure 1. Purification of PEG-modified Proteins

The purification of (A) PEG-insulin and (B) PEG-lysozyme, from unreacted PEG-5000 SE and unmodified protein by IEX chromatography. For both chromatograms the salt gradient is shown by the dotted line. Proteins were detected by absorbance at 280 nm and the protein peaks displayed on the chromatogram were analysed using thiol reducing SDS-PAGE (4-12% Bis-Tris gel, run in MES running buffer); relative molecular mass determination was estimated using Mark 12 molecular weight standard (standards from top to bottom: 200, 116, 97, 66, 55, 36, 31, 21, 14, 6 and 3 kDa (insulin gel only)) all from Invitrogen.

3.2 Protein-to-Protein Variability**3.2.1 Amino Acid Analysis to Determine Protein Amount**

Many of the traditional dye-based protein concentration assays are dependent on both protein quantity and composition. The elements of a protein that are reported to influence concentration assays include amino acid content, post-translational modifications (glycosylation) and the protein conformation [7]. To determine the effect of protein sequence on the protein concentration assays AAA was taken as the 'gold standard' method to which the dye-based concentration assays were compared. Table 2 details the calculated ϵ_{280} for the selected proteins using AAA, mass spectrometry and the uv absorbance data. The ϵ_{280} values were derived using the Beer-Lambert Law from absorbance measurements on the stock protein solution, with the amount of protein and its molecular weight determined through AAA and mass spectrometry respectively. It is likely that a significant fraction of the errors associated with the determination of the ϵ_{280} values arose from the incomplete derivatisation of amino acids due to the presence of salts in the protein buffer. Determination of the extinction coefficient was performed in the assay buffer, so as to generate values that represent the conditions that both absorbance measurements (ϵ_{280}) and the colorimetric and fluorescent measurements were performed in. The protein IgG is a mixture of purified immunoglobulins; therefore AAA was not performed in this case due to the heterogeneous nature of the protein sequence. The IgG used was a mixed source of immunoglobulins, whereas the majority of biopharmaceutical IgG's are monoclonal having a defined single peptide sequence. To obtain a commercial source of a monoclonal IgG with a published sequence at the amounts required for AAA and all the dye-based assays would be difficult, therefore a semi-purified polyclonal source was used in this study. The concentration of protein in the IgG sample was instead calculated using the manufacturer's reported extinction coefficient, and the protein's purity (IgG content) was confirmed by analysis on SDS-PAGE (data not shown). Figure 2 shows the variation in signal observed for each assay when the proteins were analysed at a fixed concentration that was within the quantitation range of each of the dye-based assays.

Protein	Mw (Da)	ϵ_{280} (L.mol ⁻¹ .cm ⁻¹)	ϵ_{280} %CV
Lysozyme	14303	34196	3.9
Insulin	5806	6300	4.7
RNase A	13680	9240	6.1
RNase B	15095*	8213	6.5

*The Mw for RNase B was derived by the relative abundance of each of the glycosylated species to estimate the mass average molecular weight.

Table 2: Amino Acid Analysis Data for the Model Proteins

The %CV for the ϵ_{280} represent the variance associated with the amount of a specific amino acid (only “robust” amino acids were used in the calculation) calculated from each protein sample during AAA. In using the Beer-Lambert law to determine the ϵ_{280} , absorbance and mass spectrometry measurements were also made, but the errors associated with these measurements were not significant compared to those associated with AAA (%CV for Abs₂₈₀ was always <0.08).

3.2.1.1 Dye-Based Protein Concentration Assays

Protein concentration assays were selected based on their different reaction mechanisms. The fluorescamine and CBQCA™ assays primarily detect secondary amines (ϵ -amino group of lysine and the N-terminal amino acid) in proteins and peptides. The PEG-5000 SE modification reagent was tested with all the protein concentration assays to determine if it generated a positive signal in the absence of protein. The CBQCA™ reagent generated a significant signal (when compared to the blank of assay buffer) in the presence of PEG-5000 SE, but not the unmodified PEG-5000. Given that both reagents react with secondary amines it is surprising that they cross-react, therefore the CBQCA™ assay could not be used to assess the degree of PEG-modification. As expected, PEGylation resulted in a decrease in intensity for both insulin and lysozyme when analysed with the fluorescamine reagent; this method was subsequently used to estimate the degree of labelling.

Variation in the response of each microplate assay to amino acid sequence was assessed by comparing the signal generated for the 7 model proteins at a fixed concentration within the quantitation range. Analysis of the variance (ANOVA) was used to examine the null hypothesis ‘that all the proteins analysed generated the same response’ when analysed at the same concentration (Data not shown). Analysis of all the assays examined suggests that the mean response from at least one protein was significantly different from the mean (Probability ≤ 0.05). Protein composition therefore appears to affect all the microplate assays analysed, this response is discussed for each of the methods in more detail below.

