Chapter 19

Methods for Protein Characterization by Mass Spectrometry, Thermal Shift (ThermoFluor) Assay, and Multiangle or Static Light Scattering

Joanne E. Nettleship, James Brown, Matthew R. Groves, and Arie Geerlof

Mass spectrometry (MS) is widely used within structural and functional proteomics for a variety of tasks including protein quality assessment, identification, and characterization. MS is used routinely for the determination of the total mass of proteins, including N-glycosylated proteins, analysis of selenomethionine incorporation, crystal content verification, and analysis of N-glycosylation site occupancy. Protocols for sample preparation, data collection, and analysis are given.

A recent development is the fluorescence-based thermal shift (ThermoFluor) assay. It uses an environmentally sensitive dye, Sypro Orange, to monitor the thermal stability of a protein and investigate factors (e.g., buffers, additives, and ligands) affecting this stability. This chapter describes the application of this method using a 96-condition in-house screen. The measurements are performed on a commercially available real-time PCR machine.

Multiangle or static light scattering (SLS) is a very powerful technique to determine the conformational state of proteins in solution, especially when used in combination with size exclusion chromatography (SEC). In the authors' experimental set-up the SLS detector is connected in-line to a standard protein purification machine (e.g., the Äkta Purifier) equipped with an analytical SEC column. The data collection and analysis are performed using commercial software.

1. Introduction

Protein characterization plays an important role in two key aspects of proteomics. The first is the quality assessment of the produced protein preparations. Obtaining protein of sufficient quality for structural and/or functional studies is one of the major bottlenecks of most proteomics projects. This is reflected in the overall statistics for the major structural genomics consortia (for an overview, see http://sg.pdb.org/target_centers.html), which show that so far the structures of less than 4% of the cloned genes could be determined. Hence, it is essential to perform an extensive quality assessment of the protein preparations prior to their application and use the results in the evaluation

From: *Methods in Molecular Biology, Vol. 426: Structural Proteomics: High-throughput Methods*
Edited by: B. Kobe, M. Guss and T. Huber © Humana Press, Totowa, NJ

of the production process *(1)*. The second is the determination of protein properties such as domains, oligomeric state, posttranslational modifications and protein–protein and protein–ligand interactions.

In most laboratories SDS-PAGE and dynamic light scattering (DLS) are routinely used to determine the purity and monodispersity of a protein. However, the application of these methods is well documented in the literature; therefore, it is not considered here. This chapter gives protocols for the sample preparation and data collection and analysis of three other protein characterization methods that are routinely used in the authors' laboratories.

1.1. Protein Characterization by Mass Spectrometry

MS is a useful analytical technique and is widely used within structural proteomics consortia *(2–6)*. Analysis of both intact proteins and protein digests allow for the verification of construct along with any modifications such as selenomethionine incorporation or methylation as well as giving information on posttranslational modifications.

The protocols in this section describe how to analyze intact proteins using minimal sample preparation and automated procedures (Section 3.1.1.); remove N-glycans prior to automated LC-MS (Section 3.1.2.); and prepare proteins for LC-MS from crystals (Section 3.1.3.). A protocol for preparation of peptides by digestion of proteins followed by automated LC-MS or LC-MS/MS is given in Section 3.1.4., followed by methodology for the determination of site occupancy of N-glycosylated proteins (Section 3.1.5.).

1.2. ThermoFluor Assay

Thermal shift (ThermoFluor) assays offer a rapid and simple technique for assessing the thermal stability of proteins and to investigate factors affecting this stability *(1,7,8)*. An environmentally sensitive fluorescent dye is used to monitor protein unfolding with respect to temperature. Melting curve analysis determines the melting temperature, T_m (the midpoint of the unfolding transition); a shift in T_m under different conditions indicates a change in stability. Commercially available real-time PCR machines have sensitive thermal control and fluorescent detection capabilities, allowing the assay to be routinely performed using low amounts of protein. Thermal shift is amenable to drug screening and screening of buffer conditions, additives, ligands, and cofactors to indicate promising crystallization or storage conditions or to assign function *(1,9,10)*.

1.3. Protein Characterization by Static Light Scattering

SLS is a noninvasive technique whereby an absolute molecular mass of a protein sample in solution may be experimentally determined to an accuracy of better than 5% through exposure to low-intensity laser light (690 nm). The intensity of the scattered light is measured as a function of angle and may be analyzed to yield the molar mass, root mean square radius, and second viral coefficient (A_2). The results of an SLS experiment can be used as a quality control in protein preparation (e.g., for structural studies) in addition to the determination of solution oligomeric state (monomer/dimer, etc.). SLS experiments may be performed in either batch or chromatography modes. However, as the measurement yields the volume-averaged molecular weight of the

sample within the laser beam, it is more powerful to utilize the technique in combination with protein purification. As the measurements are performed in a flow cell, there is no loss of sample and the SLS detector can be integrated easily into standard protein purification equipment. Due to the necessity of obtaining good baselines in both 280 nm absorption measurements (UV) and light scattering (LS) measurements, SEC represents a good choice of separation media, due to the use of only a single buffer system for the entire purification. Since the light scattering and concentration are measured for each eluting fraction, the mass and size can be determined independently of the elution position. This is particularly important for protein species with nonglobular shapes, which may elute at positions distant from that predicted by the calibration curve for the column.

2. Materials

2.1. Mass Spectrometry

2.1.1. Preparation and Automated Mass Spectrometry of Intact Protein Samples

1. Protein solution(s) to be analyzed.
2. 20 µM myoglobin solution (Sigma, Gillingham, UK). Dissolve protein in water and store at −80°C in 15-µl aliquots.
3. Ultimate HPLC with autosampler (Dionex, Camberley, UK) coupled to an electrospray ionization Q-Tof Micro mass spectrometer (Waters, Manchester, UK) (see Note 1).
4. C4 PepMap 300 µ-precolumn cartridge (Dionex) (see Note 2).
5. Wash solvent: 97.3% water, 2% acetonitrile, 0.5% formic acid, 0.2% trifluoroacetic acid.
6. Solvent A: 95% water, 5% acetonitrile with 0.1% formic acid.
7. Solvent B: 20% water, 80% acetonitrile with 0.1% formic acid.
8. Skirted PCR plate (Abgene, Epsom, UK) and Pierceable Power Seals (Greiner Bio-One, Stonehouse, UK).

2.1.2. Preparation of N-Glycosylated Samples

1. Peptide-N-glycosidase F (PNGase F, Sigma). Reconstitute the PNGase F according to the manufacturer's instructions, make 5-µl aliquots, and store at −80°C.

2.1.3. Preparation of Samples from Crystals (11)

1. Extra fine long paper wicks (Hampton Research).
2. Acetonitrile.
3. Resuspension buffer: 20 mM Tris-HCl pH 7.5, 200 mM NaCl, 8 M urea.
4. Skirted PCR plate (Anachem, Luton, UK) and Pierceable Power Seals (Greiner Bio-One).

2.1.4. Preparation and Automated Mass Spectrometry of Peptide Samples

1. Sequencing grade trypsin (Promega, Southampton, UK). Make 5-µl aliquots and store at −80°C to cut down on degradation of the enzyme due to freeze–thaw cycles.
2. Ultimate HPLC with autosampler (Dionex) coupled to an electrospray ionization Q-Tof Micro mass spectrometer (Waters) (see Note 1).
3. Ubiquitin (6 µM in water).

4. Glufibrinopeptide B (3 µM in water).
5. Jupiter™ 4µ Proteo 90 Å column (Phenomenex, Macclesfield, UK).
6. Wash solvent: 97.3% water, 2% acetonitrile, 0.5% formic acid, 0.2% trifluoroacetic acid.
7. Solvent A: 95% water, 5% acetonitrile with 0.1% formic acid.
8. Solvent B: 20% water, 80% acetonitrile with 0.1% formic acid.
9. Skirted PCR plate (Anachem) and Pierceable Power Seals (Greiner Bio-One).

