Multivariate Analysis of 2D $^{1}H,^{13}C$ methyl NMR Spectra of Monoclonal Antibody Therapeutics to Facilitate Assessment of Higher Order Structure

Luke W. Arbogast, Frank Delaglio, John E. Schiel and John P. Marino*

Institute for Bioscience and Biotechnology Research, National Institute of Standards and Technology and the University of Maryland, 9600 Gudelsky Dr., Rockville, Maryland 20850, United States

Corresponding Author

*E-mail john.marino@nist.gov

ABSTRACT

Two-dimensional (2D) ^{1}H-^{13}C methyl NMR provides a powerful tool to probe the higher order structure (HOS) of monoclonal antibodies (mAbs), since spectra can readily be acquired on intact mAbs at natural isotopic abundance, and small changes in chemical environment and structure give rise to observable changes in corresponding spectra, which can be interpreted at atomic resolution. This makes it possible to apply 2D NMR spectral fingerprinting approaches directly to drug products in order to systematically characterize structure and excipient effects. Systematic collections of NMR spectra are often analyzed in terms of the changes in specifically identified peak positions, as well as changes in peak height and linewidths. A complementary approach is to apply Principal Component Analysis (PCA) directly to the matrix of spectral data, correlating spectra according to similarities and differences in their overall shapes, rather than according to parameters of individually identified peaks. This is particularly well-suited for spectra of mAbs, where some of the individual peaks might not be well resolved. Here we demonstrate the performance of the PCA method for discriminating structural variation among systematic sets of 2D NMR fingerprint spectra using the NISTmAb and illustrate how spectral variability identified by PCA may be correlated to structure.

INTRODUCTION

Therapeutic monoclonal antibodies (mAbs) provide a platform for life-changing and life-saving clinical medicines, however their development and manufacture presents many analytical challenges. In particular, characterization of mAbs requires assessing their higher order structure (HOS), since misfolding or aggregation can lead to loss of efficacy or cause unintended and potentially life-threatening immune responses. Nuclear Magnetic Resonance (NMR) spectroscopy is being increasingly recognized as a potent tool for the HOS assessment of protein therapeutics, including mAbs.[1–6] Indeed, we have recently demonstrated the feasibility of recording two-dimensional (2D) heteronuclear ^{1}H-^{15}N and ^{1}H-^{13}C correlation spectra on mAbs,[7–10] however, questions remain as to best practices for spectral analysis in connection with determination of structural comparability of HOS.[5]

Conventionally, interpretation of multi-dimensional protein NMR data has required expert analysis of individual spectra, which involves identification of specific changes in cross peak characteristics such as position, height and linewidth. However, the utility of this approach is limited when large numbers of datasets must be analyzed as well as if individual peaks are not all fully resolved, as is the case for mAbs.[11] Fortunately, practical solutions to these limitations can be found in one-dimensional (1D) ^{1}H NMR methods, such as in application to complex mixtures, where statistical based approaches for analysis of spectra have been well established.[12] Indeed a number of 1D ^{1}H NMR based methods have been recently proposed for chemometric HOS analysis of mAbs.[3,4,13–15] The application of chemometric approaches to the problem of 2D spectral analysis of protein therapeutics is less well-established, but is actively being explored. Proposed approaches include direct point-to-point linear correlation analysis of the processed data matrix,[16] spectral covariance analysis,[17] multivariate principal component analysis on both the processed data matrix, [18,19] as well as on tables of peak positions and intensities,[5] spectral distance metrics such as nearest neighbor approaches on peak tables,[20] and total spectral Euclidian deviation.[5]

Previously, we have demonstrated that 2D ^{1}H-^{15}N and ^{1}H-^{13}C spectra of different mAb species can be analyzed and discriminated from one another using linear correlation analysis.[9,10] However, we also demonstrated the susceptibility of this method to experimental noise and resolution, and results from the study suggested a high limit of detection (LOD) for discrimination between similar structural species, thus bringing into question the suitability of the method to therapeutic mAb spectral analysis. Here we present a multivariate principal component analysis (PCA) method applied directly to the matrix of spectral intensities, which in comparison, we demonstrate can readily discriminate mAb species that manifest both large and subtle spectral differences relative the reference mAb spectrum which arise from enzymatic deglycosylation or

degalactosylation. Furthermore we demonstrate how PCA may be used to readily visualize spectral differences and map those differences back on to the protein structure.

EXPERIMENTAL SECTION

Certain commercial equipment, instruments, and materials are identified in this paper to specify the experimental procedure. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the material or equipment identified is necessarily the best available for the purpose.

Preparation of Samples.

The native NISTmAb sample, a humanized IgG1κ expressed in murine suspension culture, was prepared from the bulk primary sample PS-8670, formulated at 100 mg/mL, L-histidine, pH 6.0 by dilution to 30 mg/mL and buffer exchange into 25 mM Bis-Tris-d19, pH 6.0 using a 100 kDa centrifugal filter. The exo-degalactosylated (ExoGal) NISTmAb sample was prepared as described by Yamaguchi *et. al*.[21] Briefly, 300 μL of the bulk material was diluted to 1 mL in 50 mM sodium acetate, pH 5.5 with 5 mM $CaCl_2$ and incubated with 80 μL of β (1-4) galactosidase (8 kU/mL, New England Biolabs) for 72 h at room temperature. The reaction was quenched with concomitant purification of the mAb by removal of the β(1-4) galactosidase using Protein A affinity chromatography. After elution of the antibody from the Protein A column, it was concentrated to 30 mg/mL and exchanged into 25 mM Bis-Tris-d19, pH 6.0 using a 100 kDa centrifugal filter. The extent of the reaction was verified by intact LC-MS analysis (Figure S2). The endo-deglycosylated (DeGly) NISTmAb was similarly prepared from 320 μL of the bulk material, diluted to 2 mL in 25 mM sodium phosphate, pH 6.0 and incubation with 25 μL PNGase F (500 kU/mL, New England Biolabs) for 70 h at room temperature. The reaction was quenched with concomitant purification of the mAb by removal of the PNGase F using Protein A affinity chromatography. After elution of the antibody from the Protein A column, it was concentrated to 30 mg/ml and exchanged into 25 mM Bis-Tris-d19, pH 6.0 using a 100 kDa centrifugal filter. The extent of the reaction was verified by intact LC-MS (Figure S2). Blended samples of native NISTmAb and ExoGal NISTmAb were created by serial dilution of the ExoGal NISTmAb with native NISTmAb such that total protein concentration was kept at ca. 40 mg/mL. NMR samples were prepared from 300 μL aliquots of the mAb samples with addition of 3 % D2O and were then loaded into Bruker Biospin shaped tubes for data acquisition.

