SUPPORTING INFORMATION

On the Analytical Superiority of 1D NMR for Fingerprinting the Higher Order Structure of Protein Therapeutics Compared to Multidimensional NMR Methods

Leszek Poppe[1*], *John B. Jordan*[1], *Gary Rogers*[2] *and Paul D. Schnier*[1‡]

[1]Discovery Attribute Sciences, Amgen Inc. One Amgen Center Drive, Thousand Oaks, Ca 91320

[2]Product Attribute Sciences, Amgen Inc. One Amgen Center Drive, Thousand Oaks, Ca 91320

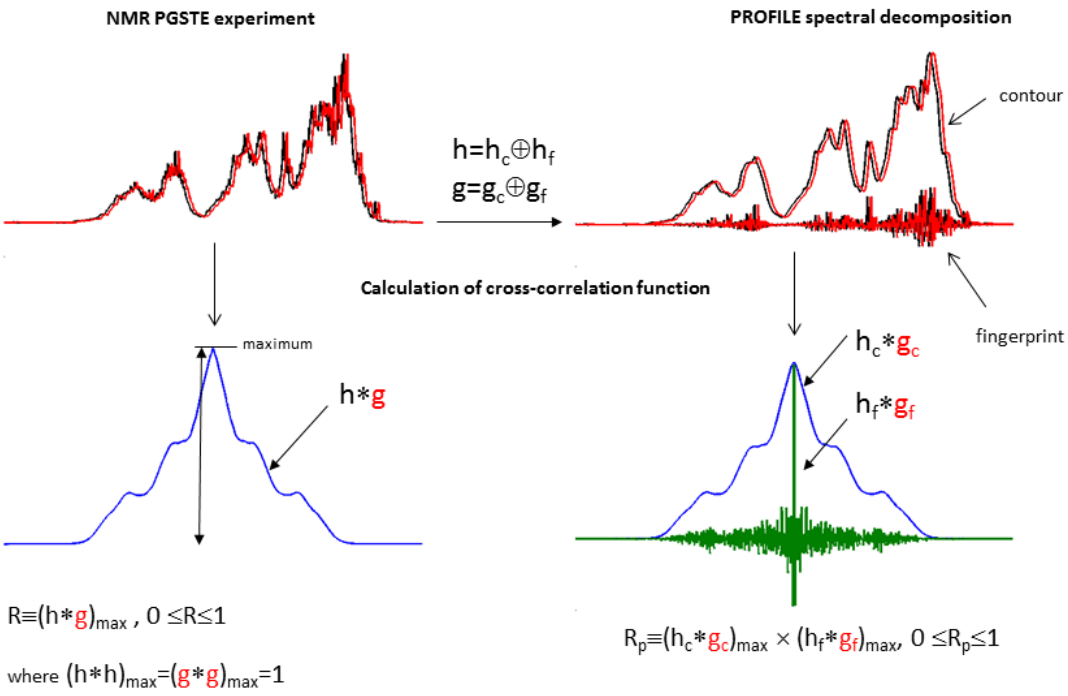

Figure S1. Principles of the calculation of cross-correlation coefficients from the 1D ^{1}H spectra and from the PROFILE spectra.

‡ Present address: Department of Protein Chemistry, Genentech, Inc., South San Francisco, CA 94080