2.1.5. Preparation of Samples for N-Glycosylation Site Analysis and Interpretation of Results

1. Sequencing grade trypsin (Promega). Make 5-µl aliquots and store at −80°C. to cut down on degradation of the enzyme due to freeze–thaw cycles.
2. Peptide-N-glycosidase F (PNGase F). Reconstitute the PNGase F according to the manufacturer's instructions, make 5-µl aliquots and store at −80°C.
3. Zwitterion Chromatography-Hydrophilic Interaction Chromatography (ZIC-HILIC) resin (HiChrom, Berkshire, UK).
4. Isopropanol.
5. Rainin Finepoint P20 filtered tips (Anachem, catalogue number: RT-20F).
6. Wash solvent: 20% water, 80% acetonitrile with 0.1% formic acid.
7. Resuspension solvent: 100% acetonitrile with 0.1% formic acid.
8. Elution solvent: 100% water with 0.1% formic acid.
9. 1 M Tris-HCl pH 7.5, 0.5 M NaCl.
10. Skirted PCR plate (Anachem), Pierceable Power Seal (Greiner Bio-One) and Adhesive PCR seal (ABgene).
11. ProteoMass™ Guanidination Kit (Sigma) (optional depending on quality of results).

2.2. ThermoFluor Assay

1. Solution of purified protein at 1 mg/ml (see Note 11).
2. Sypro Orange dye (Molecular Probes); 5,000× stock in DMSO diluted to 10× working stock in water and stored at 4°C (see Note 12).
3. 96-well thin-walled PCR plate (e.g., Thermo-Fast 96 skirted, Abgene) (see Note 13).
4. Optically clear plate seals (e.g., Microseal "B" Film, Bio-Rad).
5. Plate sealer (not essential).
6. Buffer/additive/ligand screen. For example, buffers at 20–100 mM, salts at 25–500 mM, ligands at suitable concentrations compared to protein and other additives at concentrations comparable to those used in the Hampton Research Additive Screen (see Note 14).

2.3. Static Light Scattering

1. Protein solution(s) to be analyzed (see Notes 22 and 23).
2. Äkta Purifier (GE Heathcare) applied with a size exclusion chromatography column (see Notes 24 and 25).
3 Static light scattering detector (Wyatt Technology miniDAWN Tristar) (see Note 26).
4. BSA solution (10 mg/ml in water).

3. Methods

3.1. Mass Spectrometry

3.1.1. Preparation and Automated Mass Spectrometry of Intact Protein Samples

1. Prepare 15 µl of a 20-µM solution for every protein to be analyzed by mass spectrometry in buffer solution containing only ionic salts, for example, 20 mM Tris-HCl pH 7.5, 200 mM NaCl. Buffer components such as detergents, polyethylene glycol (PEG) or high imidazole are to be avoided for best results, since components of this nature are favorably ionized within the mass spectrometer and therefore mask the protein signal (see Fig. 19.1).
2. Pipette 15 µl 20 µM myoglobin into the first well of the PCR plate. This is used for the calibration of the instrument (see Note 3).
3. Aliquot 15 µl water into the second well to wash the precolumn, followed by the sample(s) of interest, which are pipetted into the proceeding wells. Preparation protocols for some specialized samples are given in the following (see Sections 3.1.2. and 3.1.3.).
4. The PCR plate is sealed before automated running of the samples, which can take place overnight.

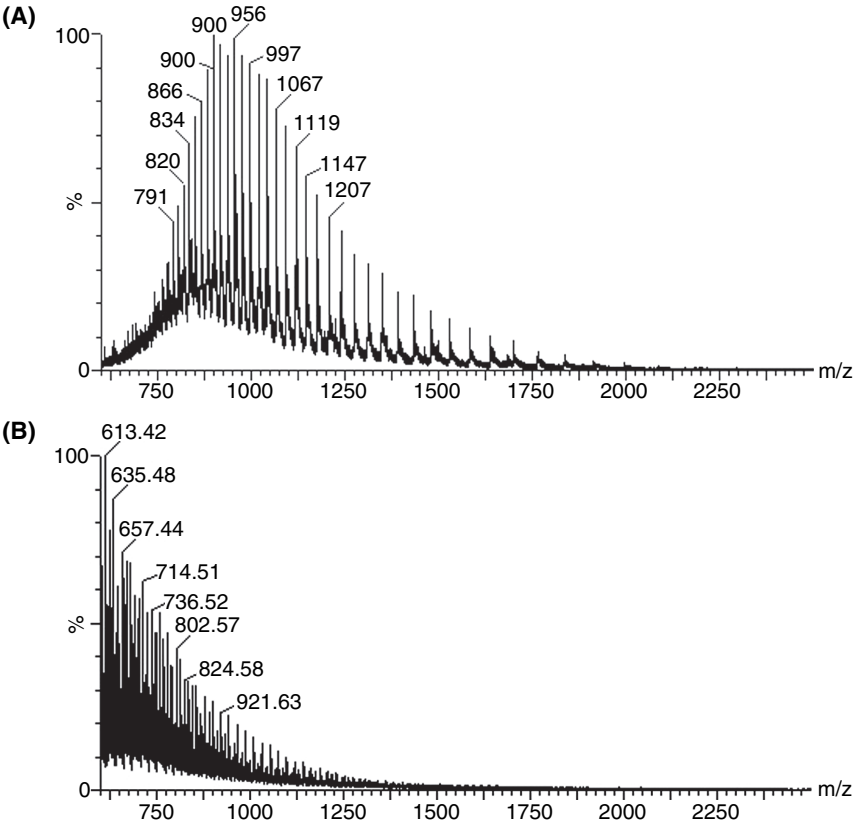

Fig. 19.1 (continued)

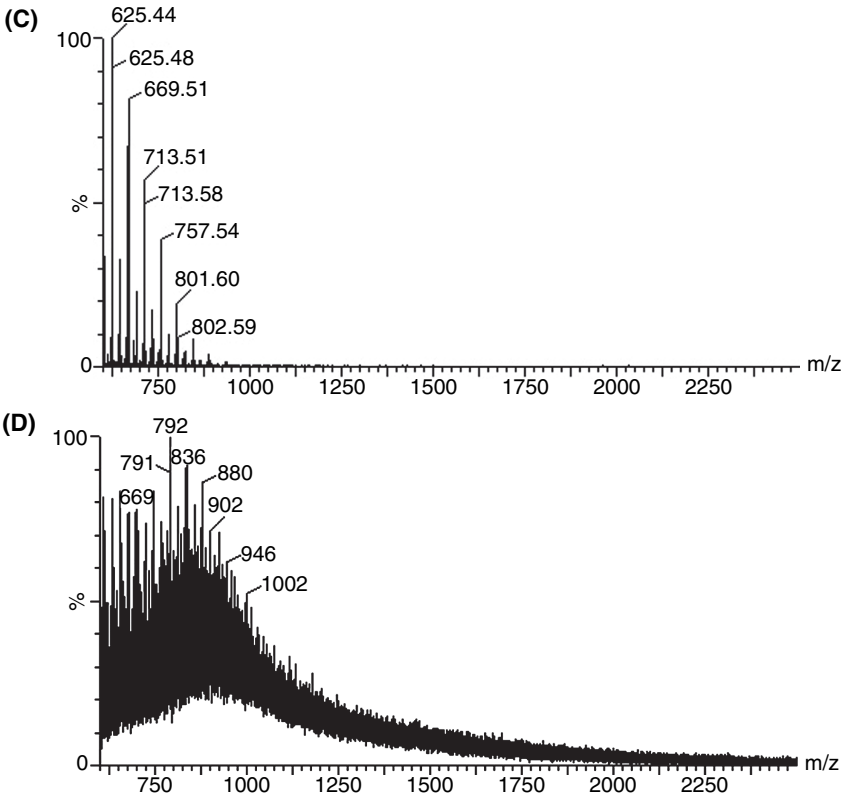

Fig. 19.1 Examples of raw data of good quality **(A)**; with 10% PEG3350 contamination **(B)**; with 1% Triton X100 (detergent) contamination **(C)**; and with 0.5 M imidazole contamination **(D)**.

