Electronic Supplementary Information

# Structure-energy relations in hen egg white lysozyme observed during refolding from a quenched unfolded state

**Theresa Y. Cho[a], Nolene Byrne[b], David J. Moore[c], Brian A. Pethica[a], C. Austen Angell[b] and Pablo G. Debenedetti[a]**

*a Department of Chemical Engineering, Princeton University, Princeton, NJ 08544, USA. E-mail: pdebene@princeton.edu;Fax: +1 609 258 0211; Tel: +1 609 258 5480*

*b Department of Chemistry and Biochemistry, Arizona State University,* 60 *Tempe, AZ 85287, USA.*

c *International Specialty Products, 1361 Alps Road, Wayne, New Jersey 07470*

## Experimental details

### *Materials*

Hen egg white lysozyme (L-6876) was purchased from Sigma Chemical Co. (St. Louis, USA) and was used without further purification. Sucrose (S-8501) was also purchased from Sigma Chemical Co. Ethylammonium nitrate (EAN) was synthesized by adding nitric acid (87920) to an equimolar amount of ethylamine (A15022) in aqueous solution, both purchased from Alfa Aesar. The mixture was stirred at room temperature for several hours. To ensure a complete reaction, a slight excess of amine was left over and was removed along with the water by heating at 80$^{o}$C in vacuum using a rotary evaporator. The product was then dried at 80$^{o}$C for two days in a vacuum oven containing $P_2O_5$ to remove any excess water. Since these reactions are very exothermic, the dropwise addition of the acid to the amine was carried out while cooling the amine solution to -78$^{o}$C using an acetone/dry-ice bath.

### *Differential Scanning Calorimetry (DSC)*

DSC experiments were conducted to determine the reversibility of lysozyme folding – unfolding in varying solvent concentrations of EAN, sucrose and $H_2O/D_2O$ using a Perkin-Elmer Pyris 1 differential scanning calorimeter equipped with an Intracooler 2P cooling accessory. Indium was used to calibrate the enthalpy of the DSC instrument, while both indium and cyclohexane were used to calibrate for the temperature. An empty sealed stainless sealed pan was used as a reference. Approximately 20 μL of a sample solution were placed in a stainless steel pan. All DSC thermal scans started at 25$^{o}$C and heated to temperatures above the sample's denaturing temperature, $T_d$, which was as high as 98$^{o}$C. This is considered as "scan 1". The samples were then cooled back to room temperature, and reheated. The reheated scan is considered as "scan 2". Both heating and cooling scans were performed at scan rates of 20$^{o}$C/min. All DSC scan results were normalized by the sample mass, and the change in enthalpy between scan 1 and scan 2 provides a quantitative look at protein aggregation. The data were analyzed using Microcal$^{TM}$ Origin$^{TM}$, version 2.1.

*Infrared (IR) Spectroscopy*

IR spectra were recorded at various temperatures using a Nicolet 6700 FTIR spectrometer equipped with a broad-band liquid nitrogen-cooled mercury cadmium telluride (MCT) detector and a Linkam FTIR600 Stage with liquid nitrogen ($LN_2$) cooling and CI94 temperature control units. Each spectrum was generated by a co-addition of 64 interferograms collected at a 4 $cm^{-1}$ resolution. The spectrometer was also under a continuous dry air purge. Samples of approximately 6 µl were sandwiched between two $AgCl_2$ windows. For quench studies, the sandwiched windows were first placed in an oven at 90°C for 10 minutes to unfold lysozyme. Then the sample was quickly quenched in a cup of $LN_2$. After quench, the sample was placed in the temperature-controlled cell at -135°C and subsequently heated at 2°C/min with 2 minute anneals every 10°C to collect spectra. To account for possible changes in spectra due to thermal effects on bond vibration, reference spectra under identical scan conditions with only the solvent present were also recorded in the same cell. Protein IR spectra of the liquid samples were corrected by subtracting the contribution of the solvent at each temperature.