LC-MS Analysis.

LC-MS analysis was performed on an Agilent Infinity II UHPLC coupled with an Agilent 6545 QTOF mass spectrometer. Intact antibody was injected (3 μg) onto a PLRP-S Column (1 mm x 50 mm, 1000 Angstrom pore size, 5 um particle size) and eluted at 0.4 mL/min according to the gradient in Table S1. Solvent A consisted of 0.1 % Formic acid in Water and Solvent B consisted of 0.1 % Formic acid in Acetonitrile. The column was held at a constant temperature of 60 °C. The QTOF was operated in 2 GHz Extended Mass Range (500 m/z to 5000 m/z) mode at an acquisition rate of 1 spectra / sec. The 1221.990637 ion was used as a reference mass throughout the run. Deconvolution of the resulting spectrum was performed using BioConfirm 8.0 using the maximum entropy algorithm.

CE-SDS Analysis.

CE-SDS under non-reducing conditions was performed according to the PA800Plus IgG Purity/Heterogeneity Assay Kit instructions. Bare fused silica capillaries (50 μm inner diameter, 67 cm total length) were purchased from Sciex Separations (PN 338451) and cut to 30.5 cm prior to use. SDS-MW gel buffer, acidic and basic wash solutions, molecular weight ladder, 10 kDa internal standard, and Tris/SDS sample buffer (pH 9.0) were purchased as a kit from Sciex Separations (PN 390953). Briefly, 100 μg of mAb in 10 μL was further diluted with 85 μL of the SDS sample buffer (100 mM Tris-HCl, pH 9.0, 1 %SDS) in Protein LoBind tubes (Eppendorf, PN

89166-278 for 0.5 mL). 2 µL of the 10 kDa internal standard and 5 µL of 0.25 M iodoacetamide in water, prepared fresh, was added to the sample solution. The sample was mixed by vortexing, then incubated for 10 min at 70 °C in a water bath (or as otherwise indicated). The sample was cooled to room temperature, vortexed, centrifuged briefly, and then transferred to sample vials for CE-SDS analysis. All samples were analyzed using a Sciex Separations PA800 plus pharmaceutical analysis system according to the manufacturers IgG Purity Analysis settings. Samples were detected at the capillary window, 20 cm from the inlet, using the PA800 plus PDA detector set to collect absorbance at 220 nm.

NMR Experiments.

NMR experiments were conducted on either a Bruker Avance III 600 MHz or 900 MHz spectrometer equipped with triple resonance cryogenically cooled TCI probe and triple axis gradient system at 50 °C. All data were recorded using a gradient-selected ^{1}H-^{13}C HSQC experiment (gsHSQC) with a recycle delay of 1.5 s. The ^{1}H carrier was set to the water resonance and the ^{13}C carrier to 21.25 ppm. Data were collected with 16, 32, 64 or 128 scans with 60x1009 complex points in the F1 and F2 domains with spectral widths of 5998.27 Hz and 12626.26 Hz, corresponding to acquisition times of 10 ms and 80 ms, unless otherwise noted. Maximum experimental durations were 6 h 55 m for data collected with 128 scans Data were processed in NMRPipe,[22] with using cosine-squared apodization in the ^{1}H dimension, cosine apodization in the ^{13}C dimension, and zero-filling in both dimensions.

PCA Analysis.

Prior to analysis by PCA, spectra were scaled to uniform maximum intensity, and used without further scaling or centering. When applying PCA to collections of spectra with differing spectral windows or digital resolutions, spectra were resampled using bilinear interpolation so that each spectral matrix had the same size and chemical shift ranges as the first spectrum in the series, and no binning was used. PCA was performed using the NIPALS algorithm[23] implemented in NMRPipe. In each case, only the first six components were extracted, and saved in the same format used for spectral data. In the examples presented, the first component alone commonly accounted for more than 90 % of the variance in the spectral series. PCA was performed over the methyl region window from -0.6 ppm – 2.4 ppm in ^{1}H and 8 pm – 25 ppm in ^{13}C. Principal component scores for each dataset were imported into MATLAB (R2015b) for cluster analysis by the K-means algorithm on the first three principal components for the respective data series. Analysis of variance in higher principal components was revealed to be dominated by experimental noise and small measurement imprecision attributed to changes in field homogeneity from sample volume shifts during acquisition as a result of high temperature data collection, and were therefore discarded from further analysis. Cluster intervals were computed as the 95 % confidence interval over the spread of the data points in each cluster, or as twice the standard deviation of the spread of the data point in each cluster, as noted in the text. Experimental Signal-to-noise (S/N) was calculated from the root-mean- square (RMS) of the spectral signal matrix and RMS of the spectral noise matrix as previously described.[9]

RESULTS AND DISCUSSION

To investigate the strengths and limitations of chemometric approaches to the analysis of therapeutic mAb 2D NMR spectra, two variants of the NISTmAb were prepared by enzymatic glycan remodeling. The NISTmAb contains three principal glycoforms present in relative abundance greater than 10 %; G0F, G1F, and G2F (Figure S1).[24] Terminal exo-degalactosylation of the heterogeneous NISTmAb glycans was accomplished by reaction with β (1-4) galactosidase and intact LC- MS analysis of the reaction product confirmed a glycan distribution predominated by the G0F glycoform (Figure S2, herein ExoGal-NISTmAb). This modification scheme was chosen with the intent to minimize any modifications to the mAb other than glycan trimming and does not result in the presence of any new glycoforms; rather, it is a redistribution of the relative populations of the glycoforms present in the native state. This is expected to only modestly perturb the ensemble structure of the

NISTmAb, though enough to result in a slightly diminished affinity to the FcγRIII receptor.[25–27] Full removal of the NISTmAb glycan, with accompanying conversion of residue asparagine-297 to aspartate, was accomplished by reaction with PNGase F. This resulted in a near homogenous deglycosylated state (Figure S2, herein DeGly-NISTmAb). Unlike the modest changes expected in the ExoGal isoform, these changes are expected to result in significant structural and dynamic alterations of the mAb.[21,27]

To further evaluate whether extended incubation during the galactosidase reactions resulted in other variant populations in the samples with respect to size heterogeneity, CE-SDS analysis was performed on the ExoGal and DeGly samples. Figure S3 and Table S2 show non-reduced electropherograms and calculated relative abundance for the individual species identified. The ExoGal samples showed no fragmentation beyond that previously observed for the native NISTmAb using CE-SDS, indicating the sample preparation did not produce any additional levels of fragmentation. The DeGly sample did show a moderate increase in antibody fragments, with a corresponding decrease of monomeric purity to ca. 85 %. Together, this set of well-defined samples provides examples of both modest and large-scale changes to the NISTmAb structure that can serve as benchmarks for testing chemometric approaches to analysis of mAb NMR spectra.