5. Each sample is automatically loaded onto the C4 precolumn through the autosampler and washed with wash solvent at 5 µl/min for 5 minutes.
6. The sample is then eluted at 5 µl/min in the reverse direction directly into the mass spectrometer using a gradient of 5% to 80% solvent B over 1 minute. The concentration of mobile phase is held at 80% Solvent B for 10 minutes before re-equilibration of the column in 95% Solvent A, 5% Solvent B.
7. Analysis of samples is by MS (as opposed to MS/MS). Data is processed by combining the data under the peak on the chromatogram (Fig. 19.2A) to give a raw data file showing % ions against mass/charge (see Fig. 19.2B). This raw data is then deconvoluted to give a single mass peak using the MaxEnt algorithm (see Fig. 19.2C). The MaxEnt algorithm is part of the MassLynx software used to control the mass spectrometer.

3.1.2. Preparation of N-Glycosylated Samples

1. Boil 15 µl of a 20-µM protein sample in buffer solution containing only ionic salts for 10 minutes to denature the protein. This can be done either in a closed microcentrifuge tube or a sealed PCR plate.
2. After the solution has cooled to room temperature, add 1 µl of PNGase F and incubate at 37°C for over 3 hours (see Note 4).
3. Perform LC-MS and data analysis as described in Section 3.1.1.

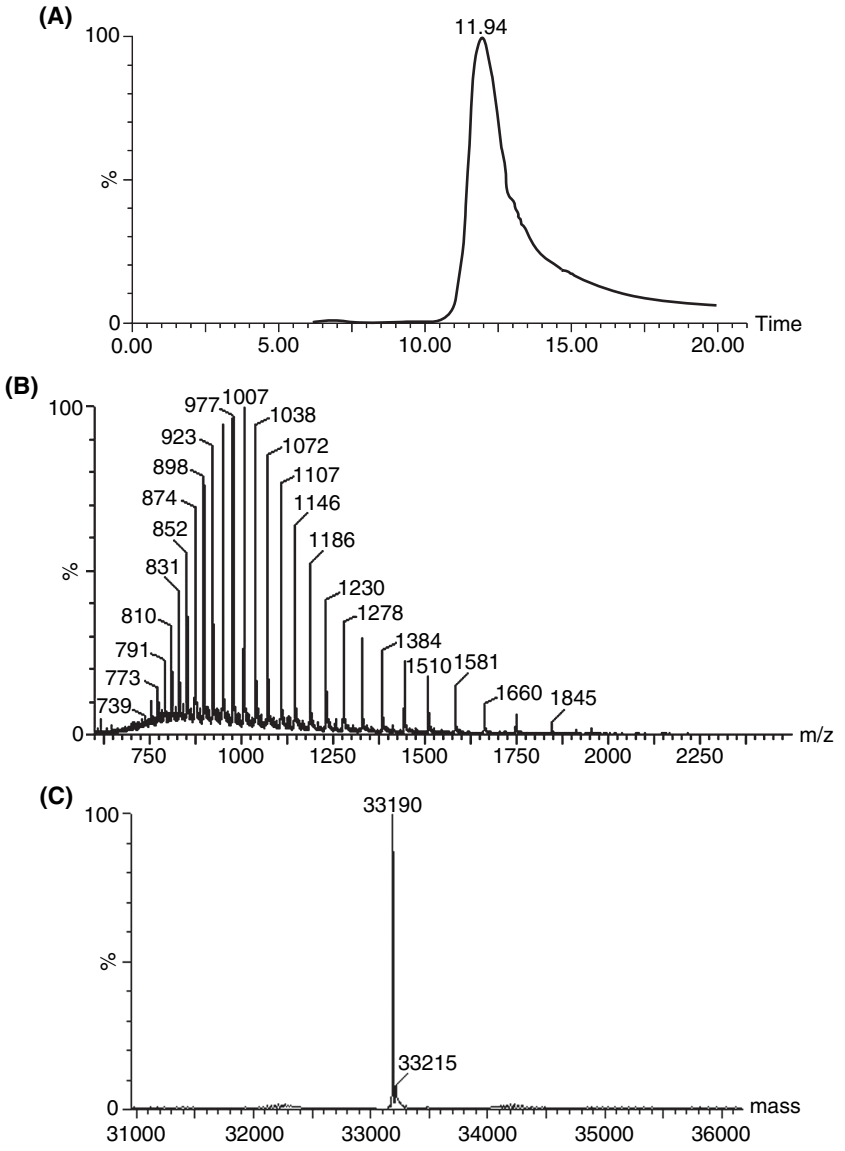

Fig. 19.2 Example data from a standard run showing the peak on the chromatogram (**A**); the raw combined data (**B**); and the deconvoluted mass peak (**C**).

3.1.3. Preparation of Samples from Crystals (11)

1. Open the well containing the crystal(s) of interest and place the crystallization plate under a microscope. This technique is limited by the concentration of protein in the crystal(s) therefore larger crystal(s) will give better data.
2. Wick away the mother liquor from around the crystal(s) using an extra fine long paper wick.
3. Carefully wash the crystal(s) in 15 µl of reservoir solution from the crystallization experiment.
4. Remove all solution from around the crystal(s) using a wick.

5. Carefully wash the crystal(s) twice with 15 µl acetonitrile, wicking away any solution in between washes. The washes remove PEG and other contaminants from the crystal(s). If any contaminants remain, this will result in less satisfactory data.
6. Dissolve the crystal(s) in 15-µl resuspension buffer and transfer to a skirted PCR plate and seal ready for mass spectrometry.
7. Analyze the sample using LC-MS as described in Section 3.1.1. (see Note 5).

3.1.4. Preparation and Automated Mass Spectrometry of Peptide Samples

1. To 15 µl of a 10-µM protein solution, add 1 µl trypsin solution and incubate at room temperature for 30 minutes (see Note 4). For glycosylation site analysis, follow the sample preparation protocol given in Section 3.1.5.
2. Aliquot calibrants into the first wells of a PCR plate. For this technique, ubiquitin (6 µM) and glufibrinopeptide B (3 µM) are used as they can be loaded into a well of the PCR plate and automatically injected onto the column as part of the run. However, accurate calibration can also be performed by direct injection using 10 µl of 0.05 µg/µl CsI in 50:50 isopropanol:water (see Note 3).
3. Pipette the sample(s) of interest into the proceeding wells and seal the plate.
4. Each sample is automatically loaded onto the Jupiter™ Proteo column through the autosampler and washed with wash solvent at 8 µl/min for 5 minutes.
5. The sample is then eluted at 10 µl/min into the mass spectrometer using a gradient of 95% solvent A:5% solvent B to 55% solvent A:45% solvent B over 40 minutes. The concentration of Solvent B is then increased to 100% over 1 minutes and held at 100% for 10 minutes before re-equilibration of the column into 95% Solvent A:5% Solvent B.
6. Peptides can be detected using either MS or MS/MS modes as required (see Note 6).
7. Analyze MS Data by combining the data under a peak of interest. MS/MS data are analyzed using the ProteinLynx Global Server™ program, which automatically processes the data and searches the online databases for proteins containing the peptides in the sample of interest.

