PCCP

ARTICLE TYPE

Cite this: DOI: 00.0000/xxxxxxxxxx

Electronic Supporting Information (ESI) to Advancing Predictions of Protein Stability in the Solid State

Maarten Batens,[b‡] Talia A. Shmool,[c‡] Jan Massant,[d‡] J. Axel Zeitler,[d‡] and Guy Van den Mooter*[a]

Supplementary Methods

Preparation of Feed Solutions

A Vivaflow 200 Laboratory Cross Flow Cassette (Sartorius, Göttingen, Germany) was used in a tangential flow diafiltration process to exchange the mAb's storage buffer with the new formulation buffer containing 15 mM L-histidine (L-his and L-hisHCl), pH 5.5, and either 25 mg mL^{-1} trehalose or 25 mg mL^{-1} glycerol prepared with ultrapure water (Type 1, (Resistivity (ρ) $\geq$ 18.2 MΩcm at 298.15 K). Following diafiltration, solutions were further concentrated and filtered (Polyethersulfone, 0.22 μm, Merck Millipore, Bedford MA, USA). The concentration of the mAb solutions was measured using UV absorbance at 280 nm with extinction coefficient 1.33 mL mg^{-1} cm^{-1} and diluted to 50 $\pm$ 0.5 mg mL^{-1} with ultrapure water. Following preparation, feed solutions were stored at 278 K before spray-drying.

Karl Fischer Titration (KF)

The residual water content of the spray-dried mAb formulations was determined using a 831 KF Coulometer with generator electrode (without diaphragm), coupled to a 774 Oven Sample Processor (Metrohm AG, Herisau, Switzerland). Samples (20 - 30 mg) were heated to 393 K in 6 mL clear glass head space vials, which were closed using septum seals with polytetrafluoroethylene inserts (Metrohm). Water contents are expressed as mass percentages unless otherwise specified.

X-Ray Powder Diffraction (XRPD)

An automated X'pert PRO X-ray diffractometer (PANalytical, Almelo, The Netherlands), equipped with a Cu radiation source (λ_{K_α} 1.5418 Å, voltage 45 kV, current 40 mA), was used to determine powder crystallinity. X-ray powder diffraction (XRPD) measurements were conducted in transmission mode using Kapton® (DuPont, Wilmington, Delaware) film. Samples were spun at a rate of 4 seconds per rotation with a counting time of 400 seconds per step and a step size of 0.0167 ° in the range from 4 ° to 40 ° 2θ. Data was analysed using version 1.7 of the Data Viewer software (PANalytical).

Differential Scanning Calorimetry (DSC)

To determine the calorimetric glass transition temperature ($T_{g,\alpha,DSC}$), a DSC 3+ differential scanning calorimeter (Mettler Toledo, Columbus OH, USA) was used to determine $T_{g,\alpha,DSC}$, which was defined as the onset temperature of the step determined with a tangential baseline, as determined by version 16.10 of the STARe Evaluation Software (Mettler Toledo). For each sample, 2 - 10 mg were placed in a 40 μl standard aluminium crucible and closed using a standard lid (Mettler Toledo) under ambient conditions. Samples were subsequently heated at a rate of 10 K min^{-1} from 233 or 153 to 353 K, for F1 and F4 or F2 and F3, respectively. For t_0, measurements were done in triplicate, for future time points, a single measurement was performed unless no clear glass transition could be determined using the obtained thermogram.

Fourier Transform Infrared Spectroscopy (FTIR)

Changes in the secondary structure of the mAb in the solid state were assessed based on relative changes in the positions of peaks in the 1700 - 1600 cm^{-1} and 1580 - 1520 cm^{-1} range, i.e. the amide I and amide II bands, respectively. Spectra were recorded (64 scans) in the 400 to 4000 cm^{-1} range with a spectral resolution of 2 cm^{-1} using a Vertex 70 Fourier transform infra-red (FTIR) spectrometer (Bruker, Billerica, MA, USA). Data was pre-processed and processed with version 7.5 of the OPUS software package (Bruker). Data pre-processing was performed on the absorbance spectra and consisted of atmospheric compensation (CO_2 and H_2O), rubber band baseline correction (128 points, with exclusion of CO_2 bands) and vector normalisation, all of which were performed using the entire measured wavenumber range.

Dynamic Light Scattering (DLS)

A Wyatt Möbiuζ Zeta Potential and DLS detector (Wyatt, Santa Barbara, CA, USA) was used to quantify the fraction of large aggregates (diameter range 2.0 nm to 2.0 μm) present in the sam-

* Corresponding author

[a] Drug Delivery and Disposition, KU Leuven, Leuven, Belgium. Tel: +32 16 33 03 04; E-mail: guy.vandenmooter@kuleuven.be

[b] Drug Delivery and Disposition, KU Leuven, Leuven, Belgium.

[c] Biological Formulation Development, UCB Pharma, Braine l'Alleud, Belgium.

[d] Department of Chemical Engineering and Biotechnology, University of Cambridge, Cambridge, UK.

‡ These authors contributed equally to this work

ples after reconstitution. The laser wavelength was 532 nm, detector angle 163.5 ° and samples were diluted to 10 mg mL^{-1} using a 60 mM L-histidine buffer. Data were recorded at 298 K, using auto-attenuation and the peak radius cut-off set at 2.0 - 2000.0 nm to reduce buffer signals. The number of acquisitions and acquisition times were optimised for each sample, based on acquired DLS correlation curves (intensity autocorrelation as a function of time). Processing was done using the Dynamics software, version 7.3.1.15 (Wyatt).