Following preparation of the isoform samples, ^{1}H-^{13}C-methyl correlation spectra were acquired on the ExoGal- and DeGly-NISTmAb and compared to the corresponding native NISTmAb spectrum (Figure 1 a,c). The spectrum of the DeGly-NISTmAb, while still quite similar to that of the native NISTmAb, contains several peak shifts and intensity changes that can be seen by direct visual comparison of the spectral overlay and corresponding deviations are manifest in the linear correlation analysis (Figure 1 b). This result is consistent with increased conformational dynamics as a result of deglycosylation as well as increased conformational heterogeneity from the presence of heavy/light chain fragment variants. The ExoGal-NISTmAb and native NISTmAb spectra on the other hand, overlay very well with one another and likewise are nearly indistinguishable from one another within the experimental signal-to-noise (S/N) by linear correlation analysis (Figure 1 d). Given the limitations of the linear correlation approach to differentiate the subtle but functionally distinct ExoGal isoform from the native state structures, we endeavored to investigate potentially more discerning chemometric approaches.

The use of PCA for analysis of NMR spectral datasets is well established, and presents many advantages for mAb HOS comparison by 2D NMR. For example, PCA can take highly complex multivariate data and reduce it down to the variables of highest interest; those with maximum variation in the analyzed dataset. This may be very useful for mAb ^{1}H-^{13}C methyl spectra, which are defined by 200,000+ digital points, a multitude of which do not represent signal. Furthermore, it is expected that mAb spectra will be highly similar, as exemplified in the comparison of the native NISTmAb and strongly perturbed DeGly-NISTmAb, and contain a large percentage of signals that do not vary across the analyte datasets. Therefore, the effective data reduction without loss of relevant informational content that results from PCA is expected to increase sensitivity to subtle spectral differences relative to direct linear analysis of the full data matrix. Furthermore, direct PCA of the spectral matrix does not require identification and measurement of specific peak attributes, which may be advantageous in overlapping spectral regions such as is found in the central region of the NISTmAb methyl spectrum. Additionally, given that PCA tends to spread the random variation from noise across all principal components, less interference from noise is expected along the lower order components where variability of greatest interest is located, making PCA more robust in sensitivity limited situations.

To investigate the performance of a PCA based approach to therapeutic mAbs, a spectral library was compiled of native, ExoGal and DeGly-NISTmAb ^{1}H-^{13}C gsHSQC spectra collected at both 900 MHz and 600 MHz. This pulse sequence was chosen in concert with conventional uniform indirect acquisition to benchmark the PCA method using high performing and well established NMR techniques. Variable numbers of scans were employed for each NISTmAb species in order to account for the effects of experimental noise in the PCA

analysis; effective S/N for each dataset is given in Table S3. All spectra were analyzed over the methyl region window from -0.6 ppm – 2.4 ppm in ^{1}H and 8 ppm – 25 ppm in ^{13}C. This spectral region contains both well resolved peaks in the periphery and overlapped signals in the central region of the window. The results of the PCA on the 900 MHz data are illustrated as a scatterplot of the 1st and 2nd principal components in Figure 2. Using this type of analysis, both the DeGly- and ExoGal-NISTmAb species are clearly distinguished from the native NISTmAb, with good separation of the 95 % confidence interval of each cluster. Interestingly, we find that low S/N data, collected using as few as 16 scans, clusters together fairly well with high S/N data, confirming our expectation that random variation from noise can be spread across higher order principle components allowing high and low S/N data to be compared effectively. Although of lower resolution, the data collected at 600 MHz gives qualitatively similar distribution of the 1st and 2nd principal components (Figure S4). Because we employ a whole data matrix approach to spectral analysis, acquisition parameters which affect the global shape of the spectrum, such as those that result in differences in the digitization of the signal, differences in line shape and resolution, or distribution of noise, must be carefully matched. Indeed, simultaneous PCA of 600 MHz and 900 MHz data results in a bimodal distribution along the 1st principal component according to field strength (Figure S5); therefore, in general it is essential that all acquisition parameters be carefully matched for datasets included in PCA.

Given that PCA applied directly to the spectral data matrix can be sensitive to experimental conditions, it is fair to consider the general robustness of the method to potential extrinsic sources of variability, including those arising from differences in instrumentation, spectral processing parameters and sample preparation. While NMR chemical shifts have been found to be highly reproducible across instruments and laboratories, PCA on data acquired on different instruments of the same field strength may prove problematic. Indeed, a previous NMR interlaboratory study on the small protein therapeutic filgrastim, displayed a slight clustering of spectra by instrument, although, experimental parameters and instrument performance were not carefully matched in this study.[18,19] Sample conditions must also be carefully considered. Because signals from formulation components can be orders of magnitude greater than those from the protein, even minor differences in formulation conditions may lead to bias in the PCA results. As such, it often may be best if formulation signals are excluded from analysis, whether by replacement with deuterated analogs or by trimming and tailoring of the region of analysis to their exclusion. However, for comparisons of samples in identical formulations conditions this is not absolutely necessary, so long as formulation conditions are carefully controlled and in fact, formulation signals themselves may then serve as useful reporters of relevant sample attributes. Spectral processing, on the other hand, is generally of less a concern as relevant processing parameters such as phasing, spectral referencing/alignment and baseline correction can largely be automated and are not expected to vary greatly for data collected systematically. Still, for the case in which spectral processing contributes to variability in the dataset, such as if autophasing were to introduce small phase offsets, the effect is to reduce intracluster precision, i.e. increase the size of cluster intervals. However, as with experimental noise, these effects are more predominant in higher order components, and the effect in the lower order components of interest is only modest (Table S4).

