# Local cooperativity in the amyloidogenic state of human lysozyme observed at atomic resolution

## Supporting Information

Anne Dhulesia *et al.*

### Fitting of Circular Dichroism data

The raw data was first normalized so that values of the beginning and end of the unfolding transition would be near 0 and 1, respectively (see Fig. S3).

#### Fitting to a two-state model

The data was analysed using a two-state model (native state N and denatured state D):

$$\mathrm{N} \rightleftharpoons \mathrm{D} \tag{1}$$

The normalized spectroscopic signal $S$ for far- or near-UV CD was fitted to equation 2, where linear baselines are assumed for both the native and denatured states ($S_N$ and $S_D$ at $T = 0$, with linear slopes $m_N$ and $m_D$; for reference see Privalov *et al.*[1]):

$$S = \frac{(S_N + m_N \mathrm{T}) + (S_D + m_D \mathrm{T}) \exp(-\frac{\Delta \mathrm{G}}{R\mathrm{T}})}{1 + \exp(-\frac{\Delta \mathrm{G}}{R\mathrm{T}})} \tag{2}$$

The values $m_N$ and $m_D$ were shared across mutants for far-UV or near-UV CD since the same baselines are expected for the three variants and are in fact visually identical. Not only does this procedure reduce the number of fitted parameters, but it is also required to get a reasonable estimate for the native baseline of near-UV CD of the I56T variant. $\Delta \mathrm{H}_m$ and $\Delta \mathrm{C}_p$ are the enthalpy change at $\mathrm{T}_m$, the melting temperature, and the heat capacity change of the transition, respectively, and are related to the change in free energy $\Delta$G according to:

$$\Delta \mathrm{G} = \Delta \mathrm{H} - T\Delta \mathrm{S} \tag{3}$$

$$\Delta \mathrm{H} = \Delta \mathrm{H}_m + \Delta \mathrm{C}_p(\mathrm{T} - \mathrm{T}_m) \tag{4}$$

$$\Delta \mathrm{S} = \Delta \mathrm{H}_m/\mathrm{T}_m + \Delta \mathrm{C}_p \ln(\mathrm{T}/\mathrm{T}_m) \tag{5}$$

$$\Delta \mathrm{G} = \Delta \mathrm{H}_m\big(1 - (\mathrm{T}/\mathrm{T}_m)\big) - \Delta \mathrm{C}_p\big((\mathrm{T}_m - \mathrm{T}) + \mathrm{T}\ln(\mathrm{T}/\mathrm{T}_m)\big) \tag{6}$$

### Fitting to a three-state model

Data was analysed using a sequential unfolding model with three states (native state N, intermediate state I and unfolded state U):

$$\mathrm{N} \rightleftharpoons \mathrm{I} \rightleftharpoons \mathrm{U} \tag{7}$$

The normalized spectroscopic signals $S$ for far- and near-UV CD were simultaneously fitted to equation 8, with linear baselines for the native and unfolded states ($S_N$ and $S_U$ at $T = 0$, with linear slopes $m_N$ and $m_U$), but with no first order dependency for the signal of the intermediate state ($S_I$ at $T = 0$), since this intermediate is populated in a narrow temperature window and that it also allows to reduce the number of fitted parameters:

$$S = \frac{(S_N + m_N \mathrm{T}) + S_I \exp(-\frac{\Delta G_1}{RT}) + (S_U + m_U \mathrm{T}) \exp(-\frac{\Delta \mathrm{G}_1 + \Delta \mathrm{G}_2}{RT})}{1 + \exp(-\frac{\Delta \mathrm{G}_1}{RT}) + \exp(-\frac{\Delta \mathrm{G}_1 + \Delta \mathrm{G}_2}{RT})} \tag{8}$$

where $\Delta \mathrm{G}_1$ and $\Delta \mathrm{G}_2$ are the free energies correponding to the transitions native/intermediate and intermediate/denatured, respectively. Melting temperatures $\mathrm{T}_{m,1}$ and $\mathrm{T}_{m,2}$ and enthalpy changes $\Delta \mathrm{H}_{m,1}$ and $\Delta \mathrm{H}_{m,2}$ are defined according to equation 6.