Turbidimetry (OD600)

Optical density at incident wavelength of 600 nm (OD600) was taken as a measure for the turbidity of reconstituted samples, an additional criterion for the presence of aggregates exceeding the inherent size limits of SEC and DLS. Measurements were performed using a Spectramax M5 multi-detection microplate reader combined with the SoftMax Pro software package version 7.0.3 GxP (Molecular Devices, Sunnyvale, CA, USA).

Supplementary Results

Stability Study

All spray-dried mAb formulations were both XRPD and DSC amorphous during the entire 52-week stability study. Further solid state characterisation at each time point did not reveal any trends over time for the calorimetric glass transition temperature ($T_{g,\alpha,DSC}$). No clear peak shifts were detected over time for the amide I or amide II bands (Fig. S1). There were no changes in reconstitution time upon storage. F3 consistently exhibited the longest reconstitution time. The large variability in the reconstitution time was likely a consequence of the nature of the method that relies on qualitative visual observation, as well as the intrinsic variability due to their particulate behaviour, e.g. particles sticking together or to the vial wall, and the increasing difficulty of correctly assessing reconstitution time end points as sample turbidity increased.

With regard to aggregation-based stability assessment, Fig. S2 summarises the HMWS content, large aggregates and turbidity of the formulations stored at 313 K as a function of time. Overall, a relative increase in the HMWS over time was observed for all formulations, with the increase over time being the most pronounced for formulation F3, i.e. the formulation with the highest glycerol content and no trehalose. In this regard it should be noted that for F3, at the 39 weeks time point, the recovery, expressed as the total area under the curve as a measure for the amount of protein that passed through the column, was only about 12 % of the recovery at t_0. This indicated that the majority of mAb species in the sample were likely too aggregated to pass through the SEC column and withheld by the guard column. This hypothesis was further substantiated by the fact that F3, stored at 313 K for 52 weeks, was too viscous to be injected in the column following reconstitution. To avoid having to omit the 39 weeks time points for formulation F3 stored at 313 K, measurements for this time point were corrected for the low recovery by adding the area of the unrecoverable fraction relative to t_0 as HMWS. The validity of this correction was corroborated by a substantial increase in the coefficient of determination (R^2) compared to the original data set, since the uncorrected % area HMWS was lower than the previous time point, breaking the linear trend. Additional confirmation of validity was given by assessing the DLS and turbidity data for the 313 K storage condition. Data of the linear fit for the HMWS content as a function of time for samples stored at 313 K is summarised in Table S1. Large aggregates and turbidity data could not be fitted linearly as a function of time.

Table S1 Gradient, d HMWS/dt, and the coefficient of determination (R^2) of the linear fit ($y = mx + c$) for the HMWS content as a function of time. Measurements for the 13 weeks storage of F4 at 278 K samples were identified as outliers and omitted. Measurements for 39 weeks storage of F3 at 313 K were corrected for low recovery, while F3 samples at the 52 weeks at 313 K could not be injected into the SEC column.

Storage condition		d HMWS/dt	± SE	R^2
		% HMWS week^{-1}		
278 K	F1	0.0115	0.0008	0.9656
	F2	0.0095	0.0014	0.8539
	F3	0.0483	0.0104	0.7288
	F4	0.0082	0.0008	0.9495
298 K	F1	0.0497	0.0034	0.9480
	F2	0.0502	0.0055	0.8724
	F3	0.1178	0.0190	0.7617
	F4	0.0409	0.0046	0.8696
313 K	F1	0.15	0.0118	0.9323
	F2	0.29	0.0116	0.9810
	F3	1.91	0.1527	0.9397
	F4	0.15	0.0075	0.9713

Increasing glycerol content resulted in a non-linear increase in d HMWS/dt (Fig. S3). In this context it should be noted that there was little to no difference between the rate of increase in HMWS of formulations F1 and F4, containing no glycerol and 1.56 % (m/m) glycerol, respectively, despite the presence of a higher total amount of potentially plasticising species, i.e. water and glycerol[1,2], in formulation F4 as formulations F1, F2, F3 and F4 all had similar residual water contents.

Relaxation Dynamics: Differential Scanning Calorimetry (DSC)

Table S2 summarises the calorimetric glass transitions for the spray-dried mAb formulations, with samples being stored at 278 K before measurement. The mean $T_{g,\alpha,DSC}$ decreased with increasing glycerol content for the spray-dried mAb formulations. There was only a relatively small difference in mean $T_{g,\alpha,DSC}$ between formulations F1 and F4. Similarly, $T_{g,\alpha,DSC}$ decreased with increasing d HMWS/dt (Fig. S4), a trend reminiscent of the one observed in Fig. S3.

Relaxation Dynamics: Terahertz Time-Domain Spectroscopy (THz-TDS)

The changes in absorption at a frequency of 1 THz with temperature for spray-dried mAb formulations are plotted in Fig. S5. For all of the formulations, this change in absorption with temperature can be observed to take place over three distinct regions and two transition temperatures, $T_{g,\beta,THz}$ and $T_{g,\alpha,THz}$ (see Table 1).