This PCA approach highlights a special strength of 2D heteronuclear NMR relative to 1D ^{1}H NMR methods, which themselves have been demonstrated to have high sensitivity to structural variation when appropriate methodology is employed.[3,13] In addition to producing scoring plots to provide statistical differentiation of HOS variation, PCA also produces spectral loading plots, a subset of the full data-matrix describing the sources of variability along each principal component. These results can be visualized as spectra and thus can reveal which particular regions of the spectrum are undergoing variation. In the case of 2D NMR, where peaks can be associated with specific residues, this provides an opportunity to identify the residues associated with observed variations. So, while both 1D and 2D methods can be used for chemometric analysis and classification, only 2D methods provide the opportunity to map sources of spectral variability to structure at

high resolution. The choice of a 1D or 2D based approach remains one of fit for purpose; with 1D methods preferred when time and not high-resolution is the defining need of the assay, whereas 2D methods will be preferred when a greater level of detail afforded by methods such as PCA is required.

Shown in Figure 3 are the loading plots for the 2nd principal components of the 900 MHz data overlaid on the average spectrum from the series. Much of the source of variation comes from the overlapped region of the spectrum; as such traditional methods of spectral analysis relying on specific peak analysis would be predicted to fail in successfully identifying the majority of the variation. Still some variation arises from well resolved peaks and in the case where an assignment of specific resonances is known this variation could be directly attributed to specific residues and regions of the structure. For example, while specific resonance assignments are not yet currently available for the NISTmAb methyl spectrum, the spectral loading plot may be compared to previously recorded spectra on the individual NISTmAb Fc or Fab domains.[8] From this analysis we are able to determine that the major sources of variation are located in the Fc domain (Figure S6) as might be expected from remodeling of the Fc attached glycan. Likewise, four well resolved signals in the downfield region of the spectrum display a strong variability amongst the three NISTmAb isoforms (Figure 3); these signals have been attributed to acetyl methyl groups from the four N-acetyl glucosamine (GlcNAc) sugar residues of the NISTmAb glycan. While this result may be trivial, arising either from direct change to the chemical environment of the GlcNAc residue for the ExoGal-NISTmAb or complete loss of GlcNAc signal for the DeGly-NISTmAb and not from perturbation of the protein HOS, it further serves to illustrate how results of PCA can be linked back to high resolution spectra which, in turn can be mapped on to structure.

Given the robust performance of PCA in discriminating the modest structural and spectral variation manifest in the native and ExoGal-NISTmAb species, it is fair to ask, what the limit of detection is for this chemometric analysis method applied to 2D NMR data. In order to gain insight into this line of inquiry, we prepared blended samples of the native and ExoGal- NISTmAb species at 100/0 %, 50/50 % 25/75 %, 15/85 %, 7.5/92.5 % and 0/100 % ExoGal:native composition and recorded a series of systematic ^{1}H-^{13}C methyl spectra on the samples. The PCA scoring plot for this series, shown in Figure 4, clearly differentiates even small changes in a blended isoform composition. Visual differentiation of the clusters is observed between the 15 %, 7.5 % and 0 % populations; however they are not clearly statistically differentiated given overlap of their 95 % cluster confidence intervals. This suggests that a practical limit, as defined by the 95 % cluster confidence interval, of a differentiation of 10 % relative isoform composition for this example. It is useful to reiterate that this limit is for a minor structural variant, whose pure spectrum cannot be visually distinguished from that the major, native species. For a variant, whose pure spectrum is visually different, such as the deglycosylated NISTmAb, we expect the limit of detection to be much lower than 10%. Additionally, if a different cluster interval is employed, such as twice the standard deviation of the spread of the cluster, statistical differentiation of the 15 %, 7.5 % and 0 % clusters can be achieved. This suggests that if more datasets for each population could be acquired with high precision, confidence in differentiation between populations with less than 10 % relative ExoGal composition could be achieved.

CONCLUSIONS

The ability to characterize the HOS of mAb therapeutics at high resolution is of considerable importance to the pharmaceutical industry. 2D NMR has been established as capable of delivering high quality, high resolution spectra to allow such characterization in reasonable time frames. The best methods for analysis of such spectra in terms of spectral and structural differences remain an open area of inquiry. Here, it has been demonstrated that multivariate approaches such as PCA are well suited to this task. Using the benchmark ^{1}H-^{13}C gsHSQC experiment and conventional acquisition, with PCA applied directly to the spectral data matrix, spectra from highly similar species have been successfully discriminated, with low limits of detection, which

could not be distinguished by visual inspection or simple intensity based statistical approaches. The major sources of spectral variation between species have been identified and provisionally ascribed to structure.

This PCA approach has also been shown to be highly tolerant of experimental noise. Although the degree to which it may be combined with rapid acquisition techniques, in particular non-uniform sampling, remains an open area of inquiry, when taken together with our previous report demonstrating rapid acquisition of ^{1}H-^{13}C spectra of mAb therapeutics at natural isotopic abundance in a matter of minutes-to few hours,[8] these results demonstrate that high quality, high resolution 2D NMR spectral data can be used to achieve high precision in distinguishing HOS variation. As such, 2D ^{1}H-^{13}C methods are competitive with other traditional, lower resolution HOS characterization techniques, while still providing the potential for atomic level assignment of structural differences. Furthermore, this allows for the possibility of benchmarking traditional lower resolution HOS characterization techniques against high-resolution NMR data to both properly assess the degree to which those techniques are sensitive to relevant differences in HOS, as well as assess whether spectral differences may be correlated to structure via corresponding 2D NMR data, thus ultimately enriching the utility of both high and low resolution approaches. As such, we expect that 2D NMR methods in combination with PCA and other chemometric analytical approaches should find robust use in statistical comparison of HOS of mAb therapeutics.

ASSOCIATED CONTENT

Supporting Information

Additional information as noted in text. This material is available free of charge on the ACS Publications website.

AUTHOR INFORMATION

Corresponding Author

*E-mail john.marino@nist.gov

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest

ACKNOWLEDGMENT

We acknowledge the support by NIST Biomanufacturing Initiative and NIST and W.M. Keck for support of Biomolecular NMR instrumentation.

