A Robust Spectroscopic Method for the Determination of Protein Conformational Composition- Application to the Annealing of Silk.

David J. Belton[1], Robyn Plowright[1], David L. Kaplan[2] and Carole C. Perry[1]*

[1] Interdisciplinary Biomedical Research Centre, Nottingham Trent University, Clifton Lane, Nottingham NG11 8NS

[2] Department of Biomedical Engineering, Tufts University, Medford, MA, 02155, USA

Corresponding author: carole.perry@ntu.ac.uk

Supplementary data.

Glossary:

Beta structure definitions –

Beta (β) aggregates – Crystalline β domains characterised by amide I band towards the low frequency end of the envelope (centred at 1615cm^{-1}) and a high frequency signal (centred at ~ 1690cm^{-1}) also termed intermolecular β sheet.

Beta (β) sheet – Non-crystalline β domains characterised by amide I signal at the upper range of the β sheet range (centred at 1628cm^{-1}) also termed intramolecular β sheet.

Beta (β) structure, nature or characteristic – all terms indicative of total beta contribution to the amide I band envelope (sheet, turn, and aggregates).

Beta (β) helix - 2 or 3 faced hydrogen bond stabilised helical stack of β strands ~ 1690cm^{-1}

Fourier self-deconvolution

This is generally accepted as the starting point for peak fitting procedures in the process of peak selection but is the process most susceptible to noise. For the purpose of this study, a gamma factor of 4.0 was chosen to enhance the peak resolution and this required a smoothing factor of 60% to reduce the noise generated to acceptable levels. Subsequent peak fitting was carried out at high sensitivity with an initial peak width of 30 cm^{-1} FWHH. A single convergence measurement was carried out to produce a fitted spectrum with minimal changes to the other parameters (position, width and shape) and statistically analysed. This was compared with the results of multiple repeats of the convergence process until equilibrium was reached. Although the agreement of the experimental and fitted spectral envelopes was excellent it was clear that the process was preferentially fitting the residual noise and consequently the reproducibility in terms of conformational output was very poor (figures 1a, b, e, f). In addition, the peak widths particularly at equilibrium did not represent those, which could be reasonably expected according to the theory or literature sources. It was clear that the poor replication was at least in part due to the poorly controlled peak widths so a second set of

measurement were made with the widths more reasonably limited to between 10 and 30 cm^{-1} (figures 1 c, d, e, f). The statistical outcome from this modified process was improved, particularly at equilibrium, but the reproducibility between replicates was still poor in terms of the pooled standard deviation for all the conformers. It was clear that there was an issue of peak pinning due to the residual noise even after a smoothing factor of 60% had been applied, and it was for this reason that we proceeded to search for solutions to the reproducibility problem without the use of Fourier self-deconvolution and automatic peak fitting.