3.1.5. Preparation of Samples for N-Glycosylation Site Analysis and Interpretation of Results

1. Pipette five samples of 20 µl of a 10-µM protein solution into five wells of the skirted PCR plate.
2. Incubate the different samples as described in step 4 and in the summary given in Fig. 19.3A. Timing of incubations can be combined as in Fig. 19.3A to allow parallelization of the experiments in the PCR plate.
3. The following types of peptide are generated by these different samples:

   - Sample 1: Nonglycosylated peptides
   - Samples 2 and 3: Both glycosylated and nonglycosylated peptides
   - Sample 4: Negative control for sample 5
   - Sample 5: Glycosylated peptides only

4. Seal the plate using an adhesive PCR seal during incubations (see Notes 4 and 7).
   - *For sample 1*: Add 1 µl of trypsin and incubate for 30 minutes at room temperature.

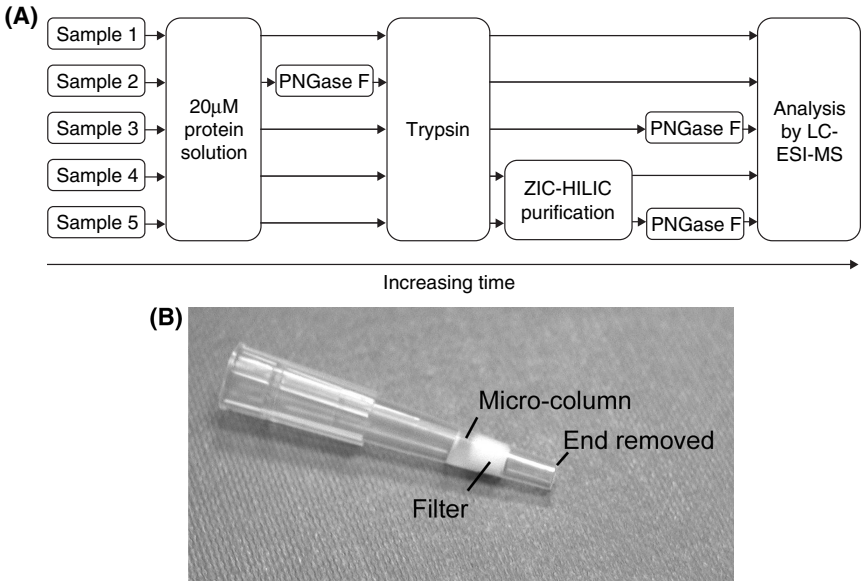

Fig. 19.3 Scheme showing the workflow for glycosylation site analysis **(A)**; and construction of the ZIC-HILIC micro-column **(B)**.

- *For sample 2*: Add 1 µl of PNGase F and incubate at 37°C for 3 hours. After the reaction is complete, add 1 µl of trypsin and incubate for 30 minutes at room temperature.
- *For sample 3*: Add 1 µl of trypsin and incubate for 30 minutes at room temperature. After the reaction is complete, add 1 µl of PNGase F and incubate at 37°C for 3 hours.
- *For sample 4*: Add 1 µl of trypsin and incubate for 30 minutes at room temperature. Then perform the ZIC-HILIC purification (see the following).
- *For sample 5*: Add 1 µl of trypsin and incubate for 30 minutes at room temperature, followed by ZIC-HILIC purification (see the following). Afterward, add 4 µl of 1 M Tris-HCl pH 7.5, 1 M NaCl to increase the pH before adding 1 µl of PNGase F. Incubate the sample for 3 hours at 37°C.

ZIC-HILIC purification:

- Resuspend the ZIC-HILIC resin in isopropanol to make a 50% (w/v) solution.
- Cut the end off of a filtered P20 tip about 5 mm after the filter and add 20 µl of resin solution above the filter to form a micro-column (see Fig. 19.3B and Note 8).
- Equilibrate the micro-column twice in 100 µl of wash solvent, using a P200 Pipetman to force the liquid through.
- Add 80 µl of resuspension solvent to the sample, which will give a final acetonitrile concentration of 80% and load the sample through the tip.
- Wash the micro-column three times in 100 µl wash solvent.
- Ensure all solvent is removed from the end of the tip before elution of the sample (see Note 9).
- Elute the sample in 20 µl of elution solvent.

5. Remove the adhesive PCR seal and replace it with a Pierceable Power Seal prior to analysis by LC-MS, which is run as in Section 3.1.4 using the MS mode of operation.
6. For analysis, perform an *in silico* digest of your protein using the Peptide Cutter tool (http://www.expasy.org/tools/peptidecutter/). Map the potential glycosylation sites onto the cleaved sequence. Glycosylation site occupancy prediction is performed using NetNGlyc (http://www.cbs.dtu.dk/services/NetNGlyc/) and NetOGlyc (http://www.cbs.dtu.dk/services/NetOGlyc/) for N- and O-glycosylation, respectively.
7. Using the calculated monoisotopic masses of the peptides containing potential glycosylation sites search for these in the MS data.
8. If the correct mass is found, this indicates the site is unoccupied. If the mass is found to be 1Da larger than calculated, the site is occupied. This is due to the PNGase F converting the asparagine to an aspartic acid during the deglycosylation reaction.
9. As indicated, the following types of peptide should be detected in the samples

   Sample 1: Nonglycosylated peptides.
   Samples 2 and 3: Both glycosylated and nonglycosylated peptides.
   Sample 4: Negative control for sample 5 (see Note 10).
   Sample 5: Glycosylated peptides only (see Note 10).

10. A consensus of the results from the five samples allows each potential site to be assigned as "glycosylated," "not glycosylated," or "partially glycosylated." The "partially glycosylated" category is for sites that appear as both glycosylated and not glycosylated in the different experiments (Fig. 19.4). It is these sites that may be mutated out to allow for secretion of a homogeneous sample.
11. If results are unsatisfactory, the signal of peptides ending in lysine can be enhanced using the ProteoMass™ Guanidination Kit *(12)*. This converts the lysine to homoarginine thus allowing the peptide to be more easily ionized.

Guanidination:

12. Add 10 µl of the Base Reagent into each well containing the digest sample and mix well.
13. Add 10 µl of Guanidination Reagent to each well and mix. Incubate the reactions at 65°C for 30 minutes.
14. Add 4 µl of 100% formic acid per well to stop the reaction instead of the 30–60 µl Stop Solution mentioned in Step 7 of the Sigma protocol. This keeps the sample volume to a minimum.
15. Run the samples by LC-MS as described previously (Section 3.1.4.).

3.2. ThermoFluor Assay

1. The thermal shift assay can be performed in any commercially available real-time PCR machine. Although certain elements of the protocol described here refer specifically to Opticon Monitor software version 3.1.32 (running a BioRad DNA Engine Opticon 2 real-time PCR machine), the general principles are applicable to other systems.
2. The software must be programmed with the necessary details for running the protocol, namely, which dye filters (see Note 12) and thermal parameters to