REFERENCES

(1) Aubin, Y.; Jones, C.; Freedberg, D. I. *Biopharm Int.* **2010**, *S28*.

(2) Visser, J.; Feuerstein, I.; Stangler, T.; Schmiederer, T.; Fritsch, C.; Schiestl, M. *BioDrugs* **2013**, *27*, 495–507.

(3) Poppe, L.; Jordan, J. B.; Lawson, K.; Jerums, M.; Apostol, I.; Schnier, P. D. *Anal. Chem.* **2013**, *85*, 9623–9629.

(4) Franks, J.; Glushka, J. N.; Jones, M. T.; Live, D. H.; Zou, Q.; Prestegard, J. H. *Anal. Chem.* **2016**, *88*, 1320–

1327.

(5) Japelj, B.; Ilc, G.; Marušič, J.; Senčar, J.; Kuzman, D.; Plavec, J. *Sci. Rep.* **2016**, *6*, 32201, 1-12.

(6) Chen, K.; Freedberg, D. I.; Keire, D. A. *J. Magn. Reson.* **2015**, *251*, 65–70.

(7) Marino, J. P.; Brinson, R. G.; Hudgens, J. W.; Ladner, J. E.; Gallagher, D. T.; Gallagher, E. S.; Arbogast, L. W.; Huang, R. Y.-C. In *State-of-the-Art and Emerging Technologies for Therapeutic Monoclonal Antibody Characterization Volume 3.*; Schiel, J. E., Davis, D. D., Borisov, O. V., Eds.; American Chemical Society: Washington, DC, **2015**; pp 17–43.

(8) Arbogast, L. W.; Brinson, R. G.; Marino, J. P. *Anal. Chem.* **2015**, *87*, 3556–3561.

(9) Arbogast, L. W.; Brinson, R. G.; Formolo, T.; Hoopes, J. T.; Marino, J. P. *Pharm. Res.* **2016**, *33*, 462–475.

(10) Arbogast, L. W.; Brinson, R. G.; Marino, J. P. In *Methods in Enzymology*; **2016**; Vol. 566, pp 3–34.

(11) Singh, S. M.; Bandi, S.; Jones, D. N. M.; Mallela, K. M. G. *J. Pharm. Sci.* **2017**, *In Press*.

(12) McKenzie, J. S.; Donarski, J. A.; Wilson, J. C.; Charlton, A. J. *Prog. Nucl. Magn. Reson. Spectrosc.* **2011**, *59*, 336–359.

(13) Poppe, L.; Jordan, J. B.; Rogers, G.; Schnier, P. D. *Anal. Chem.* **2015**, *87*, 5539–5545.

(14) Chen, K.; Long, D. S.; Lute, S. C.; Levy, M. J.; Brorson, K. A.; Keire, D. A. *J. Pharm. Biomed. Anal.* **2016**, *128*, 398–407.

(15) Kheddo, P.; Cliff, M. J.; Uddin, S.; van der Walle, C. F.; Golovanov, A. P. *MAbs* **2016**, 1–14.

(16) Amezcua, C. A.; Szabo, C. M. *J. Pharm. Sci.* **2013**, *102*, 1724–1733.

(17) Guerrini, M.; Rudd, T. R.; Mauri, L.; Macchi, E.; Fareed, J.; Yates, E. A.; Naggi, A.; Torri, G. *Anal. Chem.* **2015**, *87*, 8275–8283.

(18) Ghasriani, H.; Hodgson, D. J.; Brinson, R. G.; McEwen, I.; Buhse, L. F.; Kozlowski, S.; Marino, J. P.; Aubin, Y.; Keire, D. A. *Nat. Biotechnol.* **2016**, *34*, 139–141.

(19) Brinson, R. G.; Ghasriani, H.; Hodgson, D. J.; Adams, K. M.; McEwen, I.; Freedberg, D. I.; Chen, K.; Keire, D. A.; Aubin, Y.; Marino, J. P. *J. Pharm. Biomed. Anal.* **2017**, *141*, 229–233.

(20) Župerl, Š.; Pristovšek, P.; Menart, V.; Gaberc-Porekar, V.; Novič, M. *J. Chem. Inf. Model.* **2007**, *47*, 737–743.

(21) Yamaguchi, Y.; Nishimura, M.; Nagano, M.; Yagi, H.; Sasakawa, H.; Uchida, K.; Shitara, K.; Kato, K. *Biochim. Biophys. Acta - Gen. Subj.* **2006**, *1760*, 693–700.

(22) Delaglio, F.; Grzesiek, S.; Vuister, G.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277–293.

(23) Wold, H. In *Research Papers in Statistics*; David, F. N., Ed.; John Wiley & Sons, Inc.: New York, **1966**; pp 411–444.

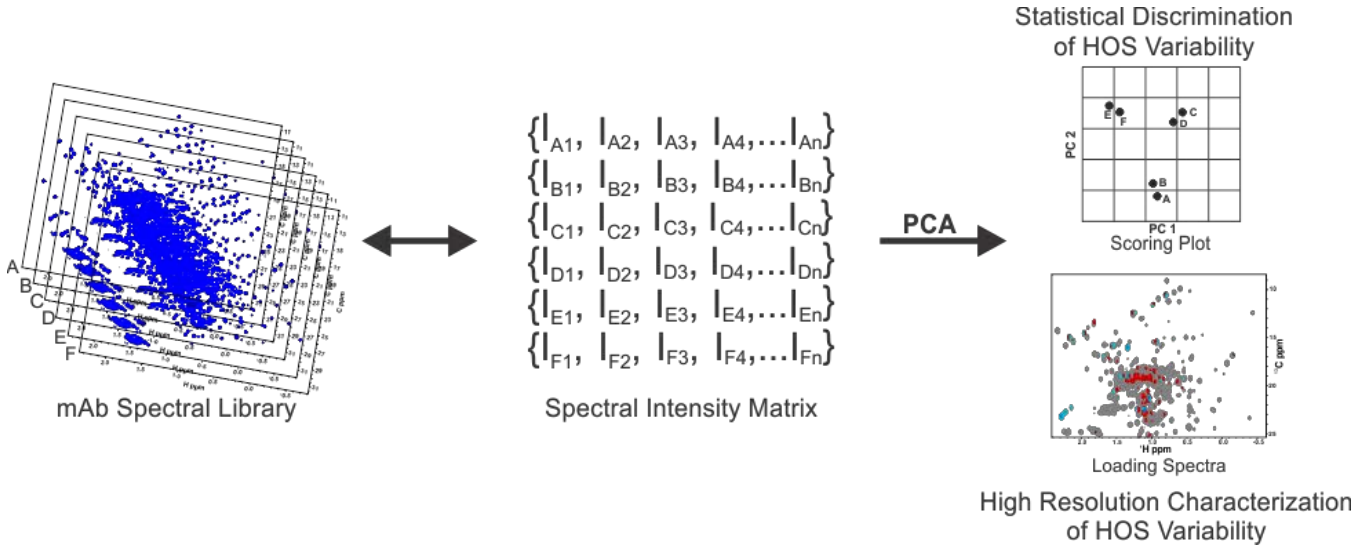
(24) Prien, J. M.; Stöckmann, H.; Albrecht, S.; Martin, S. M.; Varatta, M.; Furtado, M.; Hosselet, S.; Wang, M.; Formolo, T.; Rudd, P. M.; Schiel, J. E. In *State-of-the-Art and Emerging Technologies for Therapeutic Monoclonal Antibody Characterization Volume 2. Biopharmaceutical Characterization: The NISTmAb Case Study*; Schiel, J. E., Davis, D. L., Borisov, O. V., Eds.; American Chemical Society: Washington, DC, **2015**; pp 185–235.