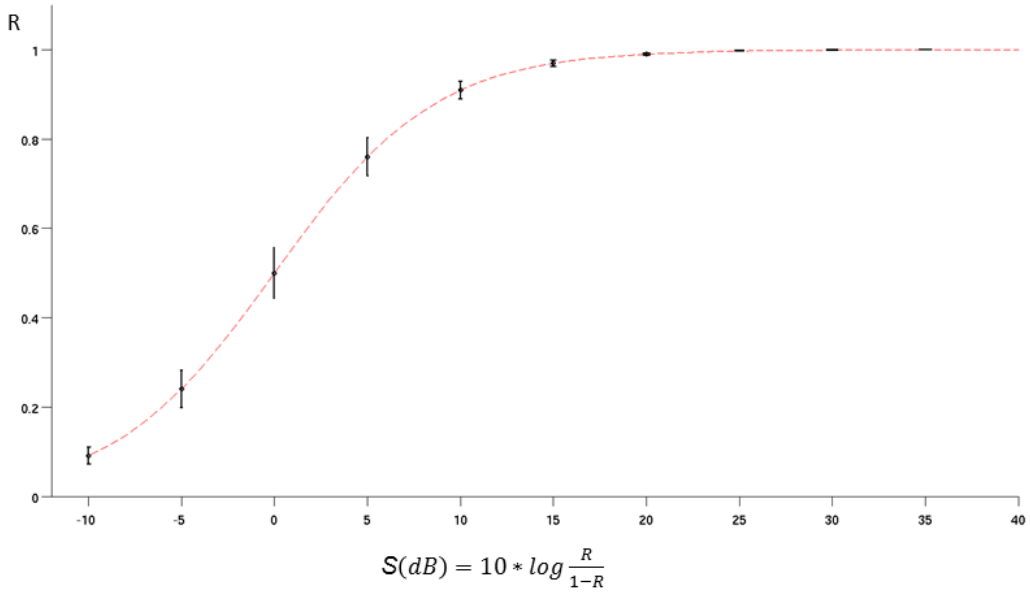

Figure S2. Relationship between the cross-correlation coefficient and the similarity S. The vertical bars correspond to 1 dB difference in S.

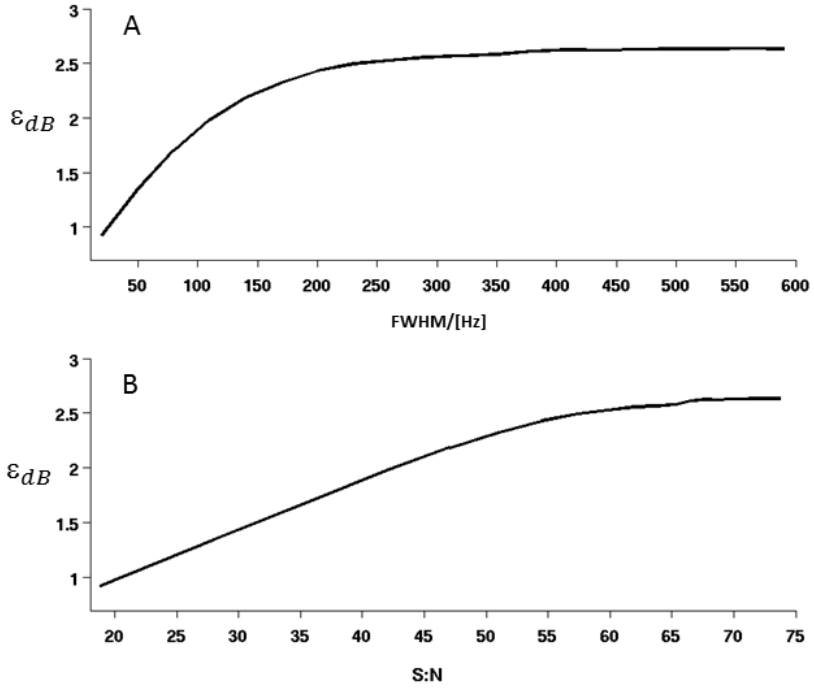

Figure S3. A – Similarity difference between the IgG1 and $^{95\%}$IgG1:$^{5\%}$IgG2 samples as a function of FWHM of the Gaussian smoothing function used to generate the contour. The original spectra were recorded with 256 scans and 50 mg/mL samples. B – the same relationship as in A but instead of FWHM the S:N in the fingerprint spectrum was used. The S:N in the original spectra was 430±30.