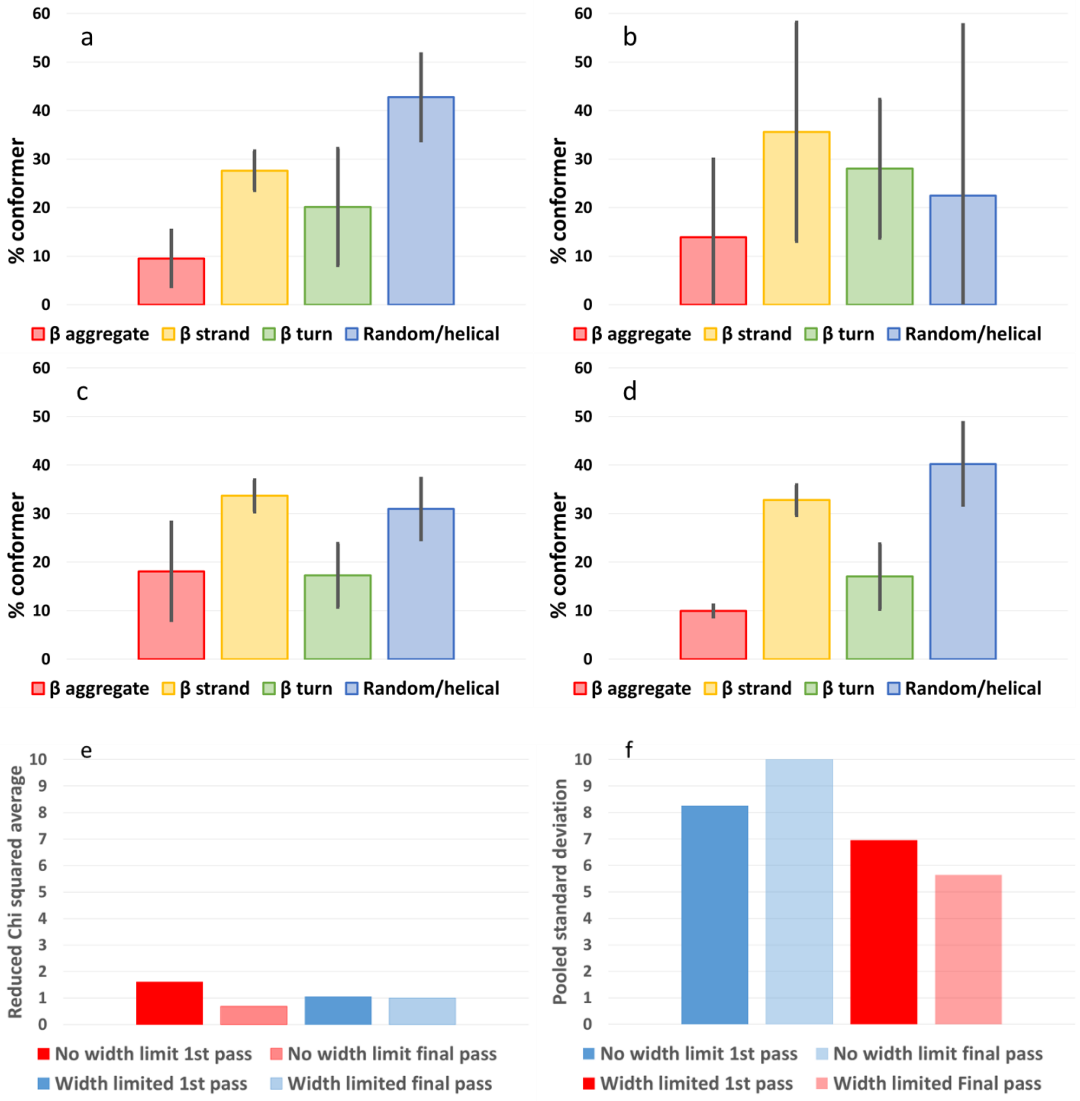


Figure S1. Conformational analysis using Fourier self-deconvolution with and without peak width limitation. No limits on peak width, first pass a), equilibrium b), peak widths limited from 10 cm^{-1} to 30 cm^{-1}, first pass c), equilibrium d), effect of parameters on goodness of fit (reduced chi squared) e), and pooled standard deviation f).

	β content (%)	**Non-β content (%)**	**Methods and [references]**	**β-content (%)**	**Non-β content (%)**
BSA	20-47[2*]	53-80[2*]	FTIR/ CD [1-3]	40[*]	60[*]

α lactalbumin	38-44 18-60[2*]	56-62 40-82[2*]	CD/Thioflavin stain/ FTIR [4,5]	45[*]	55[*]
γ globulin	55-75	25-45	FTIR/DSC [6-8]	57[*]	43[*]

Table S1. Comparison of quantitative data found in this study and by others using a range of quantitation methods. Published conformational compositions of reference proteins used in this study determined by others. * This contribution. [2*] BSA and α lactalbumin dependent on treatment (native/ molten globule / fibrillar, low to high β content)