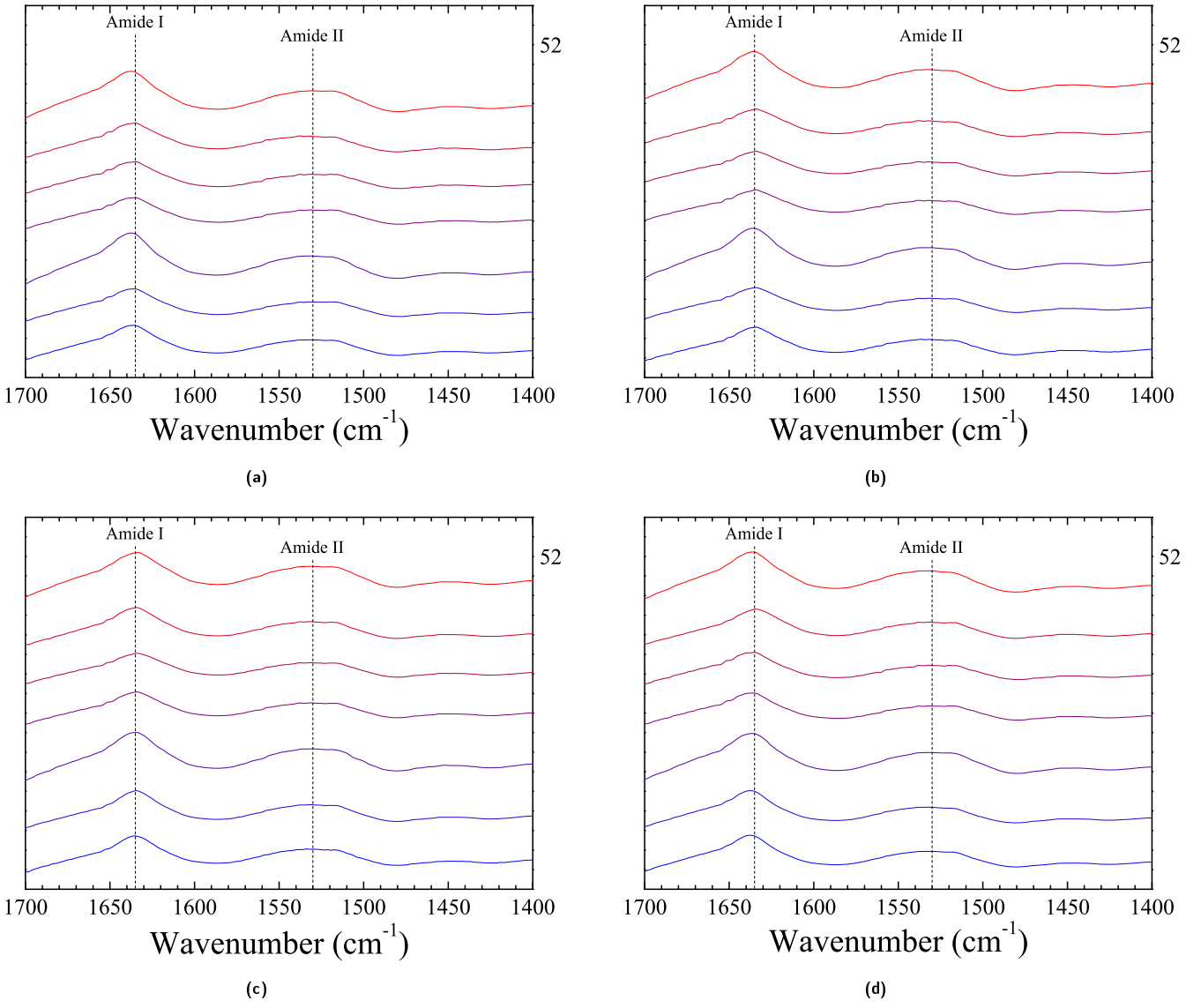

Fig. S1 FTIR spectra collected on solid state samples of formulations (a) F1, (b) F2, (c) F3 and (d) F4 stored at 313 K during 52 weeks stacked by Y Offsets to ensure that each spectrum can be viewed clearly. Dashed vertical lines at 1635 cm^{-1} and 1530 cm^{-1} serve as reference lines for the amide I (1700 cm^{-1} - 1600 cm^{-1}) and amide II absorption band (1580 cm^{-1} - 1520 cm^{-1}), respectively.