3.2.1.2 BCA Assay

The BCA assay shares a similar mechanism with the DC assay during which the production of monovalent copper ions form a complex with bicinchoninic acid, producing a purple complex that displays a characteristic absorbance at 565 nm. Reports suggest that like the DC assay, the BCA method is sensitive to protein

composition and that the Cu^{1+} complexation by BCA is also sensitive to chelating agents. The proteins analysed displayed a similar variation between composition and signal intensity with that observed for the DC assay (Figure 2). Other reports have noted this variation observed with protein composition [9, 10] when compared with AAA.

3.2.1.3 Bradford Assay

The Bradford reagent is a formulation containing the Coomassie Blue 250 dye, which displays a metachromatic shift upon binding to a protein, absorbing strongly at 595 nm. The Bradford assay signal is dependent on the protein composition, with lysine, arginine and/or hydrophobic interactions playing an important role in dye binding [7]. Two preparations of Bradford reagent from two different suppliers were analysed with the model proteins. In Figure 2, it is clear that there is a significant difference between the two preparations. Supplier 1 reagent exhibits a very weak response to proteins with a low molecular weight. The only protein that displayed a significant signal with the Supplier 1's reagent was IgG. It has been reported that the size of the protein plays an important role in the magnitude of signal generated, and many authors have reported the unsuitability of the assay to measure bioactive peptides [7]. All of the proteins that generated low absorbance signals and did not display a linear response with increasing protein amount had molecular weights of less than 16 kDa. The data using Bradford reagent from Supplier 1 appears to support the hypothesis of a molecular weight 'threshold' that must be exceeded for optimal dye-binding and to obtain the maximum response [7]. However to confirm this threshold our library of proteins analysed would need to contain more proteins with a range of molecular weight's greater than 16 kDa.

The Bradford reagent from Supplier 2 exhibits a response that varies significantly between the model proteins (PEGylated proteins were not assessed), with no apparent protein molecular weight threshold (Figure 2). The different response is most likely due to the different preparations of the Bradford reagent used; reports suggest that basic proteins generated better responses when the ratio of phosphoric acid to dye was reduced (Supplier 2 compared to Supplier 1) [18], or the acidity of the reaction was reduced by addition of NaOH [19]. The high degree of signal variation observed with the protein panel for both Bradford preparations mirrors that reported by other groups [9, 10, 20], and underlines the need for a appropriate reference protein.

3.2.1.4 CBQCA™ and Fluorescamine Assays

Both the CBQCA™ and fluorescamine data displayed a strong dependence on protein composition (Figure 2). However, when all the proteins were analysed, the signal intensity did not correlate with the number of amines in the protein per unit mass (data not shown). The relationship between amine content and signal intensity assumes that all the amines in the molecule were in a position to reactive with the modification reagent, typically surface exposed as the reaction was performed on the native protein. The fluorescamine assay was optimised for BSA giving an optimum reaction pH of 10.4; this however may bias certain proteins depending on the pKa profile of their lysine residues. The CBQCA™ assay was performed according to the manufacturer's instructions at pH 9.3, which may explain it's different reaction profile when compared to the fluorescamine data (Figure 2).

3.2.1.5 DC Assay

The DC assay is based on a reaction of the protein with an alkaline copper tartrate solution and the subsequent reduction of a Folin reagent. The composition of the protein is known to affect the signal and this is thought to be predominately due to the number of tyrosine and tryptophan residues, and to a lesser extent cystine, cysteine and histidine residues in the protein. This variation is reportedly reduced by measuring at the shorter wavelengths, therefore the absorbance was measured at 650 nm, which is still within the range recommended by the manufacturer (650-750 nm) [7]. Proteins analysed with the DC assay displayed relatively low signal variation when compared to the other dye-based assays (Figure 2).

3.2.1.6 Quant-iT™ Assay

The Quant-iT™ reagent is a detergent/dye-based preparation in which the dye binds to both the hydrophobic regions and detergent coating on the sample protein. When bound the dyes fluorescence is enhanced allowing determination of the protein concentration [21]. The apparent lack of amino acid specific binding of the Quant-iT™ reagent is suggested to result in little protein-protein variability [21]. The Quant-iT™ reagent displayed significant variation with protein composition, with both RNase A and RNase B both displaying elevated signal intensity compared to the other proteins analysed (Figure 2).

In summary the BCA, DC and Quant-iT™ display low protein variability when compared to the Bradford, CBQCA™ and fluorescamine assays. In addition there appears to be a large variation in assay response between different preparations of the Bradford reagent. Different suppliers of the BCA and DC assays are available, however they have not been analysed in this study.