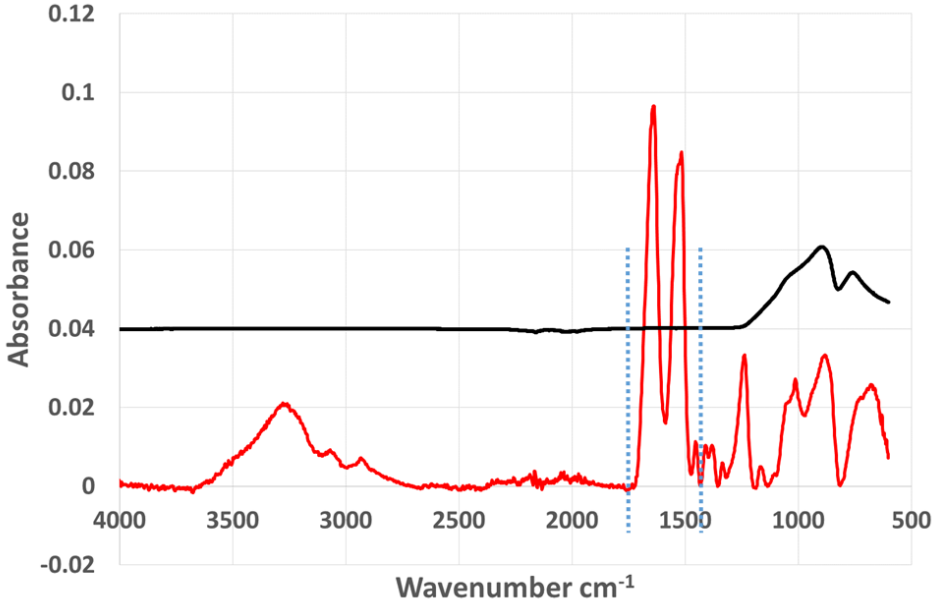


Figure S2. Full infrared spectrum of cast silk film with superimposed glass substrate spectrum showing absence of interfering absorbance over the region of analysis (1730 – 1470 cm^{-1}).

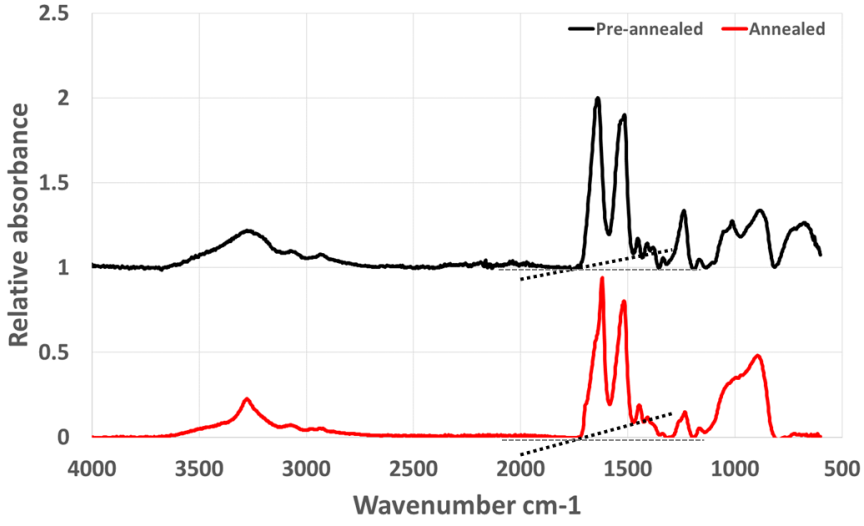


Figure S3. Comparison of conformational change pre and post annealing on baseline position.