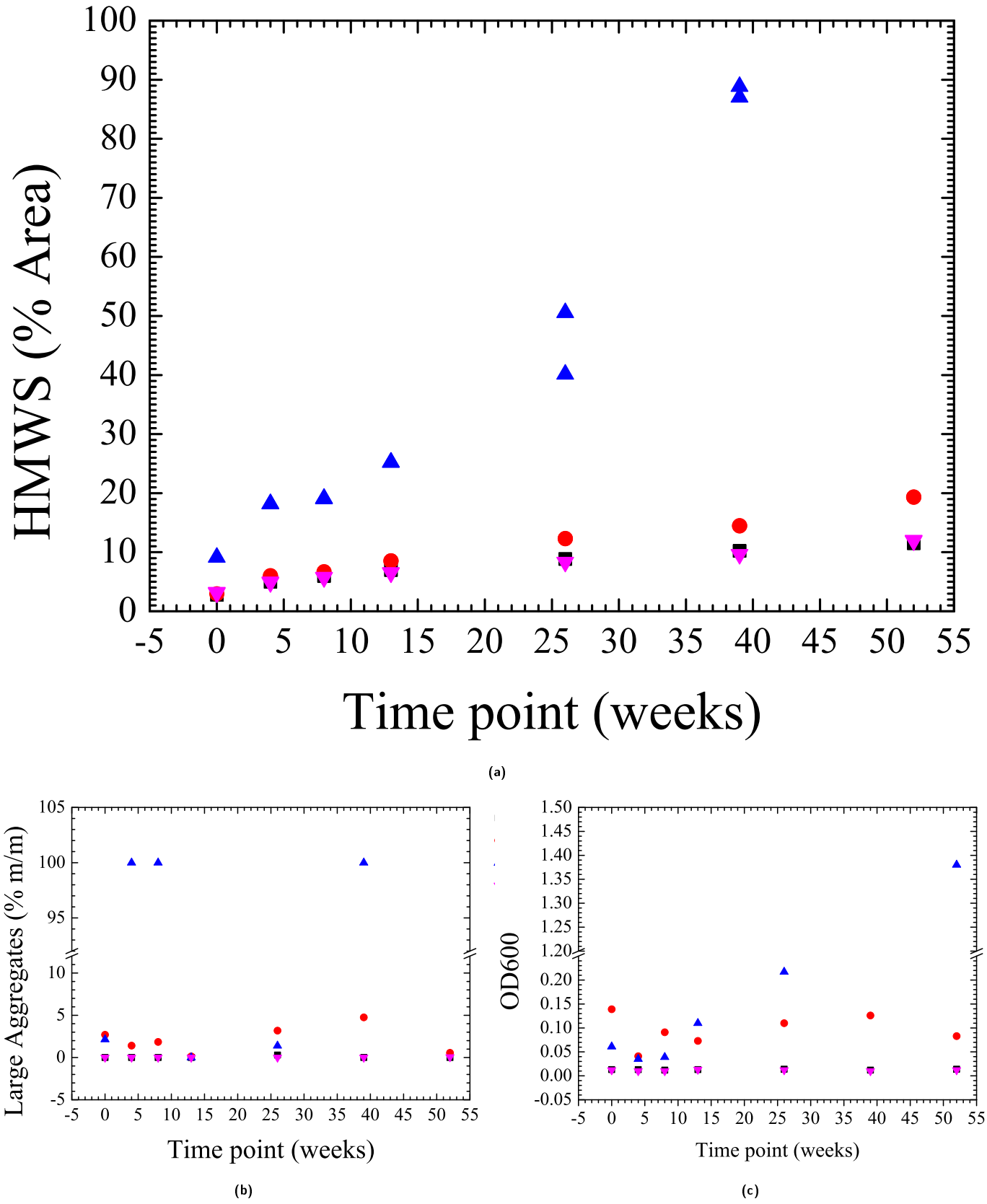

Fig. S2 (a) HMWS content based on size exclusion chromatography analysis, (b) large aggregate content based on dynamic light scattering analysis and (c) optical density at an incident wavelength of 600 nm as a function of time following reconstitution of samples stored in solid state at 313 K during 52 weeks.

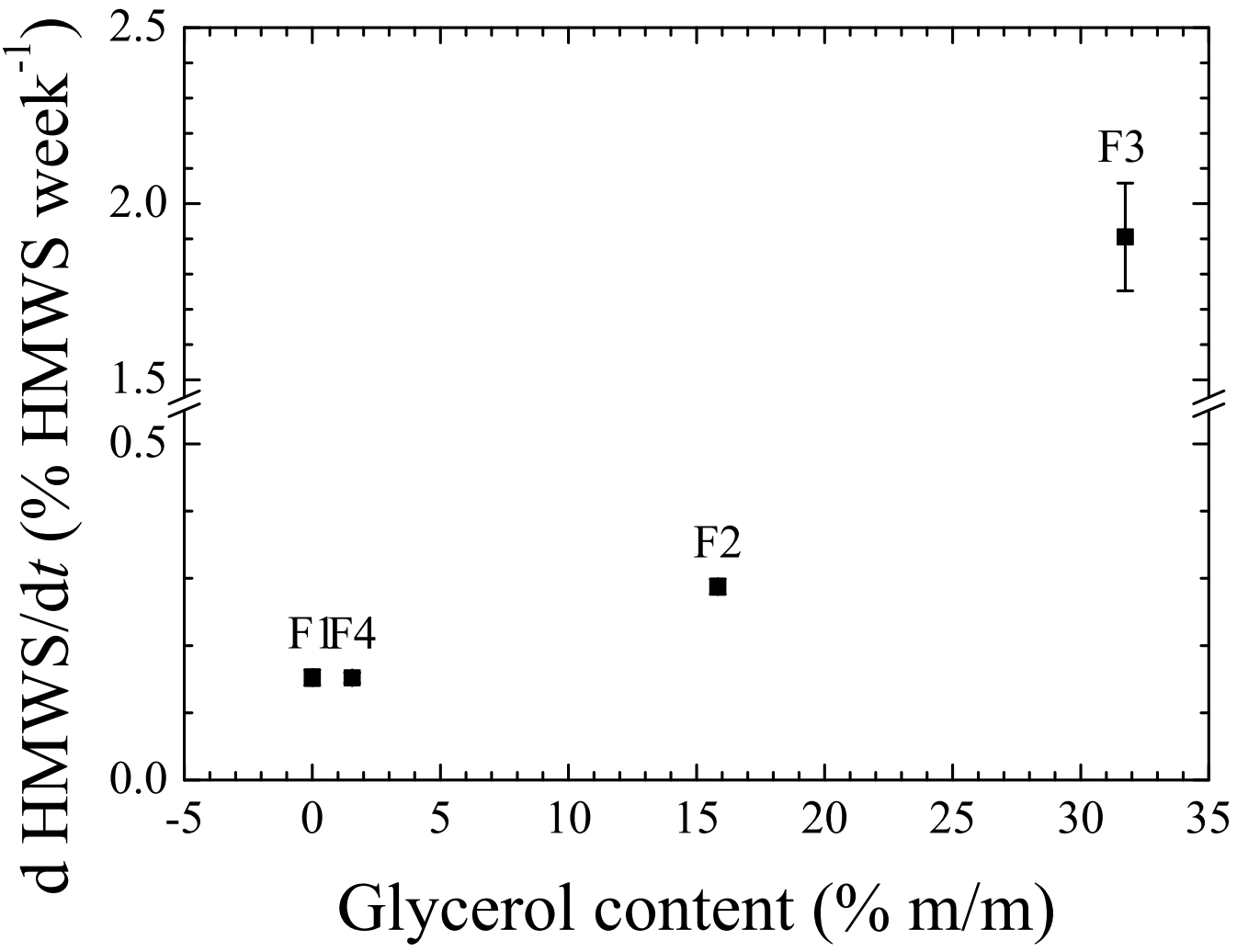

Fig. S3 Gradient, d HMWS/dt for samples stored at 313 K as a function of the glycerol content of formulations F1, F2, F3 and F4. Error bars represent the standard deviation for $n = 3$ samples.