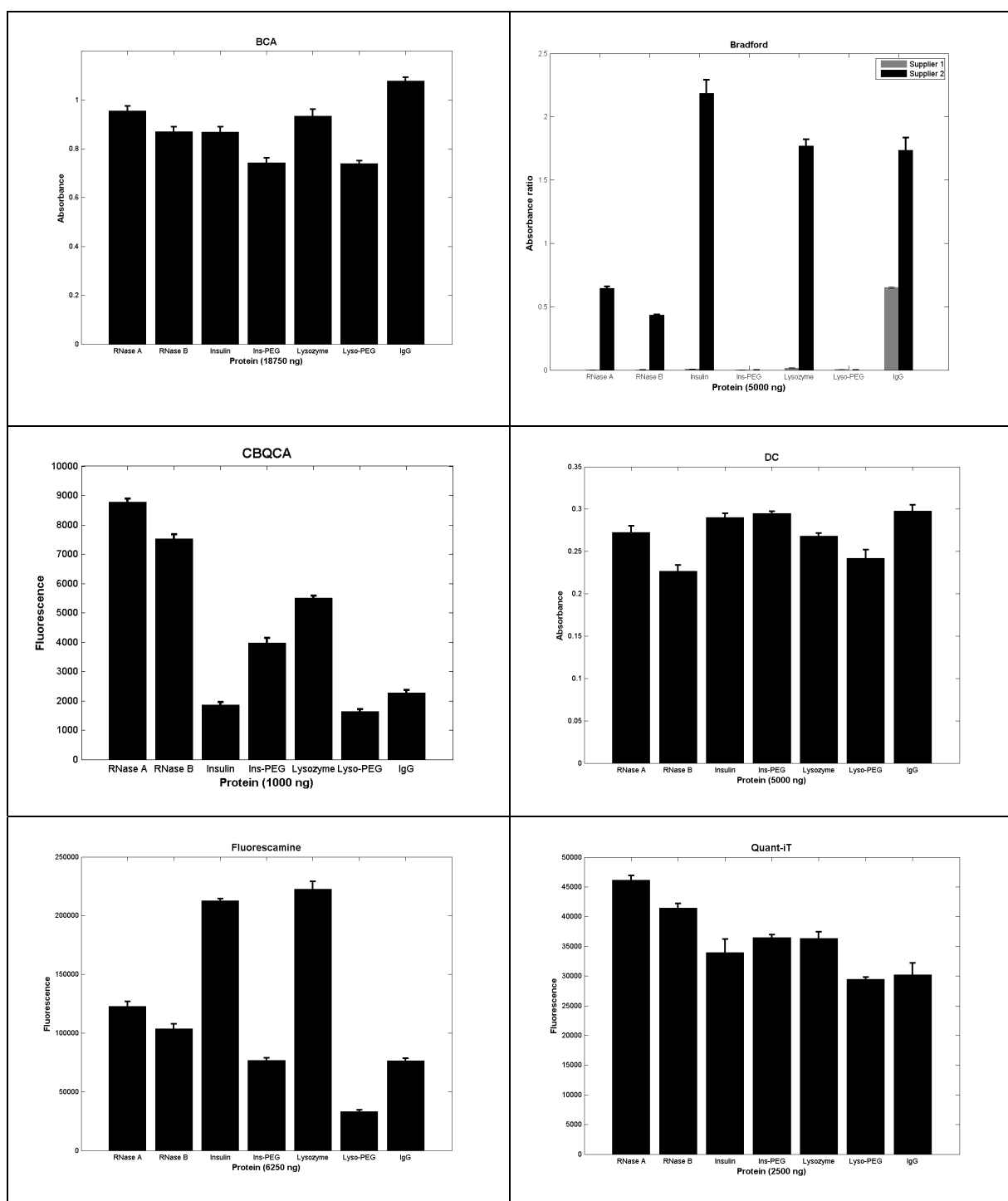


Figure 2. Effect of Protein Composition on Dye-based Protein Concentration Assays

Protein assays were performed and analysed as detailed in the Materials and Methods section. Each data point represents the average response (absorbance, or fluorescence units) at the specified protein amount. The associated variation represent the σ from the protein concentration estimate derived from the 4-independent plate assays. The amount of RNase B (ng) has been corrected so that only the protein (rather than carbohydrate) content was included in the mass calculation. The response of PEG-insulin and PEG-lysozyme with Bradford reagent from Supplier 2 was not determined.

3.3 The use of a BSA Standard to Estimate the Concentration of an Unknown

When a suitable model protein cannot be used, BSA or gamma globulins are often employed as a standard from which the unknown protein concentration is determined. BSA standard curves were prepared in every plate to account for potential variability between experiments. We therefore compared the known concentration of the model proteins (determined by AAA) with that obtained if the BSA standard curve were used to estimate the concentration of the model protein. The spread of data shown in Figure 3 suggests that caution should be taken in using BSA to estimate the amount of an unknown protein for all the microplate assays analysed, due to the large variability observed between the BSA standard and AAA estimates. Table 3 shows the inter-plate and intra-plate variation associated with the 6 dye-based assays. The good repeatability (% CV 2-8 %) suggests that the variation in concentration estimates between the AAA and BSA standard data Figure 3 are not due to repeatability between wells in a plate, or between plates but, rather the dye-based assays response being dependent on the amino acid content of the protein.