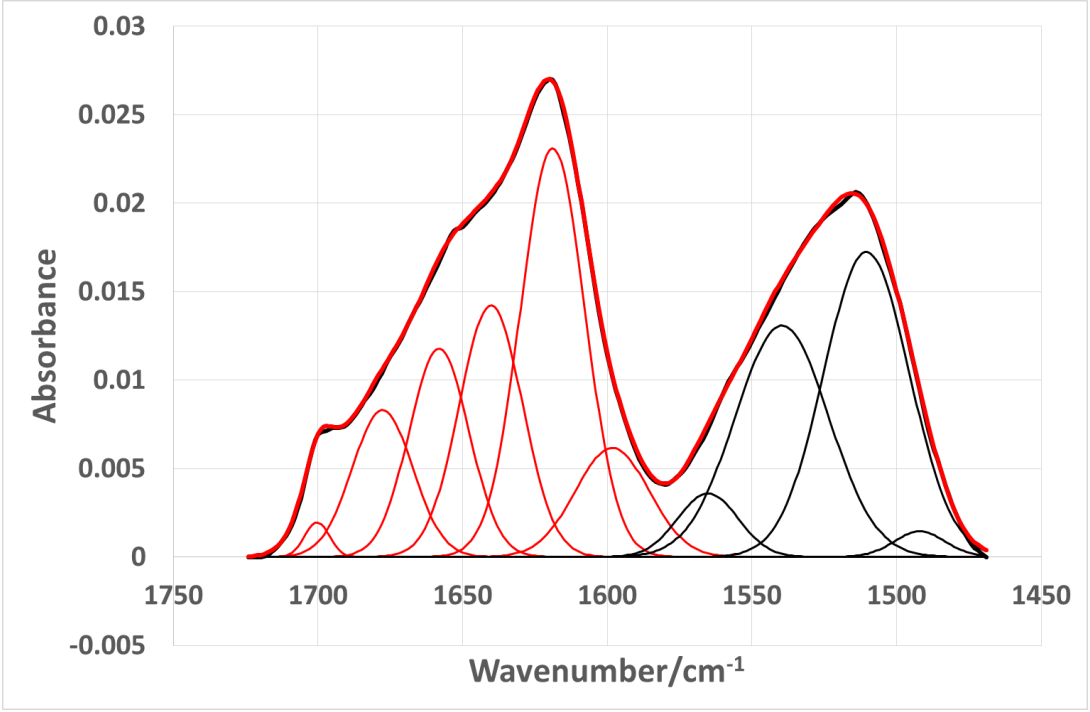


Figure S4. Full peak fitting of the amide I and II bands showing peaks omitted from the figure in the paper (shown in black).