For analysis, background spectra were subtracted at each temperature to obtain the lysozyme spectra. The subtracted spectra were then baseline-corrected and normalized to unit area. Quantitative spectral analysis was performed in two steps. First, second derivative spectra were obtained using a 2nd degree, 11-point Savitsky-Golay derivative function in Grams software. This data provided the number of peaks and peak positions needed to fit the original spectra with Gaussian peaks. In the second step, the peak number and positions were fixed, while peak intensities and widths were allowed to vary to obtain the best fit. Peak areas and positions were then used to provide a quantitative estimate of relative weights assigned to specific secondary structural elements.

**Sucrose prevents crystallization**

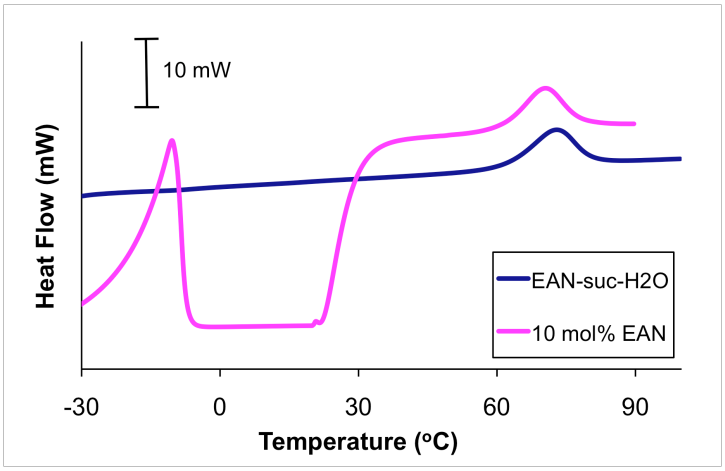


**Figure S1**: DSC scans show that the addition of sucrose prevents ice crystallization by the absence of a large endothermic peak below 0°C in the EAN-sucrose-$H_2O$ solution scan.

**Quantitative IR analysis of native lysozyme secondary structure in EAN/Suc/$D_2O$**

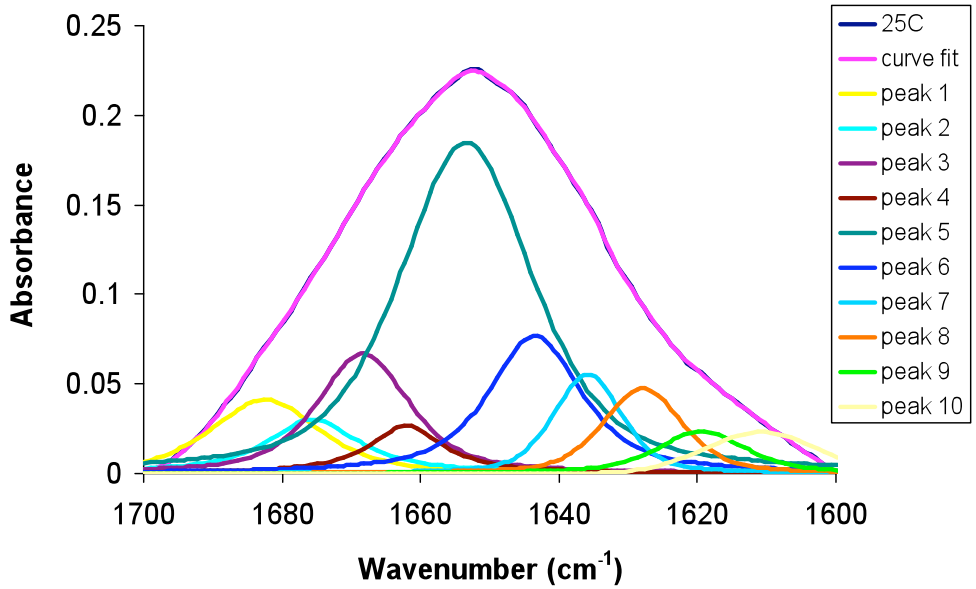


**Figure S2**:IR spectrum of lysozyme at 25°C (2 wt% lysozyme, 39 wt% EAN, 33 wt% sucrose, 25 wt% $D_2O$), after background subtraction and baseline correction. Also shown are the individual Gaussian peaks used for quantitative analysis, and the resulting fit, which is indistinguishable from the original spectrum.