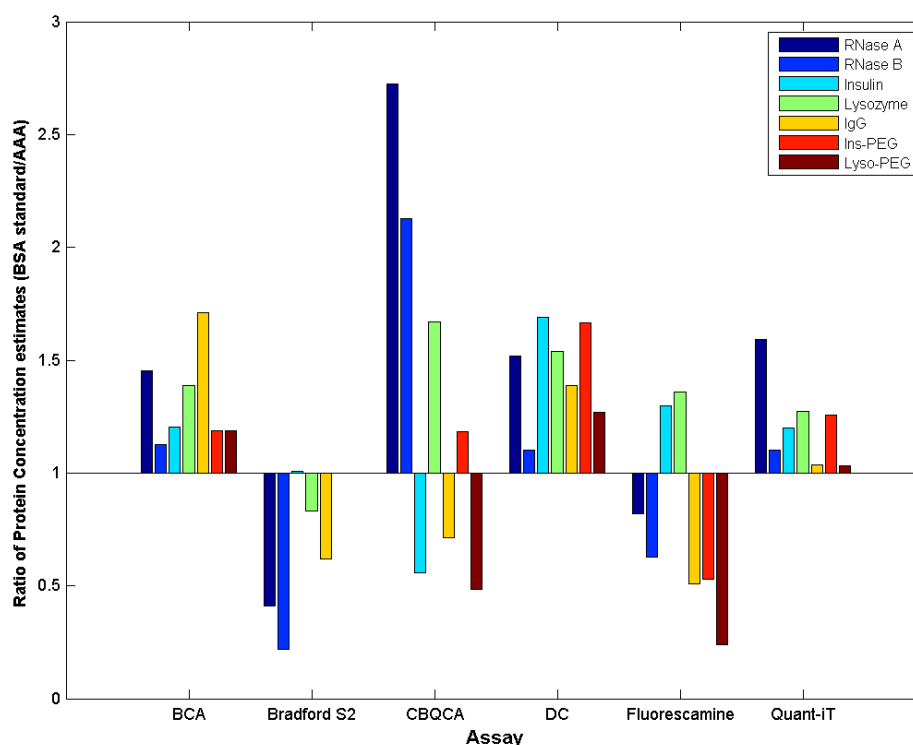


Figure 3. Errors Associated with using BSA as a Protein Concentration Standard

AAA was used to determine the concentration of the model proteins; using these estimates a calibration curve for each protein was prepared to using the dye-based assays. In the same plate a calibration curve using the BSA standard (Pierce; concentration defined by manufacturer) was also prepared and the response of this was compared to that of the model proteins to see how well the BSA standard estimated the true concentration of the model proteins. The ratio of concentration estimations refers to the concentration of protein derived using the BSA standard compared to the 'true' value using AAA, where a ratio of 1 indicates the two methods gave the same value. Bradford S2 denotes data obtained using the Bradford reagent from 'Supplier 2'.

Assay	Variation (% CV)		
	Normalised inter-plate	Inter-plate	Intra-plate
BCA	3.62	3.98	2.89
Bradford	5.81	5.39	5.26
CBQCA™	6.45	6.87	8.33
DC	6.51	4.44	4.61
Fluorescamine	4.95	5.56	2.40
Quant-iT™	5.38	6.17	6.82

Table 3: Variation Associated with the Protein Concentration Assays

The % CV values represent the average variance generated using the designated assays for the 7 proteins analysed. Intra-plate variance is the error from 4-separate reactions in a single plate; whereas inter-plate variance is the error calculated from the average value from 4-separate plate means. Normalised inter-plate variance shows the effect of normalising the data from each plate to a BSA standard on the plate-to-plate variation, when compared to the non-normalised data.

3.4 The effect of PEG-modification

Attachment of PEG moieties to proteins results in the formation of a hydration layer around the protein that greatly increases the viscosity radii of the conjugate [22] and potentially shields immunogenic sites. The possibility exists that the PEG moieties could affect the interaction of the dyes with the protein; therefore potentially generate errors in the estimation of the protein concentration using dye-based methods. This could be important if a non-PEGylated protein was used for the standard curve to estimate the amount of the PEGylated protein. The number of PEG units per protein was estimated using the fluorescamine assay, as the PEG-modification agent and fluorescamine dye would be expected to share similar reactive sites within the protein. Assay sensitivity to PEGylation appeared to be protein dependent and was rarely more than 10% of the unmodified protein signal Figure 2. Statistical analysis using the t-test (95% confidence limits) was performed to determine which assays generated a significantly different result when comparing the effects of PEG-modification. The assays DC and the Quant-iT™ (insulin) were not significantly affected by PEG-modification, whereas the BCA, Quant-iT™ (lysozyme), CBQCA™ and fluorescamine assays (Bradford assay not determined) were significantly different. It could be expected that, like the CBQCA™ and fluorescamine assays, covalent attachment of the PEG group to a lysine would decrease the signal generated by the Bradford reagent. It therefore appears that PEGylation can effect protein concentration measurements, but this is dependent on the assay and potentially the degree of modification.