### Analysis of Differential Scanning Calorimetry

The excess heat capacity change was analysed using the sequential model for 2 or 3 states (see schemes 1 and 7). The average excess heat capacity $< \Delta \mathrm{C}_p >$ was derived from the average excess enthalpy $< \Delta \mathrm{H} >$ using equations 9 and 10, where $N = 2$ for a two-state model and $N = 3$ for a three-state model:

$$< \Delta \mathrm{C}_p > = \frac{\partial < \Delta \mathrm{H} >}{\partial T} \tag{9}$$

$$< \Delta \mathrm{H} > = \frac{\sum_{i=1}^{N} \left( \exp(-\frac{\sum_{j=1}^{i} \Delta \mathrm{G}_j}{RT}) (\sum_{j=1}^{i} \Delta \mathrm{H}_j) \right)}{1 + \sum_{i=1}^{N} \exp(-\frac{\sum_{j=1}^{i} \Delta \mathrm{G}_j}{RT})} \tag{10}$$

where the enthalpy changes $\Delta \mathrm{H}_j$ follow equation 4.

The simulation of the expected DSC thermograms according to a three-state model (see dashed line in Fig. 2b, main article) derives from these equations: knowledge of $\Delta \mathrm{H}_{m,j}$, $\Delta \mathrm{C}_{p,j}$ and $\mathrm{T}_{m,j}$ allows the determination of changes in heat enthalpy $\Delta \mathrm{H}_j$ for the two transitions. The evolution of the excess heat capacity can then be simulated using equations 9 and 10.

## References


[1] Privalov, P. L. *Adv Protein Chem.*, **1979**, *33*, 167–241.

[2] Cremades, N.; Velazquez-Campoy, A.; Freire, E.; Sancho, J. *Biochemistry*, **2007**, *47*(2).

[3] Bai, Y.; Milne, J. S.; Mayne, L.; Englander, S. W. *Proteins*, **1993**, *17*, 75–86.