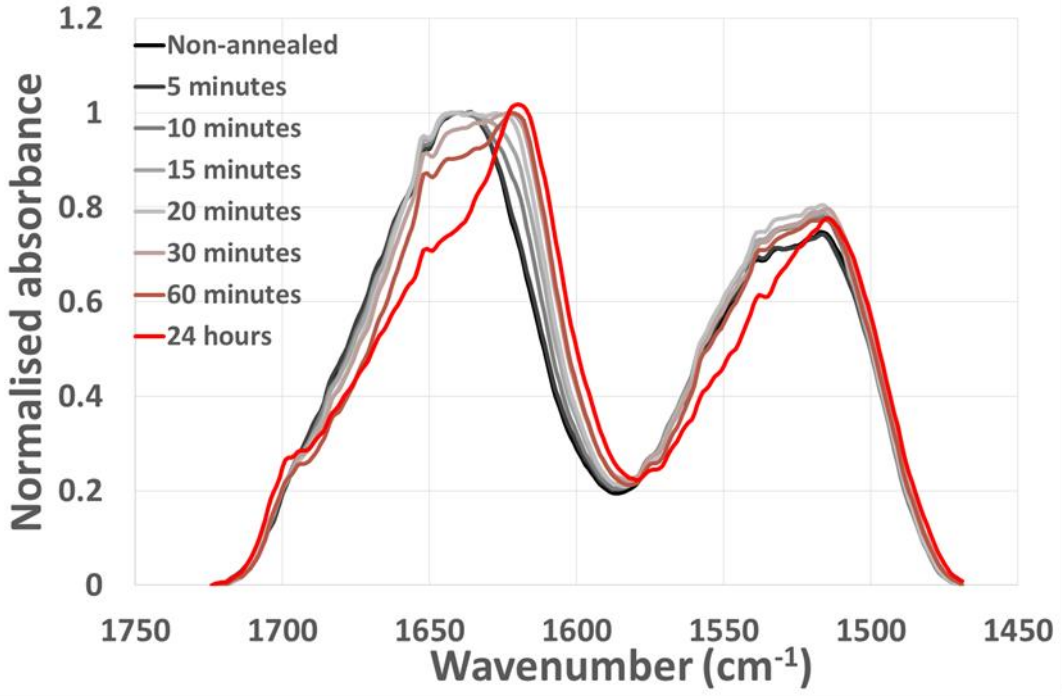


Figure S5. Spectral changes observed during water/methanol vapour annealing of regenerated silk films.

Calculation of amide I carbonyl band energy change.

Wavenumber is frequency in cm^{-1}

Convert wavenumber to frequency in Hz using:

$$f = \frac{C}{\lambda}$$

Convert frequency to energy in joule using

$$E = hf$$

This is the energy of a single photon of a given frequency

So energy change per photon

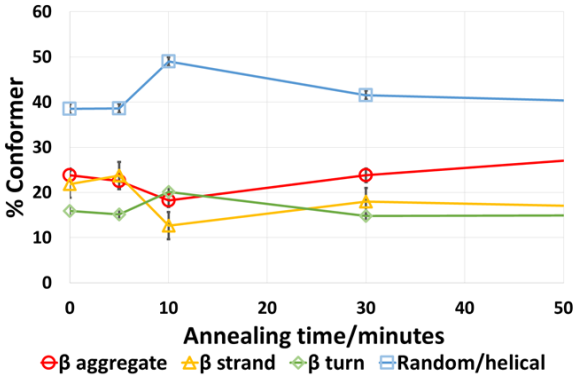
$\Delta E = h\Delta f = h(f_{ref}\text{-}f_1)$ in joule per photon

f_1 = frequency of fitted amide I carbonyl peak

f_{ref} = frequency of reference peak

If we take the non hydrogen bonded condition as the reference frequency (solutions in DMSO amide I ~ $1662cm^{-1}$ = f_{ref}) we can compare the relative hydrogen bond strength of the other conformers and correlate it with physical properties.

The overall energy change per photon is the sum of the individual energy changes multipiled by their relative intensity. Multiplying by the Avogadro constant gives the energy change per mole of carbonyl and further multiplying by the number of amino acids residues in the sequence gives it as change per mole of protein.



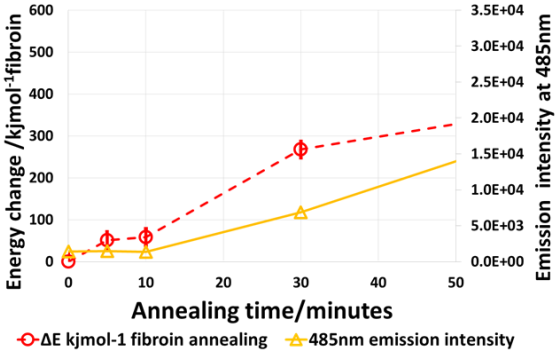


Figure S6. Conformational and energy change during the first 50 minutes of annealing.