3.5 The effect of Glycosylation

Previous studies have shown that the BCA, Bradford and Lowry protein concentration assays are sensitive to the glycosylation state of the analysed protein [9], and other techniques are sensitive to specific sugars when present in the assay buffer [7]. The

different glycoforms of RNase were chosen to investigate the effect of glycosylation on protein concentration assays, as RNase B contains a single glycosylation site, which is absent in RNase A. The oligosaccharide moiety on RNase B is reported to be GlcNAc₂Man₅₋₉ [23]. LC/ES-MS spectra (data not shown) of RNase B generated 5 major peaks each differing by 162 Da (equivalent to a mannose unit), this indicates that the protein sample contained various truncated glycoforms generating an average molecular weight of 15095 Da. The oligosaccharide therefore represents on average ~9% of the total mass of the glycoprotein. Comparison of equal molar concentrations of RNase A and RNase B suggested that the non-glycosylated form of the protein gave significantly higher signals (t-test, 95% confidence limits) for all the assays analysed (Figure 2). This difference was observed over the entire concentration range analysed for all the assays, which suggests this was not a concentration effect, or an erroneous result. The Bradford data generated using Supplier 2's formulation displayed a larger difference between the RNase B and RNase A signals, this is in agreement with the study by Fountoulakis and colleagues who observed a similar decrease in signal for various glycoproteins, when compared to AAA [9].

3.6 Quantitation range and the Limits of Detection

The quantitation range of an analytical procedure is the interval between the upper and lower concentration (amounts) of analyte, for which it has been demonstrated that the analytical procedure has a suitable level of precision, accuracy and goodness of fit to the derived mathematical response [24]. Figure 4 displays the quantitation range for all of the protein assays performed. The quantitation range was determined by fitting the plotted data to the response suggested in the manufacturer's instructions. Deviations from this response at low, or saturating, protein concentrations that reduced the R² value of the model were considered to be outside the quantitation range. Data points with CV values greater than 10% were also excluded from the quantitation range. Protein assays based on fluorescence as the signal readout generated larger quantitation ranges (typically 3 orders of magnitude), compared to those assays that used absorbance as readout (typically 2 orders of magnitude). Fluorescence read-out assays also appeared to be more sensitive to low protein concentrations; this was also observed when the limit of detection (LOD: lowest amount of analyte in a sample that can be detected but not necessarily quantitated as an exact value [24]) for each of the 7 proteins was determined for the protein concentration assays. The LOD was determined as the concentration of protein generated from the blank (no protein) $\pm 3\sigma$ of the blank, and calculated from the fitted protein concentration response. The LOD values determined were dependent on the assay used rather than the protein analysed, with the fluorescent and absorbance-based assays generated LOD values around 10 and 100 ng respectively (for volumes of protein solution analysed in each assay see Section 2.5).

