Detection of Protein Glycosylation Using Tip-Enhanced Raman Scattering

David P. Cowcher,[†] Tanja Deckert-Gaudig,[‡] Victoria L. Brewster,[†] Lorna Ashton,[†,§] Volker Deckert,[*,‡,||] and Royston Goodacre[*,†]

[†]School of Chemistry and Manchester Institute of Biotechnology, University of Manchester, 131 Princess Street, Manchester, M1 7DN, U.K.

[‡]Leibniz-Institute of Photonic Technology—IPHT, Albert-Einstein-Strasse 9, 07745 Jena, Germany

[§]Department of Chemistry, Faraday Building, Lancaster University, Lancaster, LA1 4YB, U.K.

[||]Institut für Physikalische Chemie and Abbe Center of Photonics, Friedrich-Schiller Universität, Helmholtzweg 4, 07743 Jena, Germany

S *Supporting Information*

ABSTRACT: The correct glycosylation of biopharmaceutical glycoproteins and their formulations is essential for them to have the desired therapeutic effect on the patient. It has recently been shown that Raman spectroscopy can be used to quantify the proportion of glycosylated protein from mixtures of native and glycosylated forms of bovine pancreatic ribonuclease (RNase). Here we show the first steps toward not only the detection of glycosylation status but the characterization of glycans themselves from just a few protein molecules at a time using tip-enhanced Raman scattering (TERS). While this technique generates complex data that are very dependent on the protein orientation, with the careful development of combined data preprocessing, univariate and multivariate analysis techniques, we have shown that we can distinguish between the native and glycosylated forms of RNase. Many glycoproteins contain populations of subtly different glycoforms; therefore, with stricter orientation control, we believe this has the potential to lead to further glycan characterization using TERS, which would have use in biopharmaceutical synthesis and formulation research.
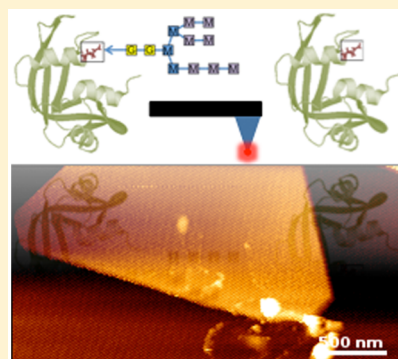
It is estimated that glycoproteins account for almost two-thirds of all protein species,[1] and with the level of research and investment into protein-based therapeutic products ever increasing,[2] the accurate characterization of post-translational modifications (PTMs) is vital for therapy. In particular, determining glycosylation status and glycan structure is becoming an important area of analytical science due to the potential adverse drug reactions for incorrect formulations and the need to have the correct protein glycoform for efficacious therapy.

In addition to the need for protein therapeutics to maintain the correct secondary and tertiary structure from the point of manufacture to their intended point of interaction with the patient, glycoproteins need to have the correct glycan attached in the correct place in order to function as intended. Incorrect glycosylation may result in misfolding, attenuation of efficacy as a result of compromised sorting/directing, ligand binding, biological activity, plasma half-life, stability, and immunogenicity.[3]

Raman spectroscopy is an increasingly popular analytical tool in the field of biomedicine[4] and has been frequently used for the characterization of biopharmaceuticals.[5] Raman has particular advantages in that it is nondestructive and can be applied through a transparent window into a vessel, giving it the potential for online use in the analysis of a dynamic system.[6] Raman has been used to characterize glycoproteins in the past,[7,8] as well as Raman optical activity (ROA),[9−12] including spectral differences as a result of chemical deglycosylation.[13] A combination of Raman spectroscopy and chemometric techniques has enabled the distinction between native and glycosylated forms of bovine pancreatic ribonuclease (RNase) protein, including quantification of the relative amounts of each form from mixtures.[14] The quantitative detection of glycated hemoglobin using Raman spectroscopy has also been demonstrated.[15]

Even though Raman microscopy (with a typical interrogation diameter of 1 μm) is a confocal technique, the number of molecules within the laser focus is still vast and the spectra recorded are thus ensemble averages of multiple layers of protein molecules. While the majority of these molecules may be correctly glycosylated, if a small percentage (<5%) are not, it is highly unlikely they would be detected.[14] From a formulation point of view, there is also no information about the specific error in the glycosylation, for example, whether the protein is
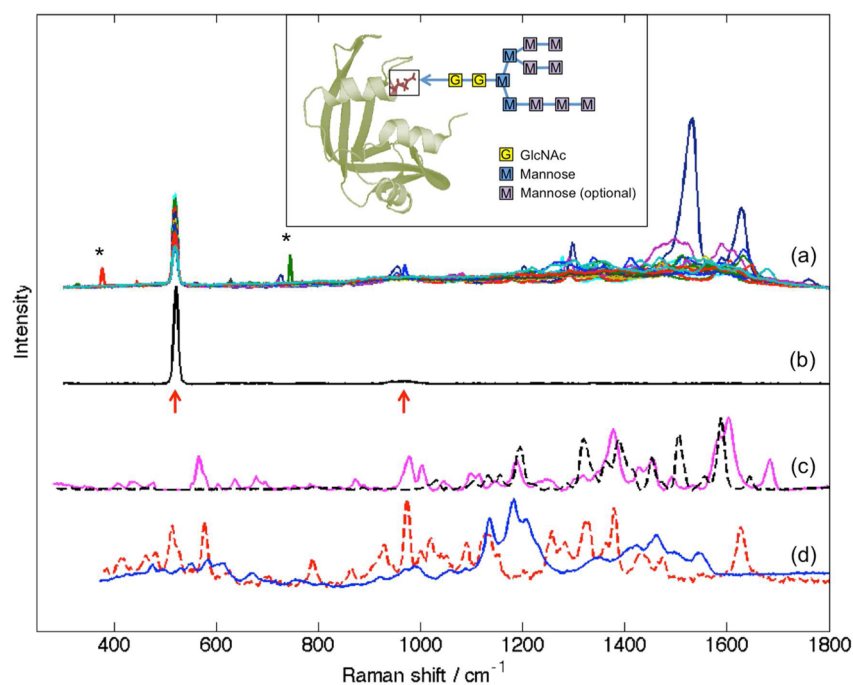
Figure 1. (a) Raw TERS spectra of RNase A from one grid with examples of cosmic rays highlighted by asterisks. (b) A typical background TERS spectrum with peaks from the silicon tip highlighted with arrows. Also shown are examples of the (c) TERS and (d) Raman spectra of mannose (solid line) and *N*-acetylglucosamine (dashed line), the sugars that make up the glycan of RNase B. The TERS spectra (c) have had the silicon background removed. Inset shows a cartoon representation of the native form of bovine RNase drawn from atomic coordinates in the PDB (5RSA) using PyMOL. Also shown is the glycosylation point (Asn34 residue) and the RNase B glycan. Optional mannose refers to the variation in number and possible arrangements of mannose between the different glycoforms that occur in RNase B.