(25) Krapp, S.; Mimura, Y.; Jefferis, R.; Huber, R.; Sondermann, P. *J. Mol. Biol.* **2003**, *325*, 979–989.

(26) Frank, M.; Walker, R. C.; Lanzilotta, W. N.; Prestegard, J. H.; Barb, A. W. *J. Mol. Biol.* **2014**, *426*, 1799–1811.

(27) Subedi, G. P.; Barb, A. W. *Structure* **2015**, *23*, 1573–1583.

For Table of Contents Only



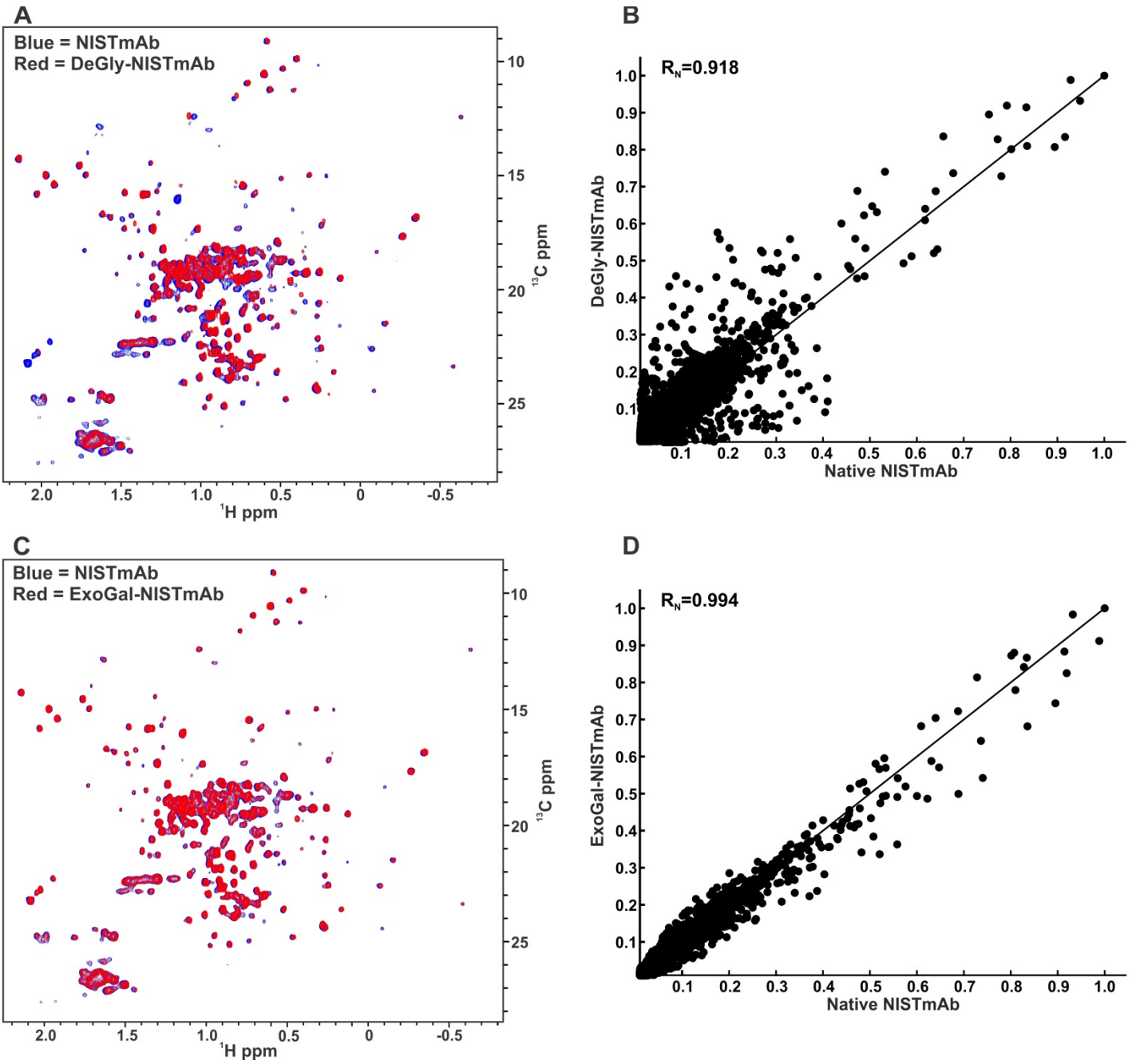


Figure 1 Comparison of ^{1}H-^{13}C methyl spectra of the NISTmAb with the DeGly NISTmAb by spectral overlay (A) and linear correlation analysis (B) as well as the native NISTmAb with the ExoGal NISTmAb by spectral overlay (C) and linear correlation analysis (D). Reported correlation coefficients, R_N, are normalized to the expectation value based on the average S/N of the datasets.