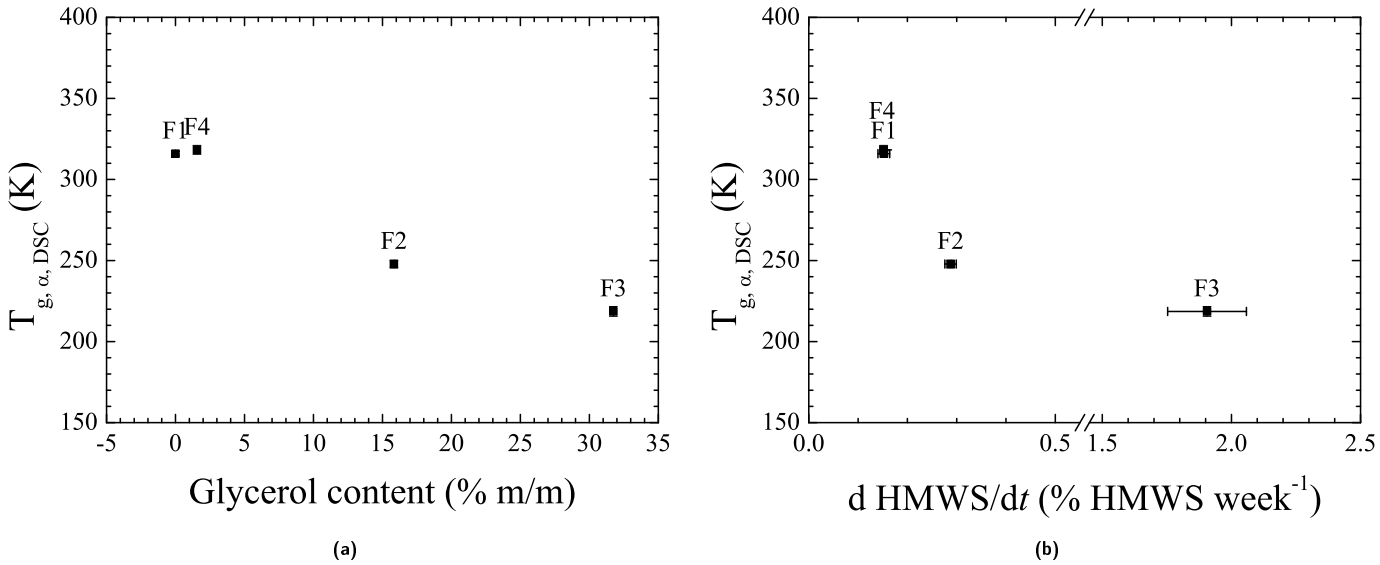

Fig. S4 (a) Mean (n=3) $T_{g,\alpha,DSC}$ as a function of the glycerol content and (b) d HMWS/dt for samples stored at 313 K of formulations F1, F2, F3 and F4. $T_{g,\alpha,DSC}$ values were determined on t_0 samples stored at 278 K. Horizontal error bars depict the standard error of the slope as a measure for the precision of the regression analysis. Vertical error bars depict the standard deviation for $n = 3$ samples.

Table S2 Mean ($n = 3$) $T_{g,\alpha,DSC}$ values determined on t_0 spray-dried mAb formulations stored at 278 K before measurement.

	$T_{g,\alpha,DSC}$	± SD
	K	
F1	316	1.4
F2	248	2.4
F3	219	3.0
F4	318	2.7

As the majority of motions are restricted/frozen below the $T_{g,\beta}$, dα_l/dT was not in scope for the present study since the focus was on the onset of motion.

All characteristic parameters of the linear fit based on terahertz analysis were clearly influenced by the composition of the spray-dried mAb formulations. A clear, positive correlation is present between the glycerol content and $T_{g,\beta,THz}$ (Fig. S6 (c)). No clear correlations could be observed with $T_{g,\alpha,THz}$ (Fig. S6 (d)).

Relaxation Dynamics: Dynamic Mechanical Analysis (DMA)

Fig. S7 depicts the normalised tan δ response for DMA measurements of each formulation. The repeat runs were found to be in good agreement with the results of the first run. Overall, two clear peaks can be observed for all formulations except F2, for which no clear tan δ α-transition peak could be defined. Peak deconvolution was found to generally result in a poor or no fit for the tan δ α-transition peaks, while good fits were obtained for the tan δ β-transition peaks, likely due to decreased sample stiffness with increasing glycerol content in the temperature region preceding the glass transition. As α-transition peaks could not be resolved in a satisfactory manner, only tan δ β-transition peaks were considered for further interpretation. A summary of the characteristic parameters of the fitted mean ($n = 2$) tan δ β-transition peaks can be found in Table 2. We would like to highlight that $T_{g,\beta,DMA}$ was defined as the temperature of the peak maximum, i.e. the transition midpoint, where $T_{g,\beta,THz}$ was defined as the transition onset.

The amplitude and area of the fitted mean tan δ β-transition peak showed a clear, positive correlation with the glycerol content (Fig. 2) and d HMWS/dt (Fig. S8) of the spray-dried mAb formulations. No clear trends were observed between the glycerol content and d HMWS/dt and $T_{g,\beta,DMA}$.

Finally, as can be seen in Fig. S7, we would like to remark that the amplitude of $T_{g,\beta,DMA}$ and $T_{g,\alpha,DMA}$ peaks was generally in the same order of magnitude for a single formulation, while DMA thermograms for small molecule amorphous solids, e.g. those reported for indomethacin by Kissi et al.[3], show $T_{g,\alpha,DMA}$ peak amplitudes several orders of magnitude larger than those of $T_{g,\beta,DMA}$. The overall low amplitude of the $T_{g,\alpha,DMA}$ peak for spray-dried mAb formulations is in agreement with reports of hard to determine calorimetric glass transitions for pure (dry) protein formulations, as these $T_{g,\alpha,DSC}$ generally are very broad and display only a small change in heat capacity during the transition[4].