glycosylated in the wrong place, or whether the incorrect glycan is attached. Lazar et al. have constructed a useful review of the current progress in the analysis of glycoproteins and their glycans, particularly in relation to mass spectrometry.[16] However, a rapidly advancing variant of Raman spectroscopy, tip-enhanced Raman scattering (TERS), may offer an alternative solution. TERS, like its more widely known relation surface-enhanced Raman scattering (SERS), exploits the phenomenon of metals with appropriately (nano)-sized features being able to localize regions of charge density at their surface when illuminated by appropriate electromagnetic radiation. These so-called surface plasmons are able to interact constructively with the electric field component of the incident light, resulting in an enhancement of many orders of magnitude to the light intensity, thus like SERS significantly amplifying the intensity of Raman scattered photons.[17] While SERS uses nanopatterned surfaces or colloids to enhance microscopic or bulk-phase measurements,[18] TERS utilizes a single nano-particle, mounted on the end of the tip of a scanning probe microscope, such as an atomic force microscope (AFM). The tip is scanned across a sample and provides the spatial resolution of AFM while collecting the vibrational structural information of an enhanced Raman measurement.[19]

TERS has already been used in a number of high-resolution biological applications, for example, in the analysis of live bacterial cells[20,21] and the investigation of individual nucleic acid strands;[22] the reader is also directed to a nice review by Treffer et al.[23] for further application areas.

In this study we set out to explore the potential of TERS to probe protein glycosylation at a near-molecular level by immobilizing a monolayer of protein molecules on a surface and taking TERS measurements. The high spatial resolution of

the technique has the potential to detect subtle structural variations that indicate incorrect glycosylation, variations that would very likely be masked by the ensemble averaging effect of normal Raman microscopy. Here we show the first steps toward realizing this goal using the same simple RNase model system that was employed with conventional Raman spectroscopy.[14] We show for the first time that TERS can be used to distinguish successfully between glycosylated and nonglycosylated proteins from the measurements of just a few molecules within a monolayer.

■ EXPERIMENTAL SECTION

Ribonuclease A and B from bovine pancreas (lyophilized powder), D-mannose, and *N*-acetyl-D-glucosamine were used as supplied from Sigma-Aldrich Ltd. (Dorset, U.K.).

Ultraflat gold nanoplates were prepared according to the method described by Deckert-Gaudig and Deckert.[24] The proteins were dissolved in deionized water at a concentration of 0.25 g L^{-1} (18.2 μM). Glass coverslips with gold nanoplate-coated surfaces were soaked in the protein solution for 19 h to enable the protein molecules to immobilize on the gold. The coverslips were then removed from the solution, rinsed three times with water, and dried under vacuum. The samples were then ready for TERS.