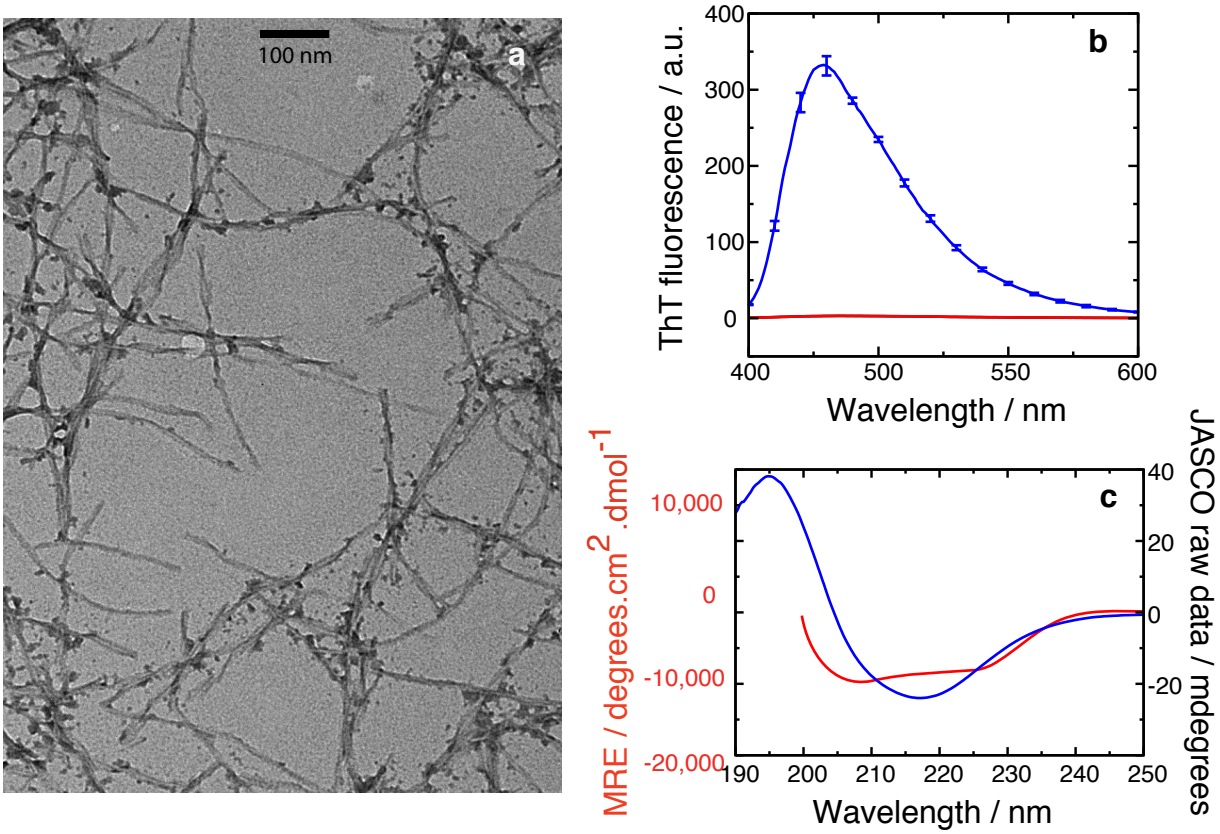


**Figure S1: Characterization of amyloid fibrils of human lysozyme.** Amyloid fibrils of wild type human lysozyme formed after incubation at ∼47.5°C, under stirring and with 2% v/v of fibril seed. (a): Transmission electron microscopy of a fibril sample. (b): Fluorescence emission spectrum of Thioflavin-T when mixed with the same amount of fibrils (blue) or monomers (red). The spectrum of a control sample without protein is identical to that of the monomeric sample. (c) Mean residual ellipticity of soluble, monomeric human lysozyme (red) and amyloid fibrils (blue) measured by far UV circular dichroism (CD) at 25°C.

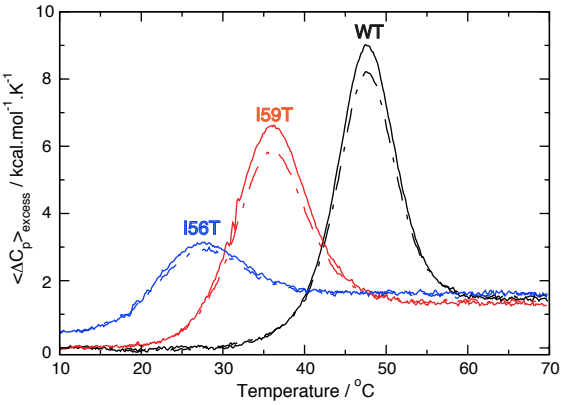


**Figure S2: Reversibility of the unfolding transition assessed by DSC.** Thermal denaturation of the three variants WT (black), I59T (red) and I56T(blue) obtained by DSC. After the first scan (plain lines), the protein is allowed to cool in the calorimetric cell before undergoing a second unfolding scan (dashed lines). The reversibility of the unfolding process is assessed by comparing the areas below the first and second melting traces: the reversibility for the WT was 95%, 90% for I59T and 93% for I56T.

**Table S1: Thermodynamic parameters resulting from fitting CD and DSC data.** Thermodynamic parameters obtained for the three variants, I56T, I59T and WT, by fitting the far-UV or near-UV CD transitions to a two-state model (first and second lines), or fitting both simultaneously to a three-state model (third line). The fitting of the DSC data to a two-state transition is shown in the fourth line. The mid-point temperature $T_m$ is expressed in Kelvins (K), the enthalpies $\Delta H_m$ associated to the transitions in kcal.mol$^{-1}$ and the heat capacities $\Delta C_p$ in kcal.mol$^{-1}$.K$^{-1}$. The errors on $T_m$ values are the largest between fitting errors and 1.0 K (instrumental error); fitting errors are reported for $\Delta H_m$ values.