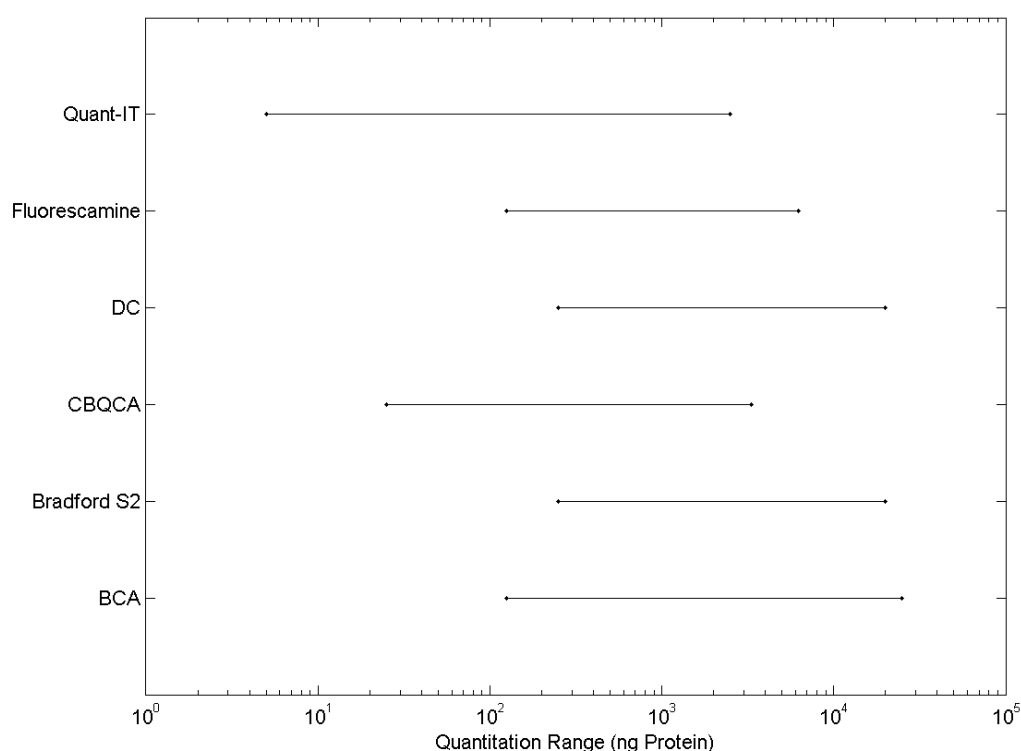


Figure 4: Quantitation Ranges for the Dye-based Protein Assays

The diamonds designate the upper and lower values for the quantitation range of each protein assay. Concentration curves for all the model proteins were performed and analysed as described in the Materials and Methods section. The Quantitation range was defined as the range of protein amounts (ng) that displayed good precision and did not show any deviation from the fitted response curve. Data for the Bradford reagent from Supplier 1 was omitted, as the quantitation range could only be determined from one of the proteins analysed and therefore this range was not considered a good representation of all the model proteins assessed within the study.

3.7 Variation Associated with the Protein Concentration Assays

The sources of variation associated with protein concentration measurements were analysed and shown in Table 3. The average % CV for all the proteins analysed using a specific dye-based assay was calculated. The intra-plate CV was typically about 5 %, this represents liquid handling errors, plate specific errors (such as edge effects), and detection errors. The inter-plate errors indicate how reproducible the assay is (repeat assays were performed using the same operator on the same day). ‘BSA normalised’ data used the BSA concentration curve to normalise data from different plates allowing comparison, this minimised the effects of assay specific artefacts (temperature, changes in reagent composition etc). Inter-plate variation (reproducibility of the assay) was typically 3-6 % CV, with the fluorescent assays displaying the highest % CV values. When the data from the inter-plate variation was normalised to the BSA standard curve typically a small decrease in the variation was observed. This suggests that artefacts such as fluctuations in room temperature and sample preparation do not appear to play a significant role in quantitation of proteins using the selected assay kits. Indeed, there appeared to be no correlation between

ambient temperature (25-29°C) and signal for all the assays performed when non-normalised data was analysed.

3.8 Assay Robustness

Assay robustness was examined by comparing the variation in protein quantitation when the assay was repeated over an extended time-scale (day-to-day variation) and when another operator conducted the assays (operator-to-operator variation). Table 4 suggests that when the amount of IgG was determined, both day and operator variation introduced significant errors, for the majority of the assays (95 % confidence limits). IgG was stable throughout the course of the experiments as determined by SDS-PAGE and protein absorbance. The day-to-day variation typically gave estimates of protein concentration that were within 5 % of the previously determined value, with the greatest variation displayed by the assays that employed fluorescence as a readout, or included multiple steps. Operator variation was used to assess if an inexperienced operator could derive the same result using the protein concentration assays. Assays with multiple steps: fluorescamine, CBQCA™ and BCA, appeared to result in the highest levels of inaccuracy when comparing operator variation, presumably due to accumulation of multiple pipetting errors.

Assay	Robustness			
	Day		Operator	
	% CV	P	Variation	P
BCA	5.2	0.002	0.93	0.02
Bradford	3.1	0.51	1.01	0.92
CBQCA™	6.6	0.005	1.2	0.01
DC	4.5	0.16	1.01	0.90
Fluorescamine	10.0	0.003	1.18	0.001
Quant-iT™	2.5	0.72	0.97	0.48

Table 4: Robustness of the Protein Concentration Assays

Concentration curves for IgG were prepared as Materials and Methods for the 6 dye-based assays and data was normalised to the BSA standard curve. Day-to-day robustness compares the variation in response generated from 4 independent concentration estimations of IgG, each repeated on a separate day using the same operator. 'P' represents the probability (ANOVA) that on the four days tested the assays would generate the same concentration estimate. Operator robustness compares the variation associated with two different operators, the 'variation' represents the ratio (normal operator/inexperienced operator) between the two resulting concentration estimates, with the assays being performed as detailed in the day-to-day testing. The 'P' value represents the probability that the operators would generate the same result (t-test, homoscedastic, 2-tailed distribution).

3.9 Sensitivity to Sample Additives

Biopharmaceutical formulations often contain additives (known as excipients). These may include bulking agents (such as mannitol and sucrose), stabilizers (inert disaccharides such as sucrose and trehalose), tonicity adjusters (such as glycine) and detergents such as polysorbate and tween. The excipients detailed in Table 5 were prepared at various concentrations with BSA to see if they affected the accuracy of the subsequent protein concentration assay. Unsurprisingly, both the CBQCA™ and fluorescamine assays were sensitive to glycine, presumably due to a reaction with the secondary amine of this amino acid. The BCA assay also appeared to be sensitive to many of the mono- and disaccharides analysed; previous analysis has shown that the BCA assay is sensitive to reducing sugars and sucrose [9, 25]. The composition of the protein formulation needs to be considered when choosing an assay, as particular assays are sensitive to certain sugars and detergents (information for detergents is typically provided by the manufacturer) that can lead to errors if the standard protein is not prepared with the same excipients.