TERS was conducted on an AFM-Raman system comprising of an AFM Nanowizard I (JPK Instruments AG, Germany), mounted on an inverted microscope (Olympus IX70, Japan). Raman was performed at 530.9 nm on a LabRam HR (Horiba Jobin Yvon, France). In depth TERS protocols and data processing are detailed in the Supporting Information.
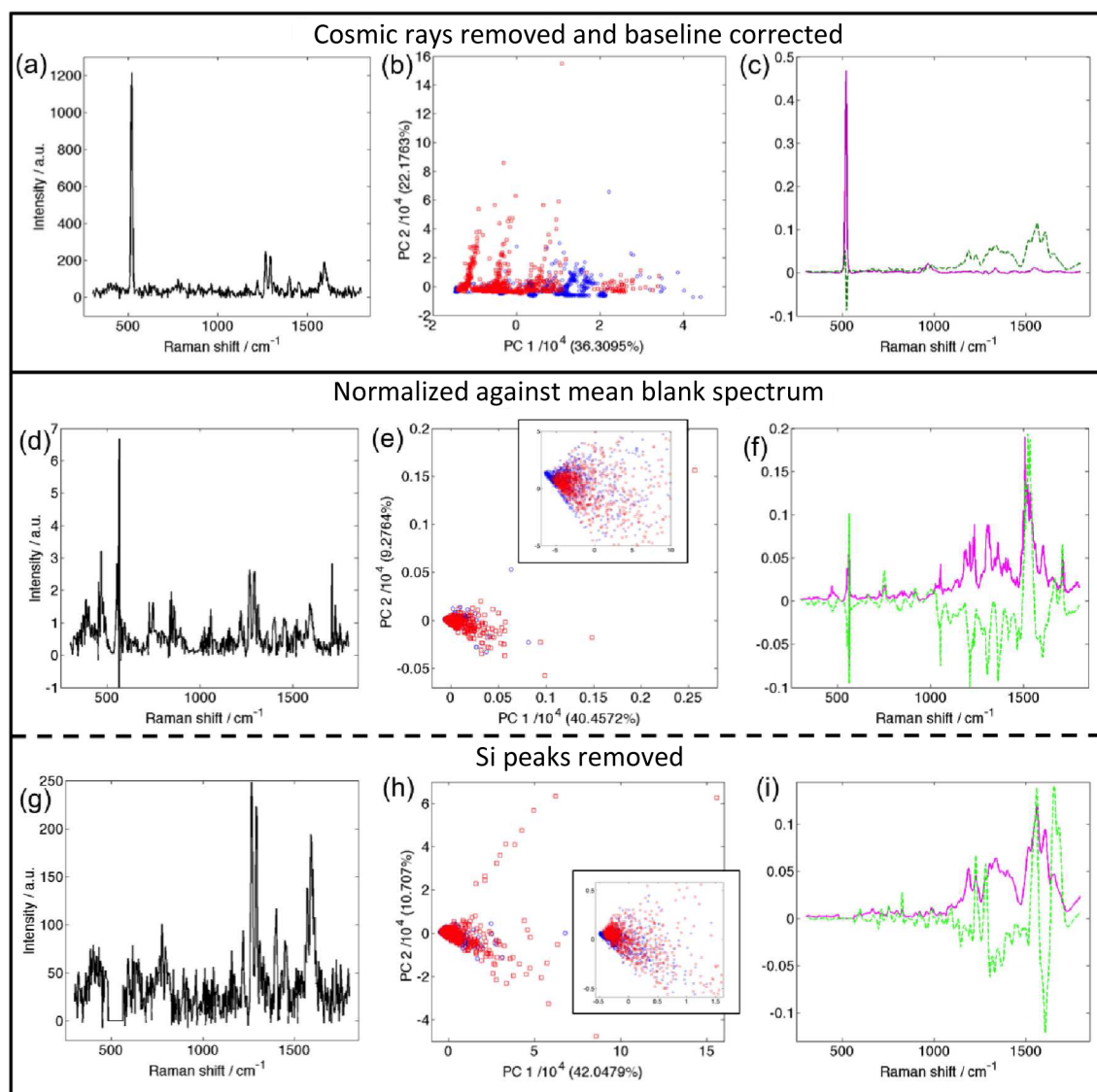
Figure 2. Following cosmic ray removal and baseline correction by the asymmetric least-squares method, (a) a typical RNase A spectrum, (b) PCA scores plots of all of the RNase A (blue circles) and RNase B (red squares) with % explained variance in brackets and (c) PCA loadings of PC1 (solid purple) and PC2 (dashed green). Following the above, (d−f) spectra normalized against the mean blank spectrum and (g−i) after the Si overtone was removed by further asymmetric least-squares baseline correction (see Figure S2) and the region containing the Si fundamental set to zero. Parts d and g are example spectra from the same measurement. Parts e and h are PCA scores plots of all of the RNase A (blue circles) and RNase B (red squares) with % explained variance in parentheses. Insets show a zoomed-in view. Parts f and i are PCA loadings plots of PC1 (solid purple) and PC2 (dashed green).

■ RESULTS AND DISCUSSION

The RNase model proteins chosen for this work were both the native form, RNase A, and the glycosylated form, RNase B, and are available commercially in high purity. More importantly, as we have measured them using Raman microscopy, we also had a solid background model in place for comparison.[14] The inset in Figure 1 shows the structure of the protein and its associated glycan. Raman and TERS spectra of the individual sugar monomers are also provided in Figure 1.

Ultraflat gold nanoplates[24] were chosen to immobilize the RNase proteins for TERS analysis as they are a suitable level for AFM measurement, are transparent to allow TERS measurement through the sample, and will bind effectively to the disulfide bridges in the RNase proteins. An example of the typical gold nanoplate topography is shown in Figure S1. The diameter of the nanoparticle at the end of the TERS tip is

approximately 20 nm, which means each measurement will interrogate just a few protein molecules at a time, assuming an RNase diameter of 3.8 nm.[25,26]

TERS measurements were performed as previously described.[24] At first the TERS tip was scanned through the laser spot and positioned at the location giving the highest reflection in the optical response image. After switching from tip-scanner to sample-scanner mode, the topography was scanned and an appropriate nanoplate was selected. TERS spectra were recorded by way of grids with 1−10 nm point separation. After each grid measurement, a reference spectrum next to the gold plate was acquired to exclude tip contamination. The acquisition time depended on the enhancing ability of the respective tip and was set to 1, 5, or 10 s.