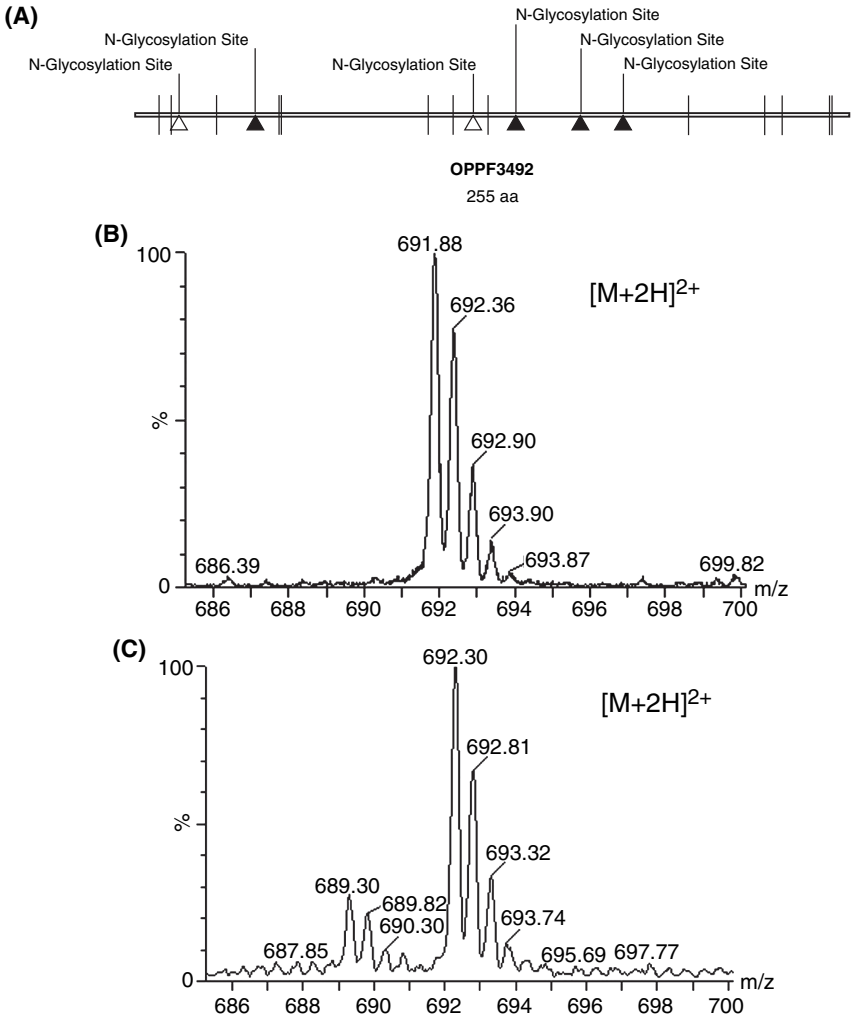

Fig. 19.4 The assigned glycosylation sites mapped onto the amino acid sequence with tryptic cleavage sites indicated by a perpendicular line, glycosylated sites by a black triangle and partially glycosylated sites by a white triangle **(A)**. Example of ESI-LC-MS data for the peptide containing the third potential glycosylation site (LSNLDPGNYSFR) showing the peak with no shift **(B)** and with a +1 Da shift **(C)**.

use. For the Opticon Monitor software, this involves specifying a plate template and running method. In the "Plate Setup" section, a new template should be programmed, choosing the appropriate plate type (clear/white) and selecting all wells to be read as samples using the SYBR Green filter (SBG1: see Note 12). In the "Protocol Setup" section, a new running method should be entered as a melting curve, with heating from 20°C to 95°C and a 15-second hold every 0.5°C, followed by a fluorescence reading. The assay volume should be entered here to allow the software to calculate the actual temperature in the reaction mixture. Also in this window, the authors set the temperature control at "Sample Calculation" with "Lid Settings" set at "Constant 101°C (shut off <20°C)." Save the templates with appropriate filenames.

3. Prepare the assay plate using a multichannel and/or repeater pipette to reduce pipetting errors. Total reaction volume is 50 µl (see Note 15). To each well, add 5 µl protein solution, 5 µl 10× Sypro Orange solution and 40 µl of whichever screen is under investigation. The order of addition is unimportant. Up to 96 additives or buffers can be tested on each plate; replicates can be used for more reliable evaluation of fewer conditions.
4. Seal the plate carefully (see Note 16) and centrifuge for 1 minute at 500*g* to mix components.
5. Place the plate into the PCR machine, close the lid and set the program running. For the Opticon Monitor software, select the appropriate plate and method, click "Run" and give a filename when prompted. After about 1 hour the assay is complete.
6. Assay progress can usually be followed in real-time. In the Opticon Monitor software, this is done via the "Status" window where the "Optical Read Status" tab allows the user to highlight wells and follow the melting curve(s).
7. When the assay is complete, the melting curves can be analyzed and T_ms calculated. Some software packages perform these calculations automatically while others require that the data be exported into statistical analysis software where curve fitting can be performed to extract T_ms. The "Melting Curve" window of the Opticon Monitor software calculates the T_m from the maximum value of the first derivative curve of the melting curve (Fig. 19.5). In the "Graphs" tab, selecting a well displays the data for that well, with radio buttons allowing display of the intensity curve, the first derivative curve or both. The "Calculations" tab at the bottom of the window displays the T_ms. An ideal melting curve would be sigmoidal but this is not always the case (see Notes 17–19) and adjustment of the "Peak Location Boundaries" in the "Graphs" tab might be necessary to define the correct region for the calculation. Care must be taken with the interpretation and it is advisable to visually inspect the melting curves before accepting the calculated T_ms. A "smooth" function which removes noise from the curves can also affect calculated T_ms and should be used with care (see Note 20).
8. Upon completion of the protocol, all melting curves can be compared and the results correlated with the various screens (see Note 21). Running multiple samples allows comparison of data from well-to-well; while large thermal shifts are obvious, such direct comparison can aid in the detection of smaller shifts. If necessary, raw data can usually be exported for curve fitting and further analysis.
9. Ideally, follow-up experiments should be performed using titrations of ligand or additive and assaying in replicate.

3.3. Static Light Scattering

3.3.1. Experimental Set-up and Calibration

1. For SLS measurements in the chromatography mode the miniDAWN Tristar (SLS detector) should be placed in the flow path of the Äkta Purifier. The instrument is typically equipped with a Superdex 200 10/300 GL SEC column. The SLS detector uses the UV signal from the Äkta Purifier to measure protein concentrations and care needs to be taken to ensure that a suitable delay volume is programmed into the ASTRA software (Wyatt Technology, Santa Barbara, CA) to correct for the difference in flow path between the UV

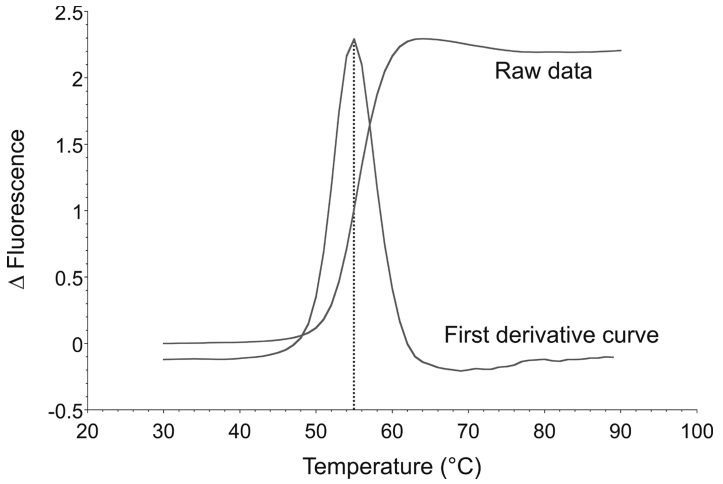

Fig. 19.5 Example thermal shift data showing the raw data curve alongside the first derivative curve. The temperature at the peak of the first derivative curve is the melting temperature, T_m *(dotted line)*.

and SLS measurements. This can be calculated from the difference in elution volume between the UV and SLS signals. A simple way to ensure the correct determination of the delay volume is available through the analysis software (see "Alignment" in the "View" menu). An overlaid display is given of both the SLS and UV signals and a right mouse click and drag between the two respective peaks allows the user to manually overlay the two signals. High precision is achievable by zooming in on the peaks (Control-left mouse to drag a zoom area). Once this delay volume is defined it should only need to be redefined if tubing length between the UV and SLS detectors is altered.

2. A 0.22-µm prefilter should be placed immediately upstream of the SLS detector in order to remove large particles (e.g., produced by the pumps), which will disturb the measurements (see Note 27).
3. The system is equilibrated with 2–3 column volumes of an appropriate buffer (see Note 28).
4. Once the system has been correctly set up and equilibrated a standard protein is used for calibration purposes. Bovine serum albumin (BSA) represents a good choice of calibration sample, as it forms a number of known oligomeric states in solution (monomer, dimer, and tetramer of 66-kDa subunits).