Assay	Concentration of excipient (mM) that results in a significant change in the assay response					
	BCA	Bradford	CBQCA™	DC	Fluorescamine	Quant-iT™
Contaminants	BCA	Bradford	CBQCA™	DC	Fluorescamine	Quant-iT™
Mannitol	100	-	1000	1000	1000	-
Glycine	100	1000	0.01	-	0.1	-
Sorbitol	10	-	1000	-	1000	-
Sucrose	-	-	-	-	-	-
Trehalose	-	-	100	-	100	-

Table 5: Effect of Excipients on the Protein Concentration Assays

The effects of excipients at concentrations between 0.01–1000 mM (except trehalose, whose solubility limit in water was 150 mM) were analysed to see what concentration of excipient resulted in a significant change in the measurement of 1 mg/ml BSA for all the dye-based assays. The concentration of excipient that resulted in signals greater than $\pm 3\sigma$ of the control samples (without added excipient) was considered to cause a significant change in the protein concentration assay's performance. The values in the table represent the lowest concentration (mM) of added excipient required to produce a significant difference in the protein estimation using the dye-based assays. Blanks indicate that the performance of the assay was unaffected at the highest concentration of excipient added. Assays were performed as detailed in the Materials and Methods section and repeated 4 times for each concentration of excipient.

4 Conclusions

A diverse set of model proteins, including glycosylated and PEG-modified proteins were analysed with dye-based assays to compare protein determination methods for biopharmaceuticals. This best practice guide describes in detail how the assays were undertaken and which performed best when analysed using specific criteria important for the biopharmaceutical industry. Table 6 shows nominal rankings for the dye-based concentration assays for various criteria analysed throughout the study.

The cost of performing each data point was also compared in Table 6. This cost estimate does not take into account the amount of protein and the user time required to prepare and run the assay, which can be significantly more expensive, these factors are addressed in the number of steps and LOD criteria's.

Where protein variability is an important factor, i.e. the standard is a different protein from the unknown sample; assays with low inter-protein variation (Quant-iT™, BCA and DC) are preferable. Where the concentration of the unknown sample protein is likely to be highly variable, or in limiting amounts (≤ 10 ng), the fluorescent assays that display lower LOD's and a large dynamic range might be more suitable. For QC measurements, or applications, where repeatability is a critical factor, the absorbance-based measurements are preferable, although the high inter-protein variability of the Bradford reagent would necessitate a careful choice of standard.

Glycosylation and PEGylation did not result in large errors for protein concentration estimation (typically $<10\%$), except where the modification resulted in loss of reactive groups, for example the CBQCA™ and fluorescamine assays. This study also highlighted the problems of using a non-cognate standard (BSA) for estimating the amount of an unknown sample; accuracy relative to BSA standards. Where applicable the use of the same, or a similar protein for the standard would therefore be recommended.

Assay	Cost per sample data pt (\$)	Number of steps	Accuracy relative to BSA standard	Repeatability	LOD	Quantitation range	Protein variation
BCA	0.05	2	2	2	3	2	1
Bradford	0.03	1	4	1	3	3	4
CBQCA™	0.27	4	4	4	1	2	3
DC	0.025	3	2	1	4	3	1
Fluorescamine	0.01	2	2	4	1	3	2
Quant-iT™	0.4	2	2	2	1	1	1

Table 6: Summary of Dye-based Protein Concentration Data

The protein concentration assays were ranked on specific criteria for suitability to measure set of model proteins analysed. A ranking of 1 indicates the assay fits the criteria well, whereas a ranking of 4 indicates the assay might not be the best choice if that criterion is an overriding factor. These nominal ranks are based on data derived from a limited number of model proteins and therefore not a recommendation of a specific protein concentration assay.

Acknowledgements

The work was funded by the UK's Department of Trade and Industry's National Measurement System through the Measurements for Biotechnology (MfB) programme. We wish to thank George Tranter and Alison Rodger for comments and reviewing the manuscript.