Data preprocessing is often necessary with Raman spectra so that one can extract the useful information from the background signal, especially when enhancement techniques such as SERS and TERS are used to probe analytes that are weak Raman scatterers, such as sugars,[27] very low concentration, or as in this case, both. This was particularly apparent here, as the raw data from the TERS measurements were highly variable. Even between spectra that were from the same grid, for both the sugar monomers and the protein samples, certain points gave no significant TERS response due to lost feedback and among those that did, there was seemingly little consistency from spectrum to spectrum in terms of peak position. Variations in band position and intensity are a common observation in TERS and can be ascribed to the small number of molecules interacting with the tip as a result of the technique's high spatial resolution. The band position is therefore much more sensitive to the specific section of the amino sequence that interacts and the associated variations in orientation, compared to the ensemble averaging observed in SERS measurements.[26,28−30] This variability can be clearly observed in Figure 1a, which shows the raw spectra from just one grid of data, and was also observed when the crystalline glycan monomers were measured (Figure 1c).

The protein sample concentration was kept low to deposit molecules at almost monolayer coverage onto the gold nanoplates; the aim being to get as close as possible to looking at just one protein molecule at a time. We might expect to see more peaks in the TERS spectrum of a sample compared to the Raman due to the relaxation of selection rules,[31] which could result in more allowed vibrational modes. For example, we believe this to be why strong peaks were seen at ∼1575 cm^{-1} in the TERS spectra of crystalline mannose but not in the Raman (Figure 1c,d).

When comparing the protein samples, one of the main regions of interest was between 800 and 1100 cm^{-1}, where peaks from the sugar molecules are expected but not strong peaks from the proteins. Despite this, the variability of peak positions within such a large data set meant it was not immediately possible to identify individual peaks that would be indicative of glycosylation status by visual inspection only. Principal component analysis (PCA) was therefore used to analyze the spectra as a whole in an attempt to extract useful information. In order to minimize the influence of random environmental factors on the PCA, cosmic rays (such as those indicated by asterisks in Figure 1a) were removed and the spectra were baseline corrected to eliminate any baseline drift and minimize fluorescence interference. PCA was then performed on the data and the resultant scores plots are shown in Figure 2b,c. From these PCA scores plots of PC 1 against PC 2 (Figure 2b), there appeared to be some clustering of spectra according to the different proteins, particularly in PC 1; while PC 2 seemed to show some variation within certain subgroups of these clusters. When the loadings were plotted however (Figure 2c), it was apparent that this separation, particularly in PC 1, was mainly due to the Si fundamental peak from the TERS tip at 520 cm^{-1} and not from any Raman features associated with the protein. In fact, PCA was able to group the samples from each respective measurement grid quite well based mainly on the features associated with the Si tip, revealing the subtle variation between measurement grids as a result of the background signal, which would need to be accounted for to gain reliable results. These data are shown in the Supporting Information (Figure S2).

An example of the background spectrum of the gold nanoplates is shown in Figure 1b, and it is clearly dominated by fundamental Si−Si stretch from the silicon tip. The Si fundamental peak was so intense that it masked any protein or carbohydrate peaks in the same region. The Si first overtone peak at ∼960 cm^{-1} appeared as a small but broad increase in the baseline and is unfortunately located in the information-rich part of the spectrum where we might expect to see differences between the protein samples, with sharper sample peaks often visible on top. This made the removal of the Si overtone somewhat more difficult without adversely affecting the useful peaks in the same spectral region. It was therefore important to remove this overtone accurately, and the three different approaches detailed below were sequentially evolved to reduce the influence of the Si peaks on the PCA model.

The first method involved fitting Gaussians under both the Si fundamental and first overtone peaks of each spectrum. The center position, width, and height of each were recorded, and the ratio of each parameter between the fundamental and overtone features were calculated. The median value of the ratios for each parameter was then used to plot a predicted first overtone peak based only on the properties of the measured fundamental peak, so as to remove the influence of any other peaks that may occupy the same region of the spectrum as the overtone. The median was chosen over the mean to negate the influence of a small number of extreme outliers. The fitted fundamental Gaussian and predicted overtone Gaussian were then subtracted from the relevant part of each spectrum.

The predicted overtone peaks fitted the general shape of the baseline very well (data not shown) but as there were overlapping features either side of this background peak, it was difficult to match up the predicted Gaussian region accurately with the continuing spectrum, leading to the insertion of false features. As such, the Gaussian-fitting method was not considered reliable.

The second method used to remove the Si features was to normalize each spectrum against the mean of the blank spectra; that is to say, the spectra of the gold nanoplates without protein sample present. The idea was to be able to remove the background peaks in a way that did not destroy any analytically useful protein information in their immediate vicinity. A typical spectrum before and after this correction procedure is shown in Figure 2a,d, where it can be seen that the intensity of a number of features in the spectrum have been emphasized as a result of this processing. The PCA plot (Figure 2e) is now completely different to that generated previously (Figure 2b), although the loadings plot (Figure 2f) still shows some influence from the Si fundamental peak. This is likely caused by the position of the same peak in the blank spectrum not matching exactly that of the sample, as evident in Figure 2d; however, this Si fundamental vibration is no longer the dominant feature in PC 1 and other peaks at higher wavenumber are having greater influence.

The final method was intended to remove the influence of the Si fundamental from the PCA model entirely while maintaining features in the region of the Si overtone. As such, the region containing the Si fundamental peak was set to zero and while there was the potential to also remove other peaks in the same region, the complete removal of the Si peak was considered more beneficial. An asymmetric-least-squares (ALS) baseline correction was applied to the region containing the overtone peak with the parameters set to fit the baseline very closely to the original data. Because of the broad nature of

the overtone peak, the fitted baseline follows its shape reliably but does not fit too closely to any sharper features that appear on top. An example of this baseline fitting is shown in Figure S3.