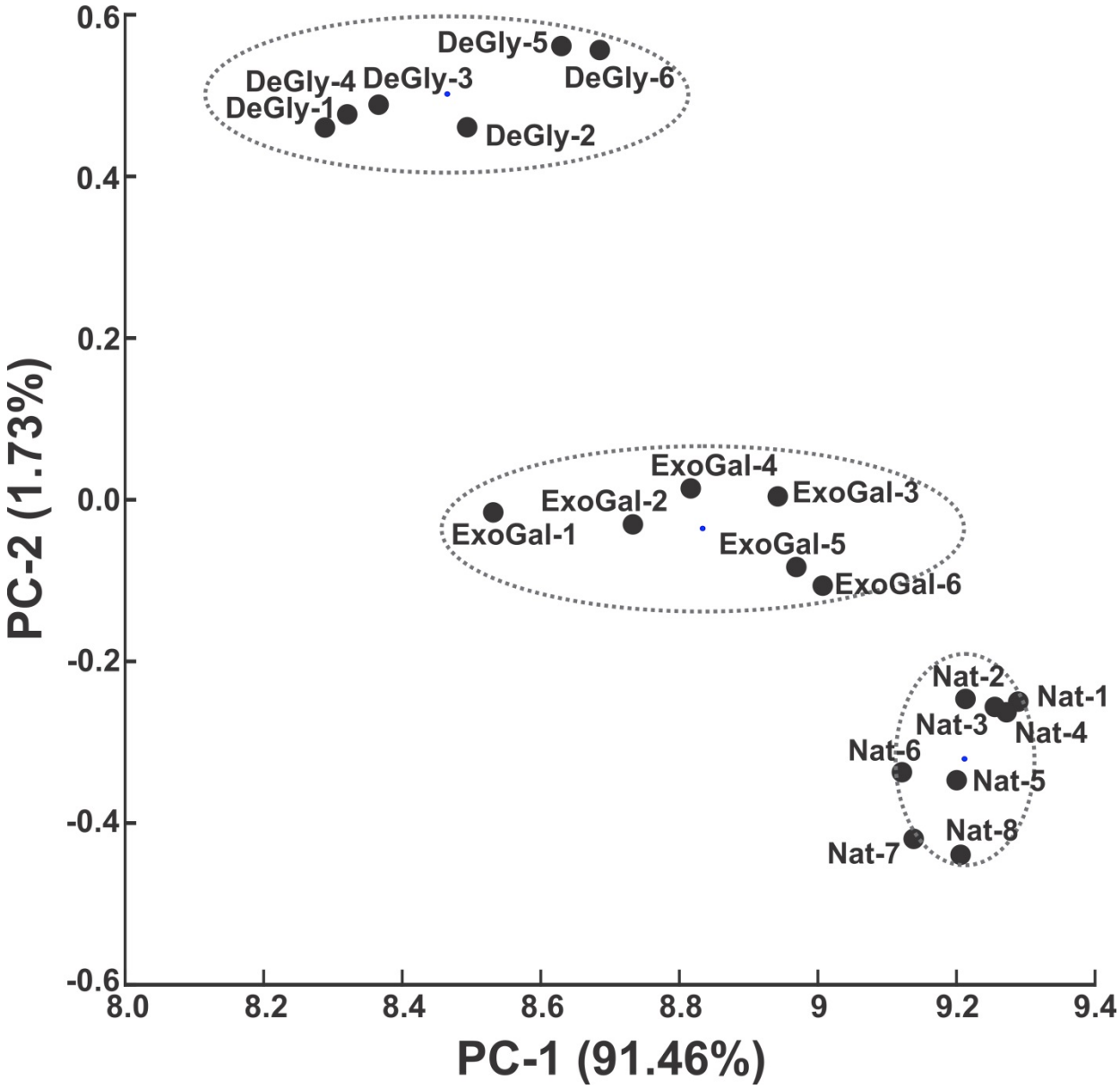


Figure 2 PCA Scatterplot from 20 ^{1}H-^{13}C-methyl spectra of NISTmAb isoforms collected at 900 MHz and 50 °C. Points in the scatterplot correspond to spectra listed in Table S2. Dashed lines correspond to the 95 % confidence interval of the spread for each cluster.