The calibration is performed as follows:

1. Inject 100 µl BSA solution and start data collection (see Section 3.3.2.).
2. Adjust the value of the "AUX1 calibration constant" in the "system set-up" window of the "Collect" menu during data analysis until the correct mass for BSA is obtained (see Section 3.3.3.). This value needs to be determined periodically to correct for changes in the intensity of the UV light.

However, any protein of "known" molecular weight and specific absorption coefficient may be used and after initial installation and calibration it is recommended that the user confirms the calibration using, e.g., lysozyme (molecular weight of 14 kDa).

3.3.2. Data Collection

1. Start a new experiment by selecting "New" from the "File" menu of the ASTRA software. A window appears with two screens monitoring the UV and LS signals.
2. Set-up the experimental parameters in the "Collect" menu. Enter in the "System set-up" window "solvent" (typically water), "flowrate" and "AUX1 calibration constant." Correct setting of the flow rate is essential as the device uses timers to define data collection windows over the required volume ranges. Hence incorrect flow rates will result in miscollection of the data. The new "AUX1 calibration constant" is calculated by correcting the value determined during the BSA calibration (see Section 3.3.1.) for the difference in the absorption coefficients of the sample protein and BSA:

$$(\mathrm{AUX1})_{\mathrm{new}} = (\mathrm{AUX1})_{\mathrm{BSA\ calibration}} \times A^{0.1\%}_{280,\,BSA} \,/\, \times A^{0.1\%}_{280,\,sample}$$

   Where $(\mathrm{AUX1})_{\mathrm{BSA\ calibration}}$ is the AUX1 calibration constant determined during BSA calibration; $A^{0.1\%}_{280,\,sample}$ is the absorption coefficient at 280 nm for a 1 mg/ml sample solution (see Note 29); $A^{0.1\%}_{280,\,BSA}$ is the absorption coefficient at 280 nm for a 1 mg/ml BSA solution ($0.68\,\mathrm{cm}^{-1}$).

   Enter in the "Collection set-up" window the following parameters: "operator," "sample ID," "injection-to-collect-delay volume," "collection duration," and "collection interval" (see Note 30). Save the experiment.
3. Fill the injection loop with 100 µl of the sample protein solution (see Note 22).
4. Start the chromatography method on the Äkta Purifier (see Note 31).
5. The data collection on the SLS detector has to be started manually: choose "single injection" in the "Collect" menu. Press "OK" to start data collection at the same time the sample is injected on the SEC column (indicated by the shift of the injection valve).

3.3.3. Data Analysis

The data obtained from the SLS detector is analyzed using the ASTRA software.

1. Choose "Baselines" in the "View" menu. Draw the baselines in the UV and the three LS signals by clicking on the traces in the linear parts to either side of the peak(s). It is worthwhile optimizing the baselines since they are crucial for the outcome of the data analysis. An option is available to define a baseline only in the second LS window, through "Auto baseline" in the "Options" menu, but the user is strongly recommended to visually check the first and third windows manually.
2. Select "Peaks" in the "View" menu. To check if the peaks in the UV and LS signals overlay choose two traces (AUX1 and one LS) by clicking on the "Data" button in the window. After a second click on "Data" both traces appear in the window. This can be repeated for the other LS signals. When the peaks overlay well select the area of the peak(s) to be analyzed by left mouse clicking and dragging over the area to be analyzed. The selected area is temporarily visualized by a gray bar and can be adjusted if needed.
3. Select "Report" in the "View" menu and choose "Summary." This window summarizes the results of the data analysis. A graphical display of the

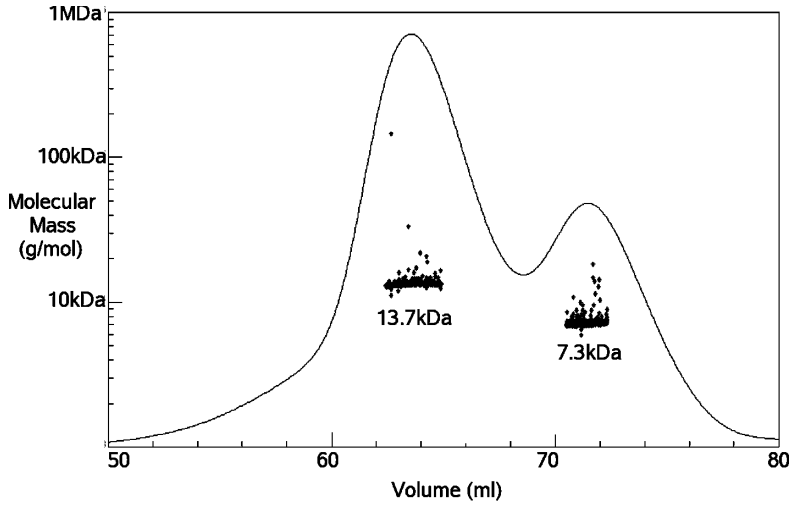

Fig. 19.6 Analysis of the solution oligomeric state of a 7.0-kDa protein by static light scattering in combination with size exclusion chromatography. The separation of the oligomers was performed on a HiLoad 16/60 Superdex 75 column (GE Healthcare). The light scattering signal *(dots)* is shown as the mass distribution in a slice in each of the two peaks in the elution profile monitored by the absorbance at 280 nm *(solid line)*. The molecular masses were calculated to be 7.3 and 13.7 kDa, respectively, which correlates well with the monomeric and dimeric protein forms.

masses present in each data slice is available through "MM vs. Volume" in the "Distribution" menu (Fig. 19.6). In order to measure only regions of the chromatogram in which a single species is eluting it may be necessary to adjust the analysis region (see Step 2) such that a horizontal line is obtained. The presence of nonhorizontal sections within the analysis area indicates the presence of two species, whose relative concentrations are not constant—leading to a change in the mass averaged molecular weights of these sections.

4. Notes

1. All protocols can be adapted for use with LC-MS systems from other manufacturers.
2. C4 precolumns can be reused, however some samples "stick" to the column and re-elute in all subsequent samples. On average each precolumn lasts for around 2 months.
3. Calibration can be performed either before running the samples or retrospectively and then applied to relevant data during analysis.
4. As PNGase F reactions and tryptic digests need to go to completion and protein denaturation is not an issue, longer incubation times—up to overnight—can be used. This will ensure a complete reaction without jeopardizing the sample.

5. When running samples prepared from crystals, it is essential to either use a new precolumn on the LC-MS or to make sure the precolumn is clean as the concentration of protein is likely to be below ideal.
6. MS analysis gives the accurate mass of a peptide, whereas MS/MS analysis gives both the accurate mass and fragmentation data. The fragmentation data can be used to give information on the sequence of the peptide. MS/MS data for a digested protein provide a mass fingerprint, which can be used to identify a protein by BLASTing the fingerprint against databases such as SwissProt.
7. This protocol can be used with digestion enzymes other than trypsin, which may lead to more informative results depending on the positioning of the cleavage sites in relation to putative glycosylation sites in the protein of interest. Other enzymes used routinely are chymotrypsin, glutamyl endopeptidase, thermolysin, and lysyl endopeptidase, although any digestion enzyme is appropriate for this method.
8. The ZIC-HILIC resin will migrate into the filter of tips from some manufacturers; however, Rainin Finepoint tips have proved successful for this method.
9. Removal of wash solvent or sample from the end of the tip can be performed using a P20 pipetman.
10. The ZIC-HILIC resin binds glycosylated peptides; however, some non-glycosylated peptides do bind such as a peptide containing a His_6 tag. Purification of the glycosylated peptides away from non-glycosylated ones increases the chance of detection
11. The authors' standard protocol uses protein at 1 mg/ml, with each well containing 5 µg protein. This amount was selected to ensure that most proteins give an adequate signal when first assayed. As little as 1.5 µg have been used per well with reproducible results.
12. The dye generally used is Sypro Orange, which fluoresces strongly when bound in the hydrophobic regions exposed in unfolded proteins. Its fluorescent properties are similar to those of SYBR Green, commonly used in real-time PCR applications, and hence the SYBR Green filters on real-time PCR machines can be used for excitation and detection.
13. Various types of plate are available depending on which PCR machine is being used (white or clear plates, skirted or non-skirted, and so on). Although the authors prefer opaque white plates, they have used clear plates with success.
14. The authors use a 96-conditions in-house screen (Table 19.1) for standard thermal shift assays which is stored as a master block at −20°C. As well as testing protein stability in several buffer types at different pHs and varying salt concentrations, the screen also assesses an assortment of metal salts, nucleotide cofactors, sugars, reducing agents, and other chemicals. When required, the block is defrosted and 40 µl are taken from each well for assaying. Each block contains enough material for five thermal shift assays.
15. The basic experimental conditions are widely applicable to different proteins. Assay volume and protein concentration can be lower (as low as 15 µl with as little as 1.5 µg protein/well), but this may not give an adequate signal for all proteins.
16. Care should be taken not to touch the seals with ungloved hands. Wear gloves to avoid smearing across the surface, which may affect optical properties.