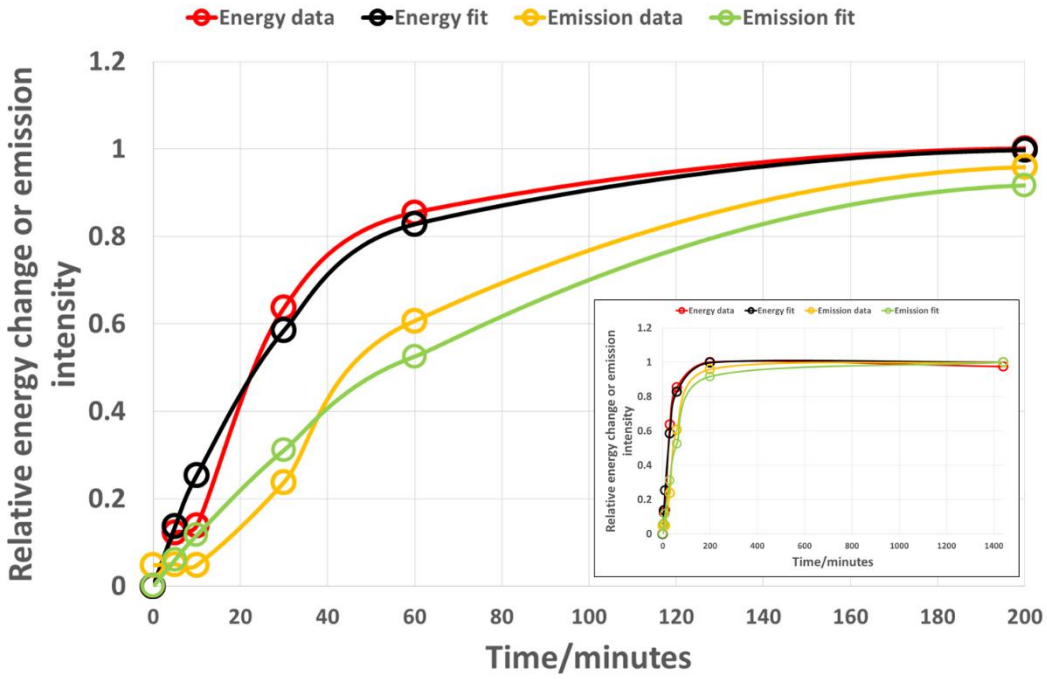


Figure S7. Exponential fitting of the growth curves for energy change and Thioflavin T emission during the annealing of silk fibroin over the first 200 minutes. Insert covers the process over 24 hours.

References for supplemental data

[1] Abrosimova, K. V,; Shulenina, O. V,; Paston, S. V 'FTIR study of secondary structure of bovine serum albumin and ovalbumin', *Journal of Physics: Conference Series*, **2016,** 769, 1-6.

[2] Reed, R. G.; Feldhoff, R. C.; Clute, O. L.; Peters, T. Fragments of bovine serum albumin produced by limited proteolysis. Conformation and ligand binding, *Biochemistry*. American Chemical Society, **1975,** 14(21), 4578–4583.

[3] Holm, N. K.; Jespersen, S. K.; Thomassen, L. V.; Wolff, T. Y.; Sehgal, P.; Thomsen, L. A.; Otzen, D. E. Aggregation and fibrillation of bovine serum albumin, *Biochimica et Biophysica Acta***, 2007,** *1774*, 1128–1138.

[4] Otte, J.; Ipsen, R.; Bauer, R.; Bjerrum, M. J.; Waninge, R. Formation of amyloid-like fibrils upon limited proteolysis of bovine a-lactalbumin, *International Dairy Journal*, **2005,** 15(3), 219–229.

[5] Goers, J.; Permyakov, S. E.; Permyakov, E. A.; Uversky, V. N.; Fink, A. L. Conformational Prerequisites for α-Lactalbumin Fibrillation, *Biochemistry*. American Chemical Society, **2002,** 41, 12546–12551.

[6] Giacomelli, C.E.; Bremer, M. G. E. G.; Norde, W.; ATR-FTIR Study of IgG Adsorbed on Different Silica Surfaces. *Journal of Colloid and Interface Science*, **1999**, *220(1)*, 13–23.

[7] Nag, M.; Das, D.; Bandyopadhyay, D.; Basak, S. Unusual denaturation trajectory of bovine gamma globulin studied by fluorescence correlation spectroscopy. *Physical Chemistry Chemical Physics*, **2015,** *17(29)*, 19139–19148.

[8] Choi, S.; Ma. C. Conformational Study of Globulin from Common Buckwheat (*Fagopyrum Esculentum Moench*) by Fourier Transform Infrared Spectroscopy and Differential Scanning Calorimetry. *Journal of Agricultural and Food Chemistry* **2005** *53(20)* 8046–5