Notes and references

1 T. Chen, A. Fowler and M. Toner, *Cryobiology*, 2000, **40**, 277–282.

2 M. T. Cicerone and C. L. Soles, *Biophysical Journal*, 2004, **86**, 3836–45.

3 E. O. Kissi, H. Grohganz, K. Löbmann, M. T. Ruggiero, J. A. Zeitler and T. Rades, *The Journal of Physical Chemistry B*, 2018, **122**, 2803–2808.

4 M. Mizuno and M. J. Pikal, *European Journal of Pharmaceutics and Biopharmaceutics*, 2013, **85**, 170–176.

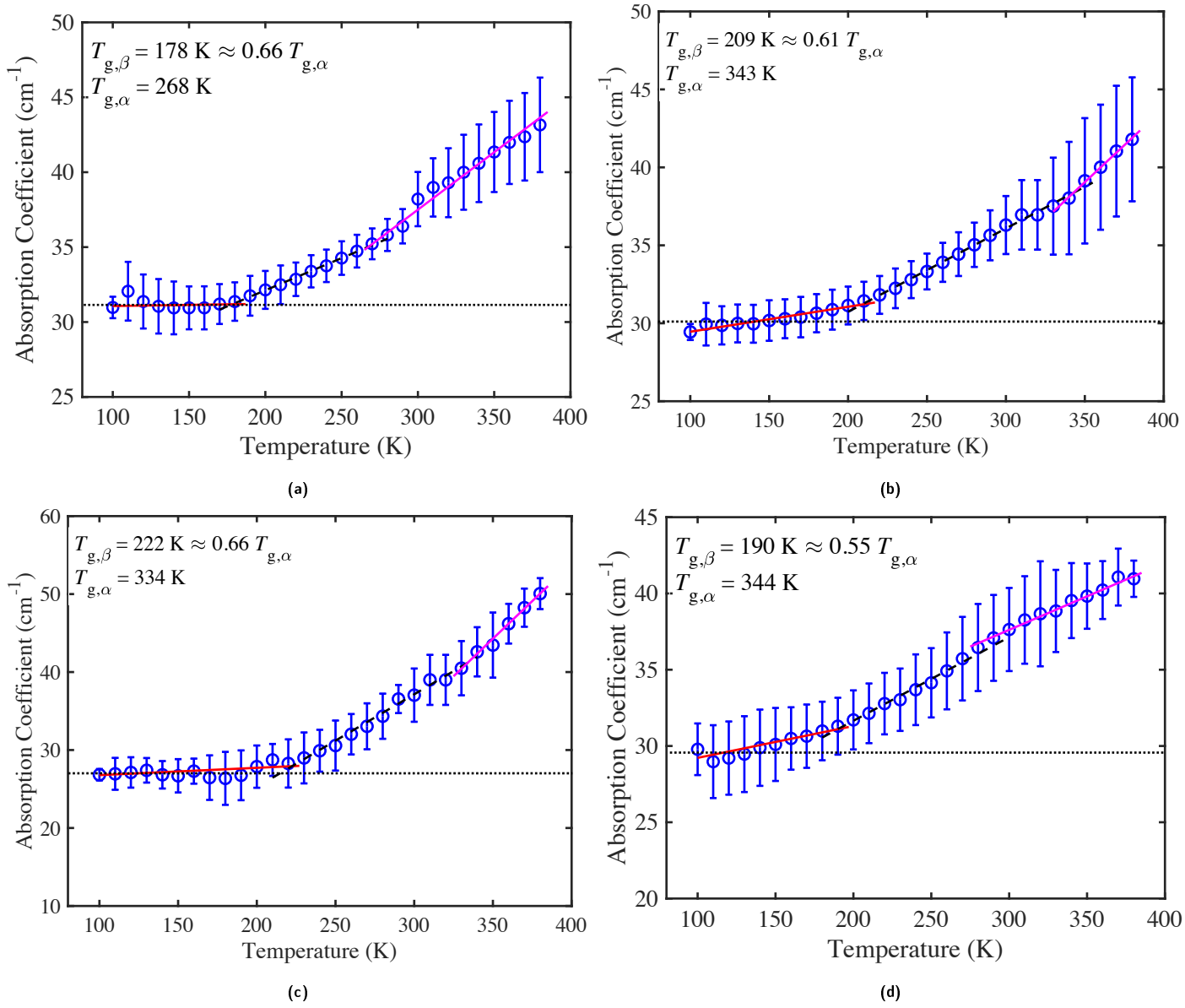

Fig. S5 Mean terahertz absorption coefficient as a function of temperature at 1 THz for mAb formulations (a) F1, (b) F2, (c) F3 and (d) F4. Solid red line, dashed black line, and solid violet line show the different linear fits for the three different regions. Error bars represent the standard deviation for $n = 3$ samples. Dotted black line shows the average value of the absorption coefficient between 100 and 150 K, reflecting an estimate for the boundary between harmonic (below the line) and non-harmonic/excess motion (above the line).