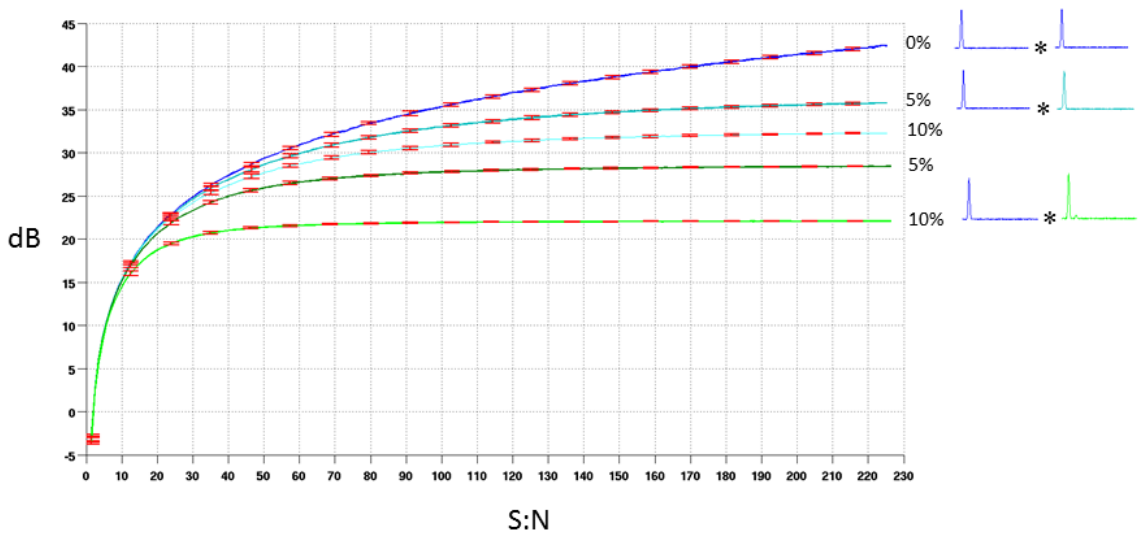

Figure S4. Calculation of the similarity between two Gaussian signals with the different levels of random noise. The signals were calculated for 1024 points as the Gaussian lines with FWHM of 33 points and different positions, where the blue colored lines correspond to 10 points difference in relative positions and the green colored lines correspond to 100 points difference. In this way the blue and green simulations reflect the peak resolution in the PROFILE or the 2D heteronuclear correlation spectra respectively. The % values correspond to the degree of mixing of the two signals. The S:N values were calculated as described in the Experimental Section. Since S:N is sligthly higher in the 0% spectrum (T), the mixed signals (M1 and M2) were simulated twice with equal levels of the random noise, and then coadded (M=M1+γM2) to match the S:N in T before the similarity S(T,M) calculation. The γ coefficient was calculated from the following formula: $\gamma = \frac{1-\sqrt{1-4\eta^2}}{2\eta}$, where $\eta = \frac{\alpha^2-1}{2\alpha}$, where $\alpha = \frac{S{:}N(T)}{S{:}N(M1)}$.