Bibliography

1. **Specifications: Test Procedures And Acceptance Criteria for Biotechnological/Biological Products (Q6B). (1999) International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. pp. 1-16.**
2. **Kelly, S.M. and Price, N.C. (2000) The use of circular dichroism in the investigation of protein structure and function. Curr Protein Pept Sci. 1(4), 349-84.**
3. **Blakeley, R.L. and Zerner, B. (1975) An accurate gravimetric determination of the concentration of pure proteins and the specific activity of enzymes. Methods Enzymol. 35, 221-6.**
4. **Anders, J.C., Parten, B. F., Petrie, G. E., Marlowe, R. L., McEntire, J. E. (2003) Using Amino Acid Analysis to Determine Absorptivity Constants. BioPharm International, 30-37.**
5. **Gill, S.C. and von Hippel, P.H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. Anal Biochem. 182(2), 319-26.**
6. **Pace, C.N., et al. (1995) How to measure and predict the molar absorption coefficient of a protein. Protein Sci. 4(11), 2411-23.**
7. **Sapan, C.V., Lundblad, R.L., and Price, N.C. (1999) Colorimetric protein assay techniques. Biotechnol Appl Biochem. 29 (Pt 2), 99-108.**
8. **Lundblad, R.L. and Price, N.C. (2004) Protein concentration determination. Bioprocess International. 2(1), 38-46.**
9. **Fountoulakis, M., Juranville, J.F., and Manneberg, M. (1992) Comparison of the Coomassie brilliant blue, bicinchoninic acid and Lowry quantitation assays, using non-glycosylated and glycosylated proteins. J Biochem Biophys Methods. 24(3-4), 265-74.**
10. **Alterman, M., et al. (2003) AAARG2003 Study: Quantitation of proteins by amino acid analysis and colorimetric assays. in ABRF 2003; Translating Biology Using Proteomics and Functional Genomics. Denver, CO.**
11. **Methods of Analysis. (2004) Council of Europe (COE) - European Directorate for the Quality of Medicines. pp. 138-141.**
12. **Barrett, G., C. (1985) Degradation of amino acids accompanying in vitro protein hydrolysis, in *Chemistry and Biochemistry of the amino acids*, G. Barrett, C., Editor. Chapman and Hall: London, New York. pp. 376-398.**
13. **Morpurgo, M. and Veronese, F.M. (2004) Conjugates of peptides and proteins to polyethylene glycols. Methods Mol Biol. 283, 45-70.**
14. **Nag, A., Mitra, G., and Ghosh, P.C. (1996) A colorimetric assay for estimation of polyethylene glycol and polyethylene glycolated protein using ammonium ferrothiocyanate. Anal Biochem. 237(2), 224-31.**
15. **Bantan-Polak, T., Kassai, M., and Grant, K.B. (2001) A comparison of fluorescamine and naphthalene-2,3-dicarboxaldehyde fluorogenic**

- reagents for microplate-based detection of amino acids. *Anal Biochem.* 297(2), 128-36.
16. Hinds, K., et al. (2000) Synthesis and characterization of poly(ethylene glycol)-insulin conjugates. *Bioconjug Chem.* 11(2), 195-201.
17. Uchio, T., et al. (1999) Site-specific insulin conjugates with enhanced stability and extended action profile. *Adv Drug Deliv Rev.* 35(2-3), 289-306.
18. Chan, J.K., Thompson, J.W., and Gill, T.A. (1995) Quantitative determination of protamines by coomassie blue G assay. *Anal Biochem.* 226(1), 191-3.
19. Stoscheck, C.M. (1990) Increased uniformity in the response of the coomassie blue G protein assay to different proteins. *Anal Biochem.* 184(1), 111-6.
20. Knight, M.I. and Chambers, P.J. (2003) Problems associated with determining protein concentration: a comparison of techniques for protein estimations. *Mol Biotechnol.* 23(1), 19-28.
21. Haughland, R., P. (2005) Protein Detection and Proteomics Technology, in *A Guide to Fluorescent Probes and labeling Technologies*, M. Spence, T, Z., Editor. Invitrogen. pp. 413-472.
22. Fee, C.J. and Van Alstine, J.M. (2004) Prediction of the viscosity radius and the size exclusion chromatography behavior of PEGylated proteins. *Bioconjug Chem.* 15(6), 1304-13.
23. Fu, D., Chen, L., and O'Neill, R.A. (1994) A detailed structural characterization of ribonuclease B oligosaccharides by ¹H NMR spectroscopy and mass spectrometry. *Carbohydr Res.* 261(2), 173-86.
24. Validation of Analytical Procedures: Text and Methodology (Q2(R1)). (1994) International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. pp. 1-13.
25. Mopper, K. and Gindler, E.M. (1973) A new noncorrosive dye reagent for automatic sugar chromatography. *Anal Biochem.* 56(2), 440-2.

5 Abbreviations

AAA	Amino acid analysis
ANOVA	Analysis of the variance
BCA	Bicinchoninic acid
BSA	Bovine serum albumin
CBQCA™	3-(4-carboxybenzoyl)quinoline-2-carboxaldehyde
CD	Circular dichroism
CV	Coefficient of variation
Da	Dalton
FTIR	Fourier transformed infrared
IEX	Ion exchange
LOD	Limit of detection
Mw	Relative molecular weight
PITC	Phenylisothiocyanate
PEG	Polyethylene glycol
PEG-5000 SE	O-Methyl-O'-succinylpolyethylene glycol 5000 N-succinimidyl ester
σ	Standard deviation
S/N	Signal/noise ratio
UV	Ultra violet