| | | $T_{m,1}$ | $\Delta H_{m,1}$ | $\Delta C_{p,1}$ | $T_{m,2}$ | $\Delta H_{m,2}$ | $\Delta C_{p,2}$ |
|---|---|---|---|---|---|---|---|
| WT | far-UV CD | - | - | - | 325 ± 1 | 50 ± 4 | 0.2 |
| | near-UV CD | 321.2 ± 1.0 | 76 ± 4 | 1.6 | - | - | - |
| | global CD | 320.5 ± 1.0 | 98 ± 7 | 1.6 | 325 ± 3 | 39 ± 10 | 0.2 |
| | DSC | 320.1 ± 1.0 | 86 ± 0.5 | 1.6 | - | - | - |
| I59T | far-UV CD | - | - | - | 318 ± 1 | 33 ± 2 | 0.2 |
| | near-UV CD | 308.5 ± 1.0 | 68 ± 3 | 1.6 | - | - | - |
| | global CD | 307.7 ± 1.0 | 69 ± 7 | 1.6 | 317.5 ± 3 | 35 ± 8 | 0.2 |
| | DSC | 307.7 ± 1.0 | 56.5 ± 0.1 | 1.6 | - | - | - |
| I56T | far-UV CD | - | - | - | 317 ± 1 | 32.3 ± 3 | 0.2 |
| | near-UV CD | 297.5 ± 1.0 | 41 ± 3 | 1.3 | - | - | - |
| | global CD | 295.3 ± 2.0 | 27 ± 3 | 1.3 | 317.2 ± 2 | 31 ± 4 | 0.2 |
| | DSC | 296.7 ± 1.0 | 35.6 ± 0.2 | 1.3 | - | - | - |

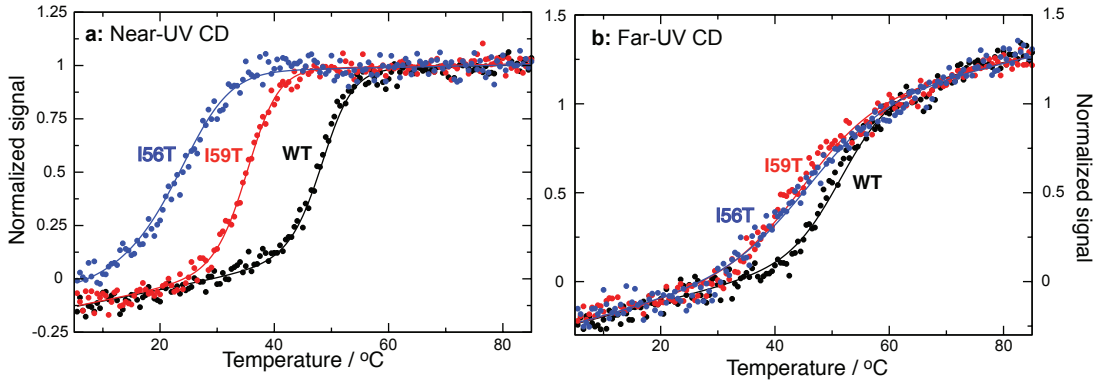


**Figure S3: Normalized CD signal.** (a): near-UV CD and (b): far-UV CD The three variants are colour-coded: WT (black), I59T (red) and I56T (blue). Plain lines correspond to the best simultaneous fit of the three variants to a two state model; baselines for the native and unfolded states are shared (see equations in Supporting Information).