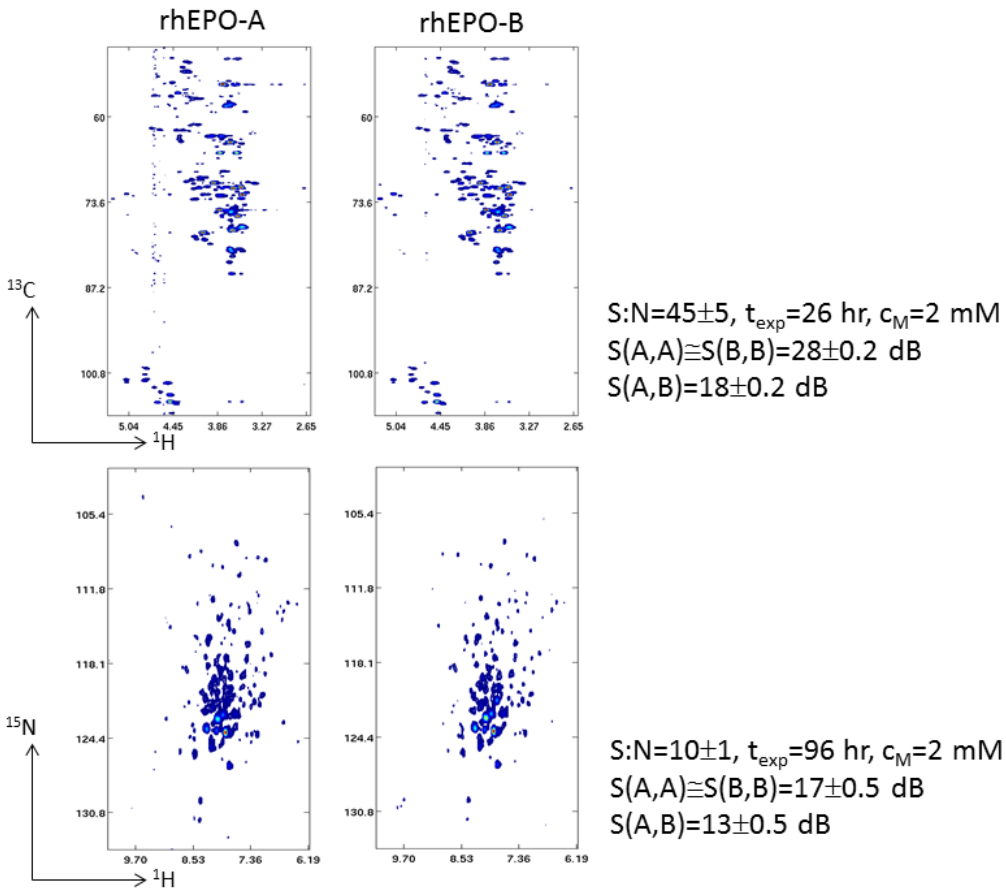

Figure S5 Similarity of the ^{1}H-^{13}C and ^{1}H-^{15}N spectra recorded for the EPO samples of the different origin (as described in the text).

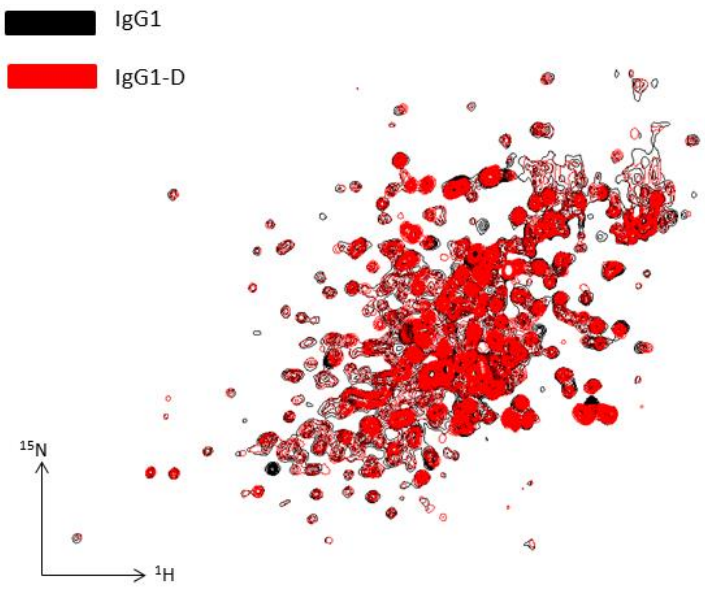

Figure S6. The overlay of the ^{1}H-^{15}N TROSY spectra of the native IgG1 (black) and the deglycosylated form (red). The similarity between both spectra is 5.3 dB.