Table 19.1 The 96 conditions of the in-house screen, which is routinely used for the thermal shift assay

	1	2	3	4	5	6	7	8	9	10	11	12
A	50 mM Sodium acetate pH 4.0	50 mM Sodium acetate pH 4.4	50 mM Citric acid pH 5.0	50 mM Citric acid pH 5.4	50 mM Sodium cacodylate pH 6.0	50 mM Sodium cacodylate pH 6.4	50 mM HEPES pH 7.0	50 mM HEPES pH 7.4	50 mM Tris-HCl pH 8.0	50 mM Tris-HCl pH 8.4	50 mM CAPSO pH 9.0	50 mM CAPSO pH 9.4
B	50 mM Sodium acetate pH4.0 100mM NaCl	50mM Sodium acetate pH4.4 100mM NaCl	50mM Citric acid pH5.0 100mM NaCl	50 mM Citric acid pH 5.4 100 mM NaCl	50 mM Sodium cacodylate pH 6.0 100 mM NaCl	50 mM Sodium cacodylate pH 6.4 100 mM NaCl	50 mM HEPES pH 7.0 100 mM NaCl	50 mM HEPES pH 7.4 100 mM NaCl	50 mM Tris-HCl pH 8.0 100 mM NaCl	50 mM Tris-HCl pH 8.4 100 mM NaCl	50 mM CAPSO pH 9.0 100 mM NaCl	50 mM CAPSO pH 9.4 100 mM NaCl
C	50 mM Sodium acetate pH4.0 200 mM NaCl	50 mM Sodium acetate pH4.4 200 mM NaCl	50 mM Citric acid pH5.0 200 mM NaCl	50 mM Citric acid pH5.4 200 mM NaCl	50 mM Sodium cacodylate pH 6.0 200 mM NaCl	50 mM Sodium cacodylate pH 6.4 200 mM NaCl	50 mM HEPES pH 7.0 200 mM NaCl	50 mM HEPES pH7.4 200 mM NaCl	50 mM Tris-HCl pH 8.0 200 mM NaCl	50 mM Tris-HCl pH8.4 200mM NaCl	50 mM CAPSO pH 9.0 200 mM NaCl	50 mM CAPSO pH 9.4 200 mM NaCl
D	50mM Sodium acetate pH4.0 500 mM NaCl	50mM Sodium acetate pH4.4 500 mM NaCl	50mM Citric acid pH5.0 500 mM NaCl	50 mM Citric acid pH 5.4 500 mM NaCl	50 mM Sodium cacodylate pH 6.0 500 mM NaCl	50 mM Sodium cacodylate pH 6.4 500 mM NaCl	50 mM HEPES pH7.0 500 mM NaCl	50 mM HEPES pH 7.4 500 mM NaCl	50 mM Tris-HCl pH 8.0 500 mM NaCl	50 mM Tris-HCl pH 8.4 50 mM NaCl	50 mM CAPSO pH 9.0 500 mM NaCl	50 mM CAPSO pH 9.4 500 mM NaCl
E	2 mM ATP	2 mM ADP	2mM AMP	2 mM AMPCCP	2 mM AMPcPP	2 mM AMPPcP	2 mM UTP	2 mM GTP	2 mM GDP	2 mM GTP-γS	2 mM TMP	2 mM FAD
F	2 mM β-NAD	2 mM β–γ MethylGTP	2mM dCMP	2 mM dGMP	ssDNA 7mer	ssDNA 9mer	1% glycerol	5% glycerol	10% glycerol	20% glycerol	1 mM DTT	5 mM DTT
G	10 mM $CaCl_2$	10 mM $MgCl_2$	10mM $MnCl_2$	10 mM $ZnCl_2$	10 mM $FeCl_3$	100 mM KCl	100 mM LiCl	200 mM NaThyiocyanate	10 mM L-Proline	10 mM Phenol	3% DMSO	10 mM $NiCl_2$
H	100 mM Glycine	10 mM Spermidine	10mM Urea	5% PEG 400	3% D(+)-Glucose	3% D-Galactose	10 mM L-Alanine	10 mM L-Methionine	10 mM L-Serine	10 mM L-Arginine	0.5% n-Octyl-β-D-glucoside	0.5% n-Dodecyl b-D-maltoside

The screen has been composed by Erika Mancini and Christian Siebold (Oxford, UK).

17. Not all proteins give good melting curves with a defined transition between folded and unfolded. This does not necessarily reflect instability, since the authors have seen poor melting curves from otherwise well-behaved and crystallizable proteins.
18. Following the complete unfolding of the protein (where fluorescence peaks), it is common to see a drop off in fluorescence, possibly due to exclusion of the dye during aggregation of denatured protein. Also, some proteins display a slight initial drop in fluorescence before the true transition begins, perhaps due to dye binding weakly in hydrophobic surface pockets and dissociating at low temperatures.
19. Sometimes, other interactions can be observed in the thermal shift profile, e.g., a plateau partway along the melting curve indicating melting of oligomers or complexes before the transition corresponding to unfolding of the individual proteins.
20. Thermal shifts are frequently significant enough to see clearly and there is often no need for accurate T_m determination. Following visual inspection, the authors generally take T_m values as calculated by the Opticon Monitor software. If necessary, data can be exported to other software for potentially more accurate T_m determination via curve fitting.
21. We have observed thermal shifts in excess of 12°C, but it is possible to see much smaller shifts of less than 1°C, especially when melting curves are compared side by side. For very low shifts, although they might turn out to be significant, experimental error must be considered as a factor. If necessary, this could be investigated by further testing in replicate.
22. The nature of the protein affects the upper and lower amounts that can be accurately analyzed. Proteins with a relatively high absorption coefficient ($A_{280}^{0.1\%} > 0.7\,cm^{-1}$) will produce significantly more deflection in the UV measurements than those with low absorption coefficients ($A_{280}^{0.1\%} < 0.5\,cm^{-1}$) for identical sample loading. It is possible to overload the UV signal used by the SLS detector, so care must be taken that sample concentrations do not become excessive. Conversely, insufficient sample concentration in the elution fractions will result in weak LS and UV signals. Such weak signals then result in large errors in the analysis. As a general rule of thumb, 100 µg to 1 mg of protein sample applied to a Superdex 200 10/300 GL column (GE Healthcare, Piscataway, NJ) result in reasonable LS and UV signals. When a preparative SEC column is used [e.g., the HiLoad 16/60 Superdex 200, GE Healthcare] 5 to 10 mg of sample should be applied.
23. The results of the SLS experiment are volume averaged. Hence, the best results are obtained when highly pure samples are analyzed by SEC since it is relatively easy to separate different oligomeric species of one protein from one another. However, when impure samples are applied increased errors appear due to insufficient separation of different protein species.
24. Other protein purification equipment can be used provided the UV signal of the instrument can be used by the SLS detector.
25. Both analytical and preparative SEC columns can be used for the analysis (see Note 22).
26. This protocol is based on the authors' experience with the Wyatt Technology mini-DAWN Tristar detector connected in-line with an Äkta Purifier. Other equipment combinations are also possible in order to

increase the accuracy of the method (e.g., the installation of a refractive index concentration measurement device). Such additional equipment will improve the accuracy of the technique, with an attendant increase in the cost of the installation. There are currently at least two companies which provide SLS detectors: Wyatt Technology (http://www.wyatt.com/) and Viscotek GmbH (http://www.viscotek.com). The reader is strongly recommended to browse the company Web pages for further information.