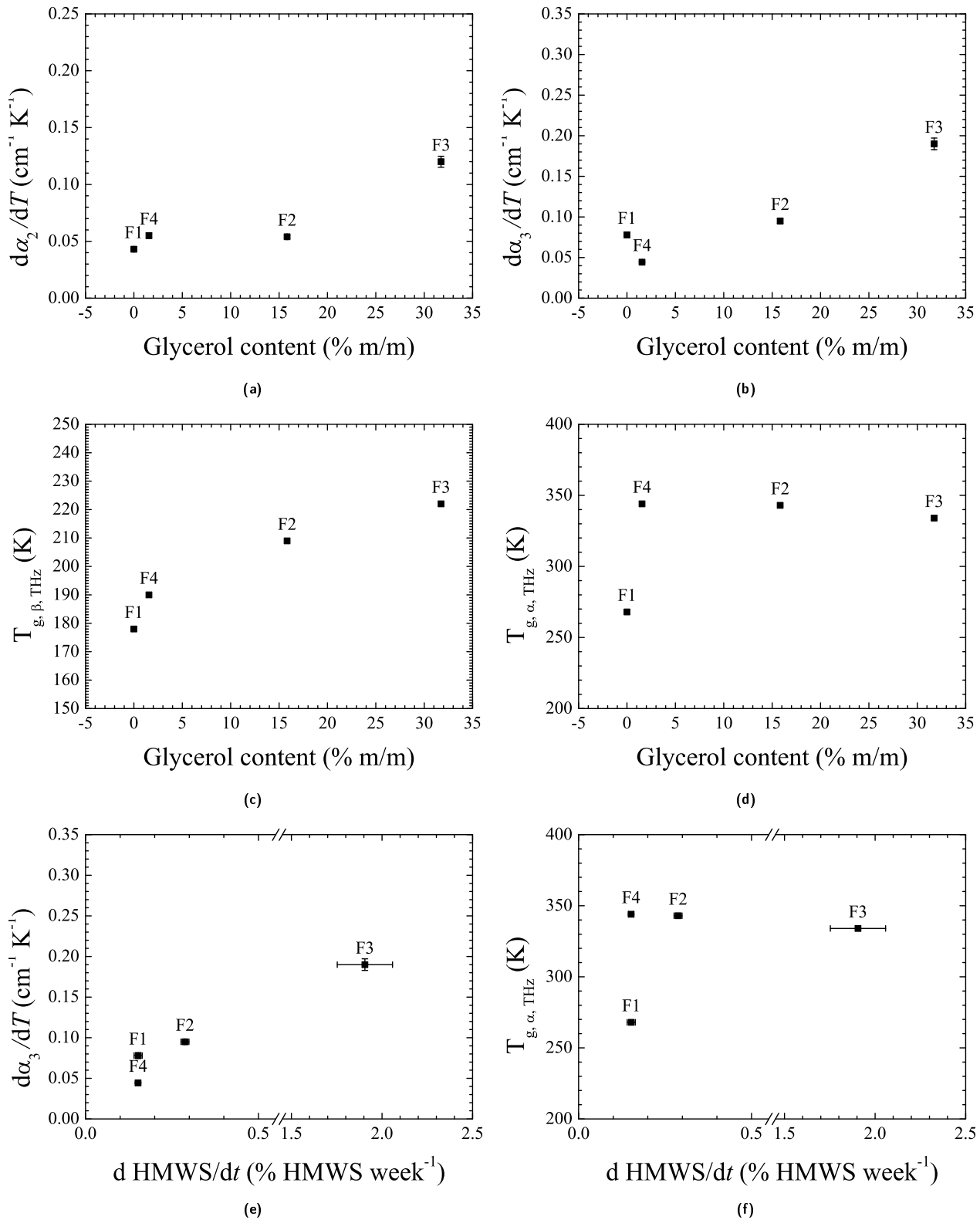

Fig. S6 Based on the terahertz analysis of the spray-dried mAb formulations, three distinct linear regions could be identified for the absorption coefficient α as a function of temperature (T). The gradient of each distinct linear region is denoted as $d\alpha_1/dT$, $d\alpha_2/dT$ and $d\alpha_3/dT$ representing regions 1, 2, and 3, respectively. The infliction points between regions 1 and 2 and regions 2 and 3 are labelled as $T_{g,\beta,THz}$ and $T_{g,\alpha,THz}$, respectively. (a) $d\alpha_2/dT$, (b) $d\alpha_3/dT$, (c) $T_{g,\beta,THz}$ and (d) $T_{g,\alpha,THz}$ as a function of the glycerol content of formulations F1, F2, F3 and F4, respectively, as well as (e) $d\alpha_3/dT$ and (f) $T_{g,\alpha,THz}$ as a function of d HMWS/dt for samples stored at 313 K. Vertical Error bars represent the standard deviation for $n = 3$ samples. Horizontal error bars depict the standard error of the slope as a measure for the precision of the regression analysis.