The example spectrum used previously is shown following this data processing, along with the subsequent PCA scores and loadings plots in Figure 2g−i, respectively. Features of the spectra are not emphasized as strongly as when they are normalized against the blank but the core cluster of the PCA scores plot shows a similar shape and the corresponding loadings show that very similar parts of the spectrum are influential; with the exception of the Si fundamental which has of course been removed.

Despite the apparent success in minimizing the influence of the silicon background, PCA had still not conclusively identified regions of the spectra where differences were observable between the two forms of the protein; as evidenced by the lack of separation of RNase A (blue circles) and B (red squares) in Figure 2e,h. One other area that had not been considered thus far was the enhancement factor of the TERS measurements. Depending on the proximity and position of the TERS tip relative to the protein molecules whose spectra were being observed, the magnitude of the TERS enhancement will have been different for each measurement. This was very apparent from the individual spectra, which differed greatly in terms of maximum intensity and this may have inadvertently influenced the PCA scores plots in Figure 2. Thus, after the removal of the Si features in the TERS spectra, data were normalized to their own total signal intensity (each spectrum was row-normalized to 1 of its total signal).

The PCA scores and loadings following this row normalization are shown in Figure S4a,b for the blank-normalized data and in Figure S4c,d for the ALS-corrected data. The fact that the PCA scores plots are now roughly circular in shape with no obvious clustering indicates that this normalizing step has been successful. However, with this bias removed, PCA did not show any separation between the two different forms of RNase protein, even when higher PCs were investigated (data not shown). The fact that the explained variance for PC1 is <10% indicates that the spectra are too variable for PCA to be useful, with no similarities between spectra, even from features unrelated to the glycosylation status.

With no obvious markers (Raman features/bands) identifiable by eye or through the use of multivariate PCA, despite strategic processing of the data, a more general approach was taken. In the previous work by Brewster et al. using conventional Raman spectroscopy,[14] after multivariate data processing using multivariate supervised learning (partial least-squares), six regions of the Raman spectra were identified that were significant in separating the two RNase types. Therefore, the area under the curve (AUC) of these same regions was measured in the preprocessed TERS spectra (in this example using the normalization against the blank method for Si peak removal), and the results represented in a box and whisker plot in Figure 3. In general, the glycosylated form of the protein has a higher AUC for the regions 850−900 cm^{-1}, 1220−1300 cm^{-1}, 1420−1490 cm^{-1}, and 1700−1800 cm^{-1} and a lower AUC for 780−820 cm^{-1} and 950−1000 cm^{-1}. Assignments for these regions are shown in the Table S1, where it is clear that the TERS was in agreement with protein conformational changes rather than the detection of the sugar *per se*.

Figure S5 shows box and whisker plots of the AUC for other regions of the TERS spectra that are not expected to contain
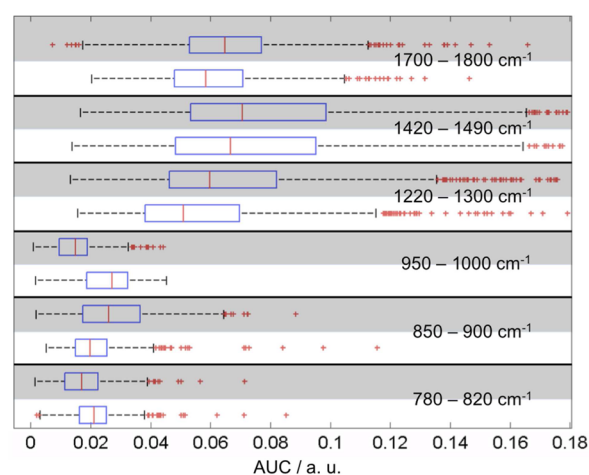

Figure 3. Box and whisker plot of the area under the preprocessed TERS spectra in the wavenumber regions indicated that showed clear differences between the two protein glycoforms using normal Raman spectroscopy. For each region, the shaded box and whisker represents RNase B. In each case the blue box represents the interquartile range (IQR), the red line the median, and the black dashed whiskers the remaining data, with the exception of any data that lie outside the IQR by more than 1.5 × IQR (±2.7σ), which are shown by red crosses. Vibration assignments for the regions are given in Table S1. In this example, the data were preprocessed using the normalization against blank Si.

any glycan specific information.[14] With the exception of the 300−780 cm^{-1} region that includes any artifacts from the removal of the Si fundamental peak, the AUCs show little difference between the two forms of the protein (the two medians are very close as are the IQRs), indicating that we may indeed be able to classify the proteins based on features within the "sugar regions" of the spectra.