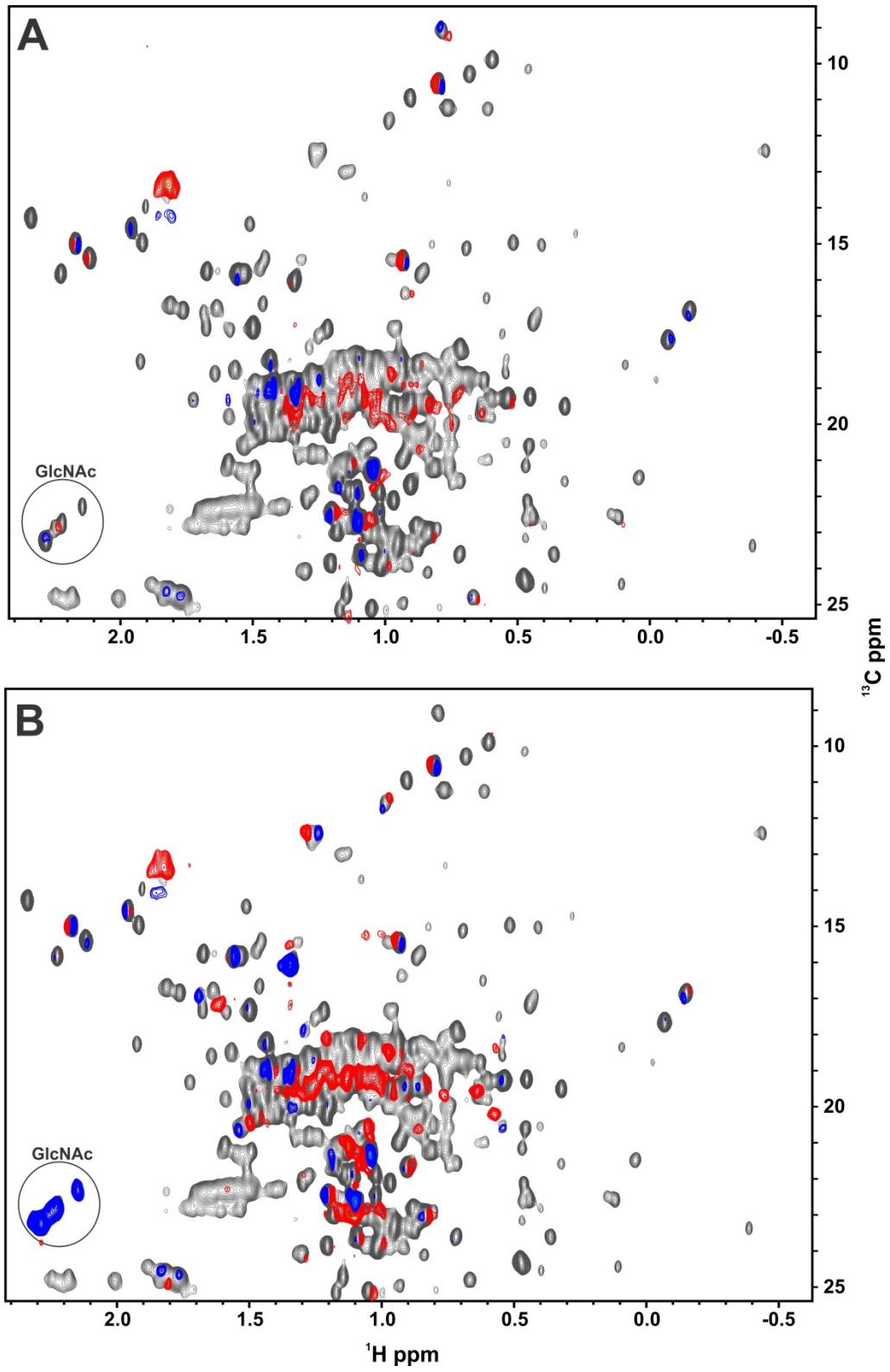


Figure 3 Spectral loading plots of the ExoGal- (A) and DeGly- (B) NISTmAb, showing the variance in the 2nd principal component (red and blue contours) of respective 900 MHz ^{1}H-^{13}C methyl spectral series overlaid on the average series spectrum (grey contours). Red/blue contours from the 2nd PC loading plot show areas of greater/lesser intensity in the ExoGal and DeGly isoform spectra relative to the average spectrum of the series.

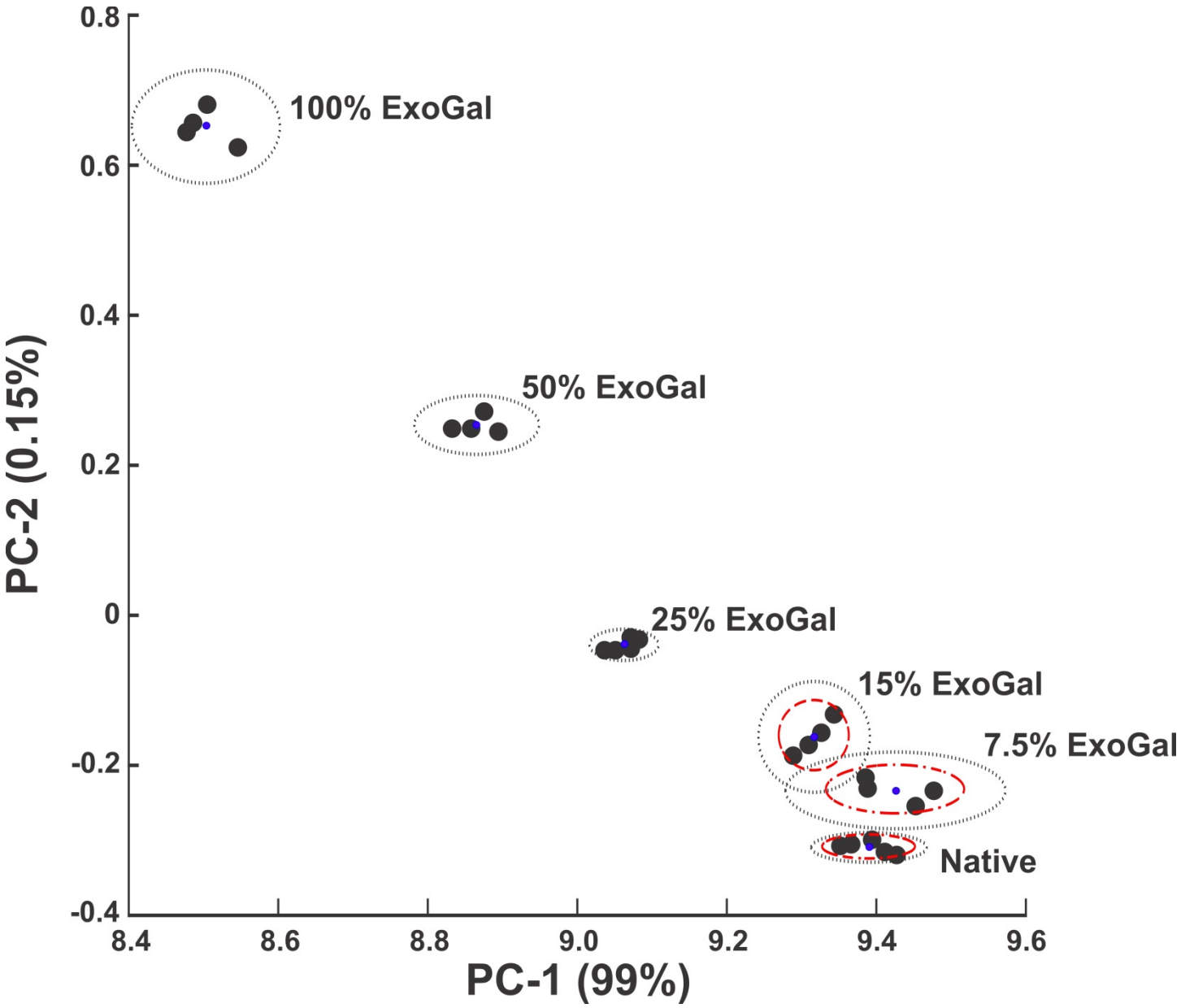


Figure 4 PCA Scatterplot from 26 ^{1}H-^{13}C-methyl spectra of blended ExoGal/native NISTmAb collected at 900 MHz and 50 °C. Dashed grey lines correspond to the 95 % confidence interval of the spread for each cluster. Dashed red lines on the 15 %, 7.5 % and native clusters correspond to the 2 σ interval of the spread in the respective clusters.