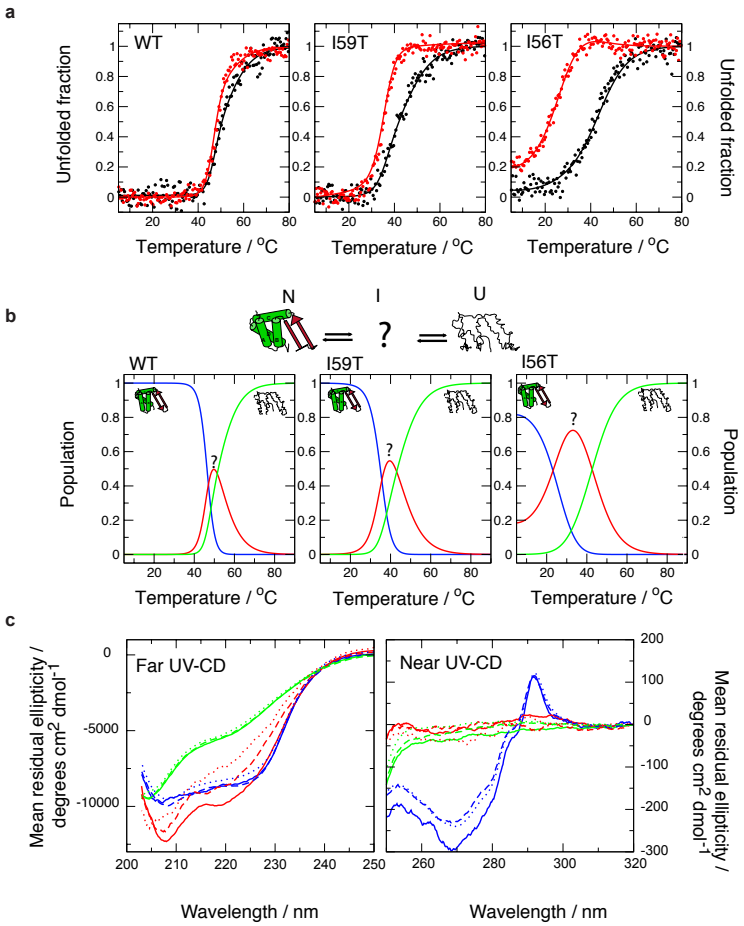


**Figure S4: Thermal unfolding of human lysozyme monitored using CD spectroscopy: analysis using a three-state model.** (a): Equilibrium thermal unfolding of WT, I59T and I56T lysozyme, followed by near-UV (270nm, red) and far-UV (222nm, black) CD at pH 1.2. The curves were fitted globally to a three-state model (the fits are represented in plain lines). (b): Population diagram derived from the three-state analysis. N, I and U refer to native (blue), intermediate (red) and unfolded (green) states, respectively. Interestingly, the structural transition between the intermediate and unfolded states (monitored by the far-UV CD transition at 222 nm) is essentially identical for the I56T and I59T variants, demonstrating that the higher population of the intermediate state for I56T relative to I59T is due to a more pronounced destabilization of the native state. In the case of I56T, this destabilization is such that the population of the native state never reaches 100%, even at the lowest temperatures studied here. (c): Far (left) and near (right) UV-CD spectra for the native (blue), intermediate (red) and unfolded (green) states, obtained by deconvolution of the CD spectra recorded at different temperatures, based on the knowledge of the populations of the different states. Data for the WT protein (dotted lines) and its I59T (dashed lines) and I56T (plain lines) variants are presented. The deconvolution of the spectra into pure native (N), intermediate (I) and unfolded (U) spectra, in the case of the three state model, was carried out assuming that, at any temperature, the observed spectroscopic signal $Y(\lambda, T)$ at a given wavelength $\lambda$ is a linear combination of the spectroscopic signatures of each state $i$, $Y_i(\lambda, T)$, modulated by their population $p_i$:[2] $Y(\lambda, T) = Y_N(\lambda, T) * p_N(T) + Y_I(\lambda, T) * p_I(T) + Y_U(\lambda, T) * p_U(T)$ An additional linear temperature dependency was introduced for $Y_N(\lambda, T)$. The spectrum of the native state is presented at 25°C. The spectra used for the deconvolution were recorded every 5°C, from 5°C to 90°C, and were averages of 3 and 6 scans for the far- and near-UV regions, respectively. Each spectrum was acquired in wavelength steps of 0.2 nm with a scanning speed of 50 nm/min, a response time of 4 seconds and a bandwidth of 2 nm.