These six sugar regions (Figure 3, Table S1) were isolated and PCA performed on each region individually. As illustrated for the 950−1000 cm^{-1} and 1700−1800 cm^{-1} regions, when all the spectra are used (Figure 4a,d, respectively), there appears to be some separation between RNase A (blue circles) and B (red squares). Furthermore, when the spectra that fell outside the interquartile range of the box and whisker plot in Figure 3 were removed (the less typical spectra), the clustering is improved dramatically, with a complete separation between the two protein types observed for the 950−1000 cm^{-1} region (Figure 4b,e). Typical representative spectra, selected from the centers of the clusters, are shown in Figure 4c,f and the features responsible for the separation, tentatively assigned to β-sheet/disordered secondary structure/phenylalanine and side-chain carbonyl vibrations, respectively, (vide infra) are also highlighted in yellow. In our previous experiments using Raman spectroscopy of RNase A and RNase B, we also observed disordered protein structure as there were differences in Amide III vibrations that could be assigned to the disordered loops of the RNase protein in proximity to Asn34 (the glycosylation site) as well as differences in the Amide I which could be due to β-structure or tertiary structure changes on protein glycosylation.[14]

Extra peaks attributed to the glycan in the spectra of RNase B were not definitively observed, a possible explanation for which might be the weak scattering nature of the mannose and N-acetylglucosamine (GlcNAc) residues. From the TERS spectra of neat mannose and GlcNAc, it is evident that even in
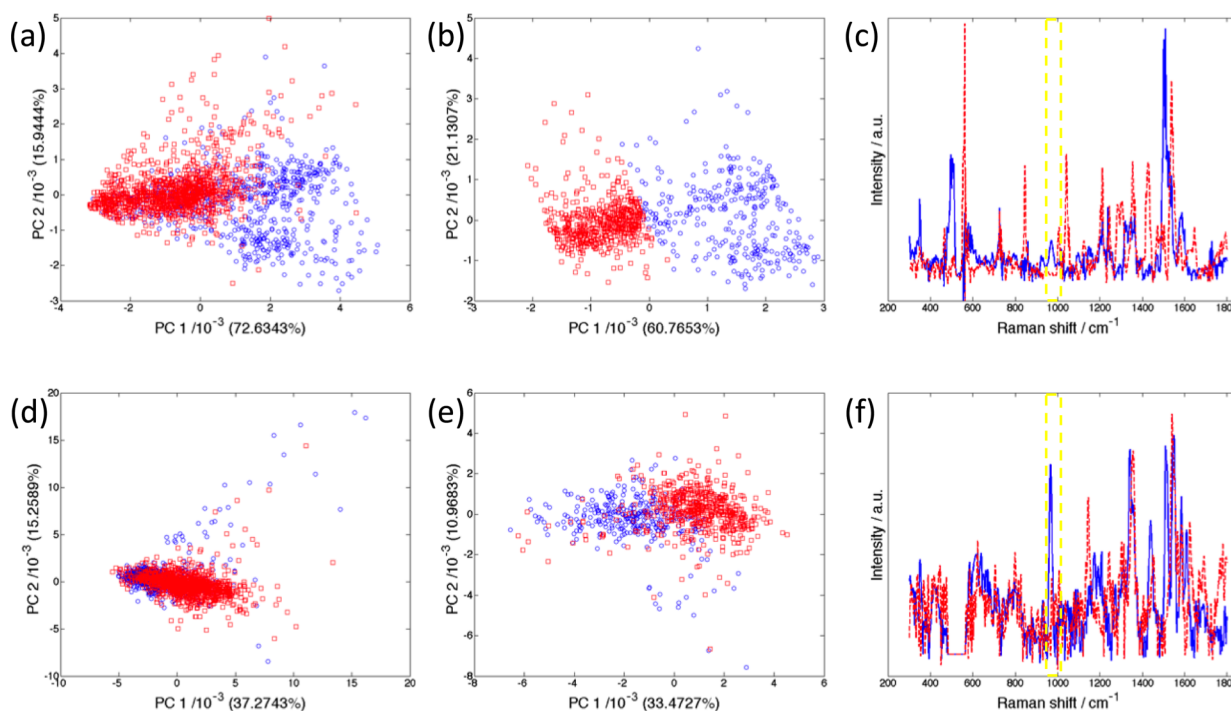
Figure 4. Following all of the preprocessing (a) shows PCA scores of the area under the curve of 950−1000 cm^{-1} and (b) shows PCA scores of the interquartile range of the same data, with RNase A shown by blue circles, RNase B shown by red squares, and % explained variance in parentheses. Part c shows representative spectra from each group in part b with RNase A in solid blue, RNase B in dashed red, and the selected region highlighted in yellow. Parts d−f show the same information as parts a−c, but following the ALS Si-removal preprocessing route and the area under the curve of 1700−1800 cm^{-1}. Parts a and d comprise 643 RNase A spectra and 1116 RNase B spectra, while after selection of the spectra within the IQR (b, e) 321 and 558 spectra are retained for the RNase A and B, respectively.

crystalline form these molecules do not always show highly enhanced signals in the characteristic carbohydrate region (900−1100 cm^{-1}). Instead, it is thought that the distinction we observe results from differences in the way the two forms of the protein adsorb onto the gold substrate. More specifically that the native form of the protein (RNase A) can adsorb in many different orientations, but the presence of the glycan in RNase B somewhat restricts the number of possibilities. We have used AFM to scan the topology of the gold surface that has proteins deposited on its surfaces, and Figure S6 indicates that there are features of approximately 3 nm in height which would correspond to the diameter of RNase A and RNase B proteins. We have made previous measurement from blank gold nanoplates and these show a roughness below 1 nm (Figure 3 in ref 24). This indicates that our TERS measurements are likely to be generated from single protein molecules. We have not been able to characterize the structure of the protein monolayer on the gold, but there are four disulfide bridges within the molecule that are able to adsorb onto the surface[32] as well as numerous side-chain functional groups. It is known that sugar molecules do not readily interact with gold surfaces[33] and therefore both sterically and chemically, the glycan is highly likely to direct adsorption on the gold substrate so that the glycan itself is preferentially on top of the protein surface (exposed to the atmosphere), thus giving it the potential to mask at least some of the protein vibrations.