27. The addition of a prefilter immediately upstream of the SLS detector is a sensible step to prevent particles entering the device and affecting measurements. A standard 0.22-μm paper filter supported by a metal frit is sufficient for this purpose, although the backpressure of the system needs to be monitored. Increases in the backpressure may indicate that the filter has become clogged and needs replacing. Care also needs to be taken that the column itself suffers no damage due to increases in backpressure at standard operating flow rates. In the authors' experience it is permissible to add the measured backpressure of the filter and SLS detector to the maximum recommended backpressure rating of the column(s) used. This value can be determined by selecting a column bypass and noting current pump pressure (typically 0.3 MPa). Thus the Superdex 200 10/300 GL column (rated for a maximum backpressure of 1.5 MPa) can now be run at 1.8 MPa.
28. The SLS measurement is highly sensitive to baseline errors in both the UV and LS signals and as a result a thorough equilibration is needed for precise SLS measurements. Typically 2–3 column volumes of buffer represents a suitable equilibration volume for size exclusion columns. The components of the buffer should also not present too strong a background in either absorption of the UV signal or the LS signal. High concentrations of reducing agents (e.g., dithiothreitol) or glycerol should also be avoided if possible. If essential for the experiment these buffers should be prepared immediately prior to the experiment. As is true for all size exclusion experiments the buffer should be well filtered and degassed prior to use.
29. Inaccurate absorption coefficients represent the major source of errors in the data analysis. While absorption coefficients generated from the linear sequence are, in the main, sufficiently accurate it is recommended that UV absorption spectra are recorded of the sample in both the size exclusion buffer and a denaturing buffer in order to experimentally establish the extinction coefficients of the folded proteins. The presence of cofactors (e.g., nucleotides), which have significant absorption at 280 nm, also results in inaccurate concentration estimates from the UV signal. Additionally, samples that show significant absorption at the wavelength used for the SLS measurements (690 nm) result in errors in the LS measurements.
30. The maximal number of data points that can be collected per experiment is 14,400. Hence the minimal collection interval that can be chosen depends on the collection duration. For an experiment performed with a Superdex 200 10/300 GL column a typical collection duration is 30 minutes, which allows a collection interval of 0.125 seconds.
31. To allow simultaneous sample injection and start of the data collection it is advisable to program a short column equilibration step (e.g., 0.1 column volumes for the Superdex 200 10/300 GL column) in the chromatography method.

Acknowledgments

The work described here was supported by the MRC and the European Commission as SPINE, contract-no. QLG2-CT-2002-00988.

References

1. Geerlof, A., Brown, J., Coutard, B., Egloff, M.-P., Enguita, F. J., Fogg, M. J., Gilbert, R. J. C., Groves, M. R., Haouz, A., Nettleship, J. E., Owens, R. J., Ruff, M., Sainsbury, S., Svergun, D. I., and Wilmanns, M. (2006) The impact of protein characterization in structural proteomics. *Acta Crystallogr. D Biol. Crystallogr*. **62**, 1125–1136.
2. Yokoyama, S., Hirota, H., Kigawa, T., Yabuki, T., Shirouzu, M., Terada, T., Ito, Y., Matsuo, Y., Kuroda, Y., Nishimura, Y., Kyogoku, Y., Miki, K., Masui, R., and Kuramitsu, S. (2000) Structural genomics projects in Japan. *Nat. Struct. Biol*. **7**, 943–945.
3. Jeon, W. B., Aceti, D. J., Bingman, C. A., Vojtik, F. C., Olson, A. C., Ellefson, J. M., McCombs, J. E., Sreenath, H. K., Blommel, P. G., Seder, K. D., Burns, B. T., Geetha, H. V., Harms, A. C., Sabat, G., Sussman, M. R., Fox, B. G., and Phillips, G. N., Jr. (2005) High-throughput purification and quality assurance of *Arabidopsis thaliana* proteins for eukaryotic structural genomics. *J. Struct. Funct. Genomics* **6**, 143–147.
4. Pantazatos, D., Kim, J. S., Klock, H. E., Stevens, R. C., Wilson, I. A., Lesley, S. A., and Woods, V. L., Jr. (2004) Rapid refinement of crystallographic protein construct definition employing enhanced hydrogen/deuterium exchange MS. *Proc. Natl. Acad. Sci. USA* **101**, 751–756.
5. Jawhari, A., Boussert, S., Lamour, V., Atkinson, R. A., Kieffer, B., Poch, O., Potier, N., van Dorsselaer, A., Moras, D., and Poterszman, A. (2004) Domain architecture of the p62 subunit from the human transcription/repair factor TFIIH deduced by limited proteolysis and mass spectrometry analysis. *Biochemistry* **43**, 14420–14430.
6. Potier, N., Billas, I. M., Steinmetz, A., Schaeffer, C., van Dorsselaer, A., Moras, D., and Renaud, J. P. (2003) Using nondenaturing mass spectrometry to detect fortuitous ligands in orphan nuclear receptors. *Protein Sci*. **12**, 725–733.
7. Pantoliano, M. W., Petrella, E. C., Kwasnoski, J. D., Lobanov, V. S., Myslik, J., Graf, E., Carver, T., Asel, E., Springer, B. A., Lane, P., and Salemme, F. R. (2001) High-density miniaturized thermal shift assays as a general strategy for drug discovery. *J. Biomol. Screen*. **6**, 429–440.
8. Lo, M. C., Aulabaugh, A., Jin, G., Cowling, R., Bard, J., Malamas, M., and Ellestad, G. (2004) Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Anal. Biochem*. **332**, 153–159.
9. Carver, T. E., Bordeau, B., Cummings, M. D., Petrella, E. C., Pucci, M. J., Zawadzke, L. E., Dougherty, B. A., Tredup, J. A., Bryson, J. W., Yanchunas, J., Jr., Doyle, M. L., Witmer, M. R., Nelen, M. I., DesJarlais, R. L., Jaeger, E. P., Devine, H., Asel, E. D., Springer, B. A., Bone, R., Salemme, F. R., and Todd, M. J. (2005) Decrypting the biochemical function of an essential gene from *Streptococcus pneumoniae* using ThermoFluor technology. *J. Biol. Chem*. **280**, 11704–11712.
10. Matulis, D., Kranz, J. K., Salemme, F. R., and Todd, M. J. (2005) Thermodynamic stability of carbonic anhydrase: measurements of binding affinity and stoichiometry using ThermoFluor. *Biochemistry* **44**, 5258–5266.
11. Nettleship, J. E., Walter, T. S., Aplin, R., Stammers, D. K., and Owens, R. J. (2005) Sample preparation and mass-spectrometric characterization of crystal-derived protein samples. *Acta Crystallogr. D Biol. Crystallogr*. **61**, 643–645.
12. Brancia, F. L., Oliver, S. G., and Gaskell, S. J. (2000) Improved matrix-assisted laser desorption/ionization mass spectrometric analysis of tryptic hydrolysates of proteins following guanidination of lysine-containing peptides. *Rapid Commun. Mass Spectrom*. **14**, 2070–2073.