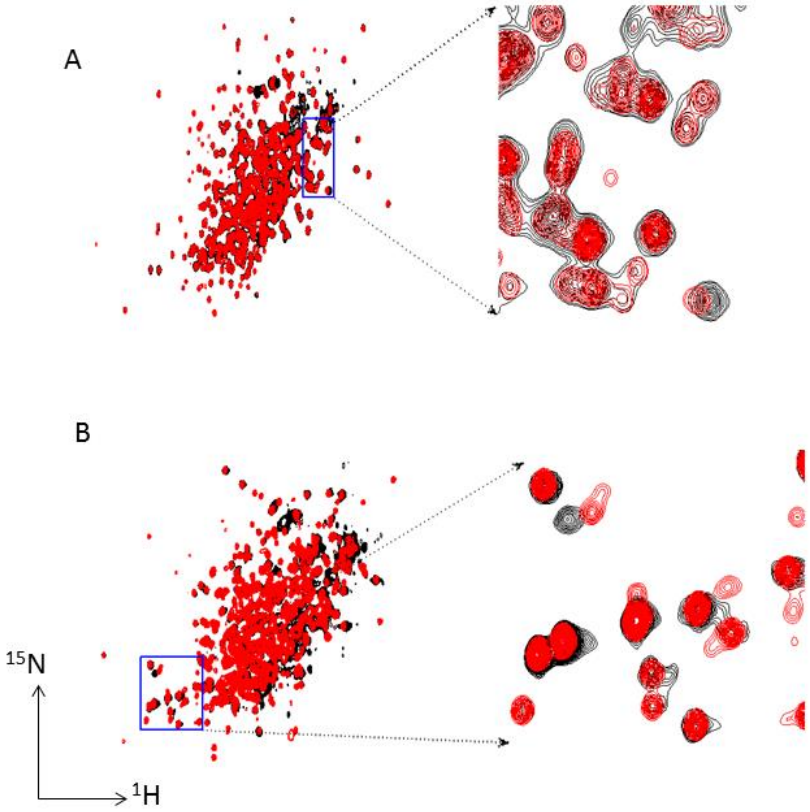

Figure S7 Overlays of the 2D ^{1}H-^{15}N TROSY spectra for the intact IgG1 (A) and IgG2 (B) antibodies with the combined spectra from the corresponding Fc and Fab fragments. The spectral similarities are 5.2 dB in A and 0.52 dB in B.

Table S1.List of the samples and experimental conditions of this study.

Protein[1]	[mM]	Shigemi /volume	T (K)	experiment	t_{exp} (hr)	S:N[2]
Epo-A(3 lots)	0.3	4mm/180μL	305	PROFILE	2.33	95
Epo-B(3 lots)	0.3	4mm/180μL	305	PROFILE	2.33	95
Epo-A(2 lots)	2	4mm/180μL	305	PROFILE	0.12	150
Epo-B(2 lots)	2	4mm/180μL	305	PROFILE	0.12	150
Epo-A(2 lots)	2	5mm/300μL	305	TROSY(^{15}N)	96	10
Epo-B(2 lots)	2	5mm/300μL	305	HSQC(^{13}C)	26	45
IgG1 (blends)	0.3	4mm/180μL	318	PROFILE	0.24	64
IgG1-D	0.3	4mm/180μL	318	PROFILE	0.24	63
^{15}N-IgG1 (blends)	0.1	4mm/180μL	318	PROFILE	2	62
^{15}N-IgG1 (blends)	0.1	4mm/180μL	318	TROSY(^{15}N)	6	20
^{15}N-(Fab)2 (IgG1)	0.1	4mm/180μL	318	TROSY(^{15}N)	24	53
^{15}N-Fc (IgG1)	0.1	4mm/180μL	318	TROSY(^{15}N)	24	60
^{15}N-IgG1	0.1	4mm/180μL	318	TROSY(^{15}N)	24	45
^{15}N-(Fab)2 (IgG2)	0.1	4mm/180μL	318	TROSY(^{15}N)	24	45
^{15}N-Fc (IgG2)	0.1	4mm/180μL	318	TROSY(^{15}N)	24	60
^{15}N-IgG2	0.1	4mm/180μL	318	TROSY(^{15}N)	24	40

[1] the Epo samples were in 20 mM phosphate buffer, all the IgG samples were in 10 mM acetate buffer with 9% sucrose.

[2] S:N was calculated as described in the main text. Listed are mean values with standard deviations not exceeding 10% (obtained from the different samples)