The major distinguishing feature is an extra peak in the spectrum of RNase A in the 950−1000 cm^{-1} region. In this region one might expect to observe peaks associated with the protein β-sheet,[34] disordered secondary structure,[35] and possibly shifted phenylalanine ring-breathing.[36] The lack of this feature in the glycosylated protein spectra could be due to

masking of the weakly scattering glycan if it is attributed to a β-sheet or phenylalanine vibration, or alternatively it is possible that due to its relatively unhindered side-chain functionality, the native form of the protein undergoes partial unfolding upon adsorption on the gold substrate, leading to the appearance of an additional peak attributed to disordered secondary structure. Unfortunately we have no data to confirm either hypothesis at this stage.

Also as a consequence of fewer carbonyl groups from the amino acid side-chains of RNase B being able to interact with the gold substrate (as they are facing the wrong way) compared to RNase A, they are more able to be "seen" by the TERS tip. This provides a tentative explanation for the increased signal intensity in the 1700−1800 cm^{-1} region for RNase B compared to RNase A (Figure 4d−f). While the TERS intensity in this region is generally higher for RNase B than RNase A, peaks are still observed in some RNase A spectra in this region, as would be expected if multiple orientations were possible: in some cases the RNase A will have adsorbed in the preferred orientation for RNase B purely by chance. This would also explain why there is some overlap in the PCA plot (Figure 4e).

The crystal structure of RNase B has previously been characterized[37] and interactions between the glycan and the amino acid side-chains were not detected, thus we do not cite this as a possible explanation for the spectral differences observed.

It was clear that like conventional Raman spectroscopy, TERS does contain enough information to allow the differentiation between two proteins that differ in whether they are glycosylated or not. The data processing steps required to obtain this useful information may be quite involved; however, care has been taken to eliminate bias with the use of exclusively

unsupervised multivariate methods, meaning at no point were the analyses developed based on prior knowledge of the sample groups to which the TERS data belonged to.

The aim of the experiment had been to detect directly the glycans on the glycosylated protein molecules in order to look for differences in the glycans themselves, but this does not appear to have been achieved here. While it is possible that this information is present, the high level of variability in peak positions makes it extremely difficult to make a definitive assignment of spectral features to the sugars in the glycans. The TERS spectra of the crystalline mannose and GlcNAc show that sugar bands in the 800−1100 cm^{-1} region are only detected if the bands in the 1400−1600 cm^{-1} region are very strongly enhanced and also that the spectra of the individual sugar monomers are highly dependent on orientation with respect to the tip.

We do believe this research is a valuable first step toward demonstrating the potential of TERS to characterize glycosylated proteins, given that this is the first time that anyone has been able to use TERS to distinguish between glycosylated and native forms of a protein without any control of protein−surface interaction. The next step would be to apply orientation control with the design of functionalized substrates that would force the proteins to orientate themselves on the surface in a reproducible way, thereby minimizing the variation in peak position from similar vibrations. With this level of control in place, it would become more likely that structural changes as a result of incorrect glycosylation would be detectable.

CONCLUSION

We have demonstrated for the first time that TERS can be used to distinguish between glycosylated and native forms of proteins. This is the first step toward the goal of being able to characterize correct glycosylation of just a few protein molecules at a time using this powerful high spatial resolution technique. We have shown that although TERS data of protein samples can be very complex, carefully considered data processing steps can reveal useful information. While TERS does not compete with the high-throughput potential of conventional Raman for manufacturing process monitoring and quality control, after careful design of surface substrates to control sample orientation, TERS has the potential for use in glycoprotein synthesis and formulation research. This will be an area of future work as will the utility of TERS to assess glycosylation status on single protein molecules as this may enable the assessment of the same protein with variable glycosylation status (i.e., mixtures of different sugars in the glycan on protein; for example, RNase B has a variable number of mannoses as illustrated in Figure 1) which are typical of manufacture.

ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.analchem.5b03535.

Full details of instrumentation used and TERS measurement, full data analysis procedures, tentative Raman band assignments (Table S1), as well as five figures detailing AFM topology of gold nanoplate (Figures S1 and S6) and additional PCA and data analyses (Figures S2−S5) (PDF)

AUTHOR INFORMATION

Corresponding Authors
*E-mail: roy.goodacre@manchester.ac.uk.
*E-mail: volker.deckert@ipht-jena.de.

Notes
The authors declare no competing financial interest.