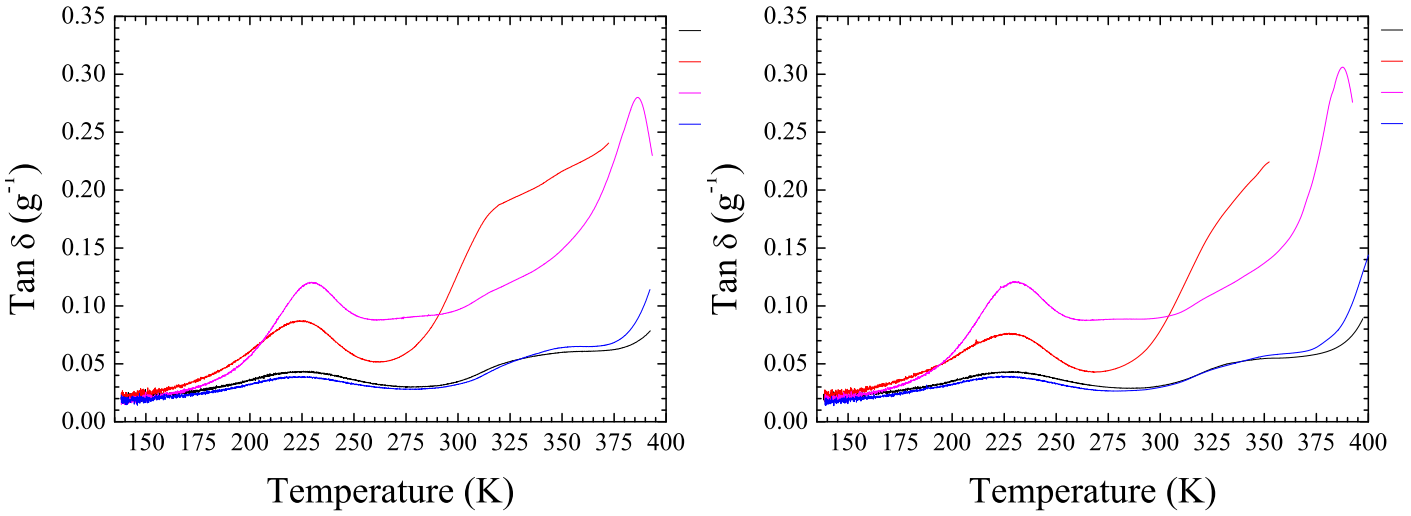

Fig. S7 Overlay of the $\tan\delta$ responses corrected for sample mass as a function of temperature for all spray-dried mAb formulations based on (a) the first and (b) second DMA run of each formulation.

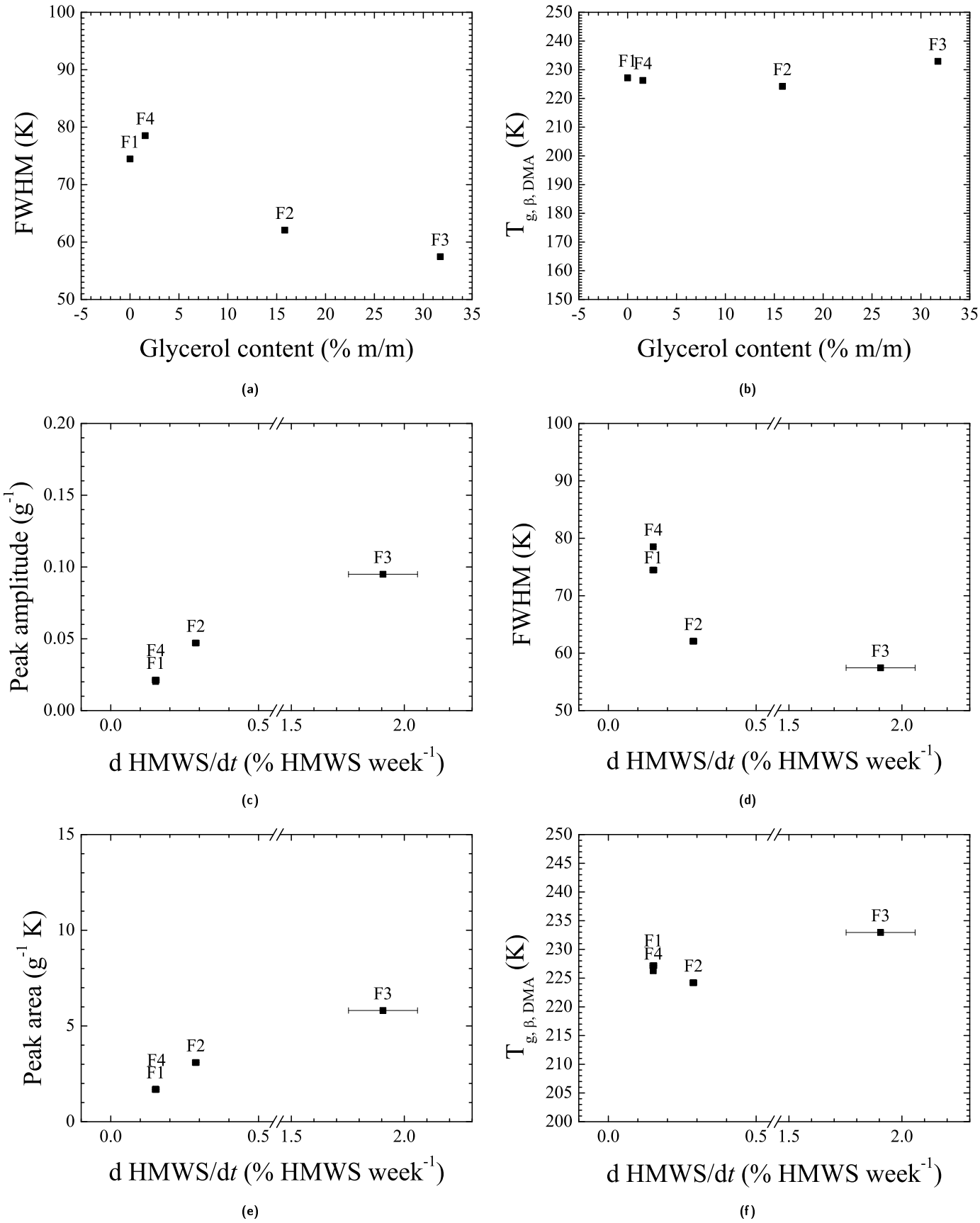

Fig. S8 (a) The full width at half maximum (FWHM) and (b) $T_{g,\beta,DMA}$ of the fitted mean ($n = 2$) $\tan\delta$ β-transition peak based on DMA as a function of the glycerol content, as well as the $\tan\delta$ β-transition peak's (c) amplitude, (d) FWHM, (e) area and (f) $T_{g,\beta,DMA}$ as a function of d HMWS/dt for samples stored at 313 K. Error bars depict the standard error of the slope as a measure for the precision of the regression analysis.