**Table S1:** Quantitative analysis of secondary structure

| | X-ray[1] Secondary Structure % | FTIR[2] Secondary Structure % (Literature Values) | FTIR Peak Area (Experimental) |
|---|---|---|---|
| α-Helix | 45 | 40 | 42 |
| β-Sheet | 19 | 19 | 19 |
| turn | 23 | 27 | 26 |
| random | 13 | 14 | 13 |

The peak positions identified by the second derivative analysis were 1682, 1675, 1668, 1662, 1653, 1644, 1635, 1627, 1619 and 1610 $cm^{-1}$. Lysozyme is predominantly α–helical, as evident by the large peak at 1653 $cm^{-1}$. The higher wavenumbers above 1653 $cm^{-1}$ are assigned to β–turns, while the peak at 1644 $cm^{-1}$ is related to random structures. The lower wavenumbers are associated with β–sheet structures.

We note that some amino acid side-chains absorb in the Amide I region. However, such side-chain modes are not directly sensitive to protein backbone conformation, and normally give rise to a very weak signal compared to the backbone carbonyl mode (Amide I). Many of these modes, furthermore, are shifted out of the 1600 - 1700 $cm^{-1}$ region in $D_2O$. For example, the band positions ($cm^{-1}$) for ASN, GLN, ARG, and LYS in $D_2O$ are[3]: ASN [1648 (CO stretch)], GLN [1635-1654 (CO stretch); 1163 ($NH_2$ in-plane bending); 1409 (CN stretch)]; ARG [1608 ($CN_3H_5^+$ asymmetric stretch); 1586 ($CN_3H_5^+$ symmetric stretch)]; LYS [1201 ($NH_3^+$ asymmetric in-plane bending); 1170 ($NH_3^+$ symmetric in-plane bending)].

## Quenched spectra of folded lysozyme

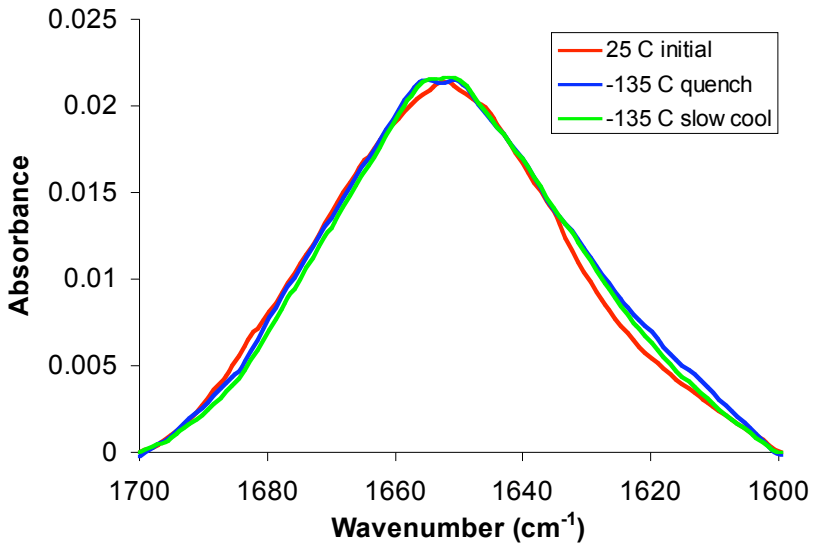


**Figure S3**: Spectra of folded lysozyme in 40 wt% EAN, 33 wt% sucrose and 27 wt% $D_2O$ at 25°C before any thermal treatment, and at -135°C after slow-cooling (20°C/min) and quenching in liquid nitrogen. As can be seen, the IR spectra are very similar.