REFERENCES

(1) Apweiler, R.; Hermjakob, H.; Sharon, N. *Biochim. Biophys. Acta, Gen. Subj.* **1999**, *1473*, 4−8.
(2) Hamrang, Z.; Rattray, N. J. W.; Pluen, A. *Trends Biotechnol.* **2013**, *31*, 448−458.
(3) Walsh, G.; Jefferis, R. *Nat. Biotechnol.* **2006**, *24*, 1241−1252.
(4) Ellis, D. I.; Cowcher, D. P.; Ashton, L.; O'Hagan, S.; Goodacre, R. *Analyst* **2013**, *138*, 3871−3884.
(5) Wen, Z.-Q. *J. Pharm. Sci.* **2007**, *96*, 2861−2878.
(6) Marison, I.; Hennessy, S.; Foley, R.; Schuler, M.; Sivaprakasam, S.; Freeland, B. The Choice of Suitable Online Analytical Techniques and Data Processing for Monitoring of Bioprocesses. In *Measurement, Monitoring, Modelling and Control of Bioprocesses*; Mandenius, C.-F., Titchener-Hooker, N. J., Eds.; Springer: Berlin Heidelberg, Germany, 2013; Vol. *132*, pp 249−280.
(7) Cui, Y.; Turner, G.; Roy, U. N.; Guo, M.; Pan, Z.; Morgan, S.; Burger, A.; Yeh, Y. *J. Raman Spectrosc.* **2005**, *36*, 1113−1117.
(8) Kopecky, V.; Ettrich, R.; Hofbauerova, K.; Baumruk, V. *Biochem. Biophys. Res. Commun.* **2003**, *300*, 41−46.
(9) Bell, A. F.; Ford, S. J.; Hecht, L.; Wilson, G.; Barron, L. D. *Int. J. Biol. Macromol.* **1994**, *16*, 277−278.
(10) Zhu, F.; Isaacs, N. W.; Hecht, L.; Barron, L. D. *J. Am. Chem. Soc.* **2005**, *127*, 6142−6143.
(11) Johannessen, C.; Pendrill, R.; Widmalm, G.; Hecht, L.; Barron, L. D. *Angew. Chem., Int. Ed.* **2011**, *50*, 5349−5351.
(12) Ashton, L.; Pudney, P. D. A.; Blanch, E. W.; Yakubov, G. E. *Adv. Colloid Interface Sci.* **2013**, *199−200*, 66−77.
(13) Kikuchi, G. E.; Baker, S. A.; Merajver, S. D.; Coligan, J. E.; Levine, M.; Glorioso, J. C.; Nairn, R. *Biochemistry* **1987**, *26*, 424−431.
(14) Brewster, V. L.; Ashton, L.; Goodacre, R. *Anal. Chem.* **2011**, *83*, 6074−6081.
(15) Barman, I.; Dingari, N. C.; Kang, J. W.; Horowitz, G. L.; Dasari, R. R.; Feld, M. S. *Anal. Chem.* **2012**, *84*, 2474−2482.
(16) Lazar, I. M.; Lee, W.; Lazar, A. C. *Electrophoresis* **2013**, *34*, 113−125.
(17) Moskovits, M. *J. Chem. Phys.* **1978**, *69*, 4159−4161.
(18) Smith, W. E. *Chem. Soc. Rev.* **2008**, *37*, 955−964.
(19) Bailo, E.; Deckert, V. *Chem. Soc. Rev.* **2008**, *37*, 921−930.
(20) Neugebauer, U.; Rösch, P.; Schmitt, M.; Popp, J.; Julien, C.; Rasmussen, A.; Budich, C.; Deckert, V. *ChemPhysChem* **2006**, *7*, 1428−1430.
(21) Neugebauer, U.; Schmid, U.; Baumann, K.; Ziebuhr, W.; Kozitskaya, S.; Deckert, V.; Schmitt, M.; Popp, J. *ChemPhysChem* **2007**, *8*, 124−137.

(22) Bailo, E.; Deckert, V. *Angew. Chem., Int. Ed.* **2008**, *47*, 1658−1661.

(23) Treffer, R.; Bohme, R.; Deckert-Gaudig, T.; Lau, K.; Tiede, S.; Lin, X. M.; Deckert, V. *Biochem. Soc. Trans.* **2012**, *40*, 609−614.

(24) Deckert-Gaudig, T.; Deckert, V. *Small* **2009**, *5*, 432−436.

(25) Ramm, L. E.; Whitlow, M. B.; Mayer, M. M. *J. Immunol.* **1985**, *134*, 2594−2599.

(26) Deckert-Gaudig, T.; Kämmer, E.; Deckert, V. *J. Biophotonics* **2012**, *5*, 215−219.

(27) Walton, A. G.; Blackwell, J. *Biopolymers*; Academic Press: London, 1973.

(28) Ichimura, T.; Watanabe, H.; Morita, Y.; Verma, P.; Kawata, S.; Inouye, Y. *J. Phys. Chem. C* **2007**, *111*, 9460−9464.

(29) Treffer, R.; Lin, X.; Bailo, E.; Deckert-Gaudig, T.; Deckert, V. *Beilstein J. Nanotechnol.* **2011**, *2*, 628−637.

(30) Kurouski, D.; Deckert-Gaudig, T.; Deckert, V.; Lednev, I. K. *J. Am. Chem. Soc.* **2012**, *134*, 13323−13329.

(31) Berweger, S.; Raschke, M. B. *J. Raman Spectrosc.* **2009**, *40*, 1413−1419.

(32) Di Felice, R.; Selloni, A.; Molinari, E. *J. Phys. Chem. B* **2003**, *107*, 1151−1156.

(33) Stuart, D. A.; Yonzon, C. R.; Zhang, X.; Lyandres, O.; Shah, N. C.; Glucksberg, M. R.; Walsh, J. T.; Van Duyne, R. P. *Anal. Chem.* **2005**, *77*, 4013−4019.

(34) Howell, N.; Li-Chan, E. *Int. J. Food Sci. Technol.* **1996**, *31*, 439−451.

(35) Ashton, L.; Barron, L. D.; Hecht, L.; Hyde, J.; Blanch, E. W. *Analyst* **2007**, *132*, 468−479.

(36) Tuma, R. *J. Raman Spectrosc.* **2005**, *36*, 307−319.

(37) Williams, R. L.; Greene, S. M.; McPherson, A. *J. Biol. Chem.* **1987**, *262*, 16020−16031.