**Evolution of α-helix signature during upscan through the unfolding temperature**

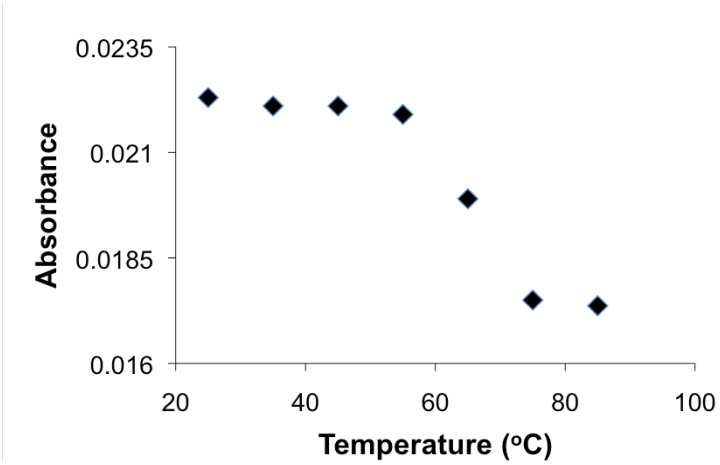


**Figure S4**: Changes in absorbance level at 1650 $cm^{-1}$ (α-helix) for thermal unfolding of lysozyme in 40 wt% EAN, 33 wt% sucrose and 27 wt% $D_2O$. Note that the unfolding temperature range seen in the IR is similar to that found calorimetrically (Fig. 1).

**IR evidence of prevention of aggregation in EAN/Sucrose/$D_2O$**

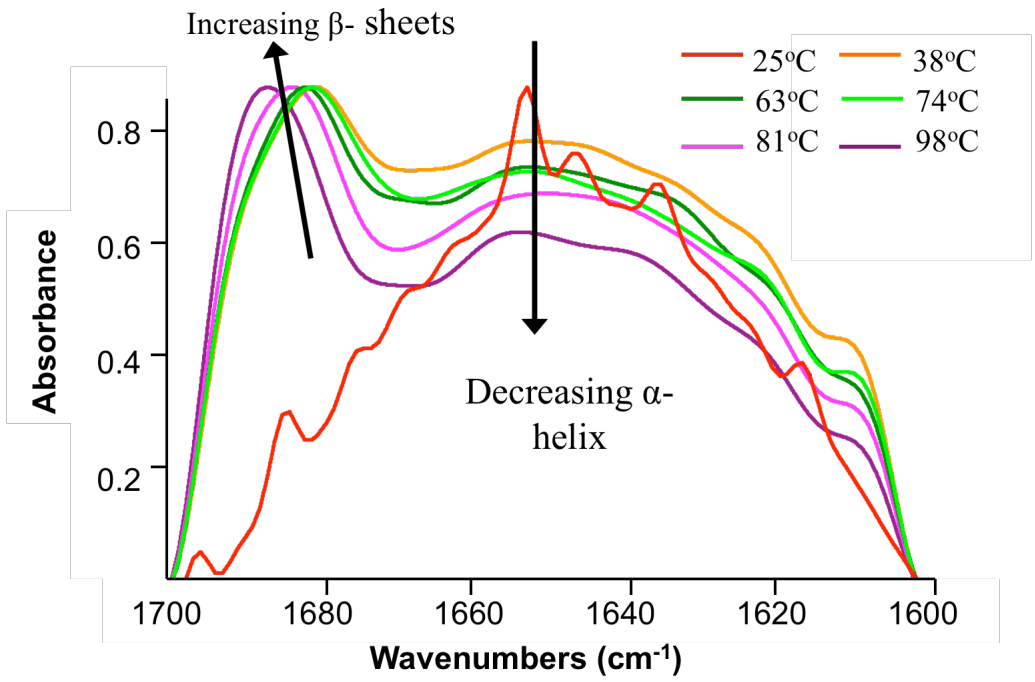


**Figure S5**: IR spectra of lysozyme in 80 wt% Trehalose and 20 wt% $D_2O$, heated from 25°C to 98°C. Irreversible aggregation is evident in the appearance of a large peak around 1680 $cm^{-1}$, which is not evident in the unfolded spectra of lysozyme in 40 wt% EAN, 33 wt% sucrose and 27 wt% $D_2O$ (Fig. 3).

## References


1. A. Dong, Huang P. and Caughey W.S., *Biochemistry*, 1990, **29**, 182.

2. S.Q. Luo, Huang C. Y. F., Mcclelland J. F. and Graves D. J., *Anal. Biochem*, 1994, **216**, 67.

3. A. Barth, *Prog. Biophys. & Molec. Biol.*, 2000, **74**, 141.