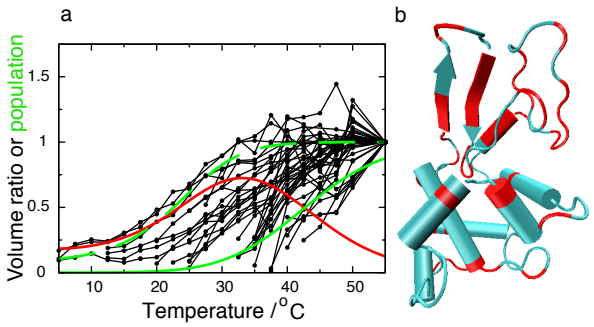


**Figure S5: Relative temperature increase in peak volume of NMR resonances from the denatured state.** (a): Black thin lines represent the temperature dependence of normalized $^{1}$H-$^{15}$N HSQC cross-peak volumes of the denatured state of I56T, normalized relatively to the volume of each peak at 55°C. The figure shows data for 40 non-overlapping and assigned residues: these residues (red) are mapped onto the protein native structure in panel (b). The population of the denatured state as derived by near-UV CD according to a two-state model fitting is overlaid in green (dashed line), in agreement with Fig. 2. The populations of the intermediate and unfolded states, as derived by global fitting the far and near-UV CD to a three-state model, are shown in red and green solid lines, respectively, in agreement with figure 2. This figure shows the same result as for glycine residues only (right column in Fig. 4, main article), but for more residues, and of all types.

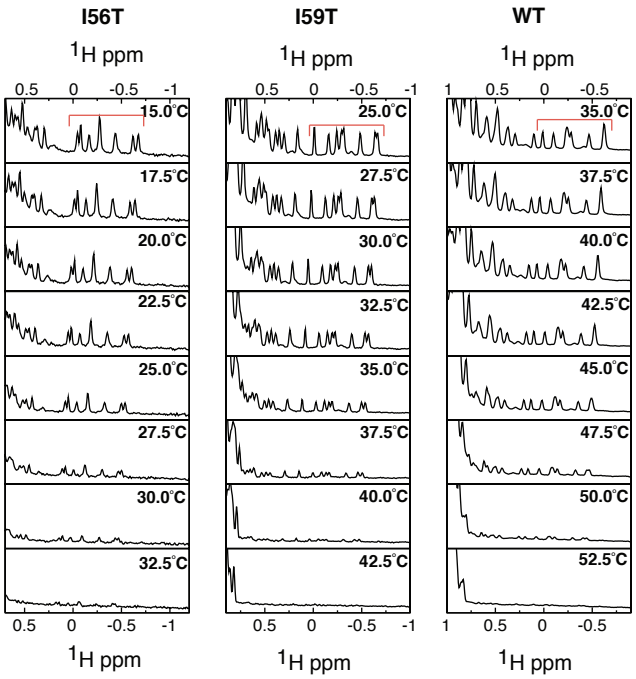


**Figure S6: Raw NMR 1D data: temperature evolution.** Selected region of the $^{1}$H 1D NMR spectrum showing the temperature decay in signal arising from buried methyl groups of the protein native state. The integrated regions are shown in red (see light blue dots in main article, left column in Fig. 4 and Methods).

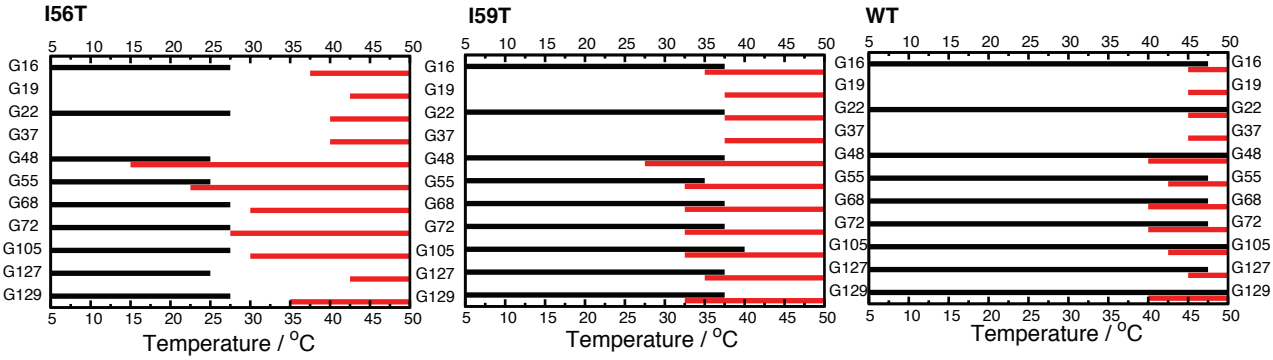


**Figure S7: Temperatures of disappearance and appearance of NMR resonances in HSQC spectra.** Temperatures at which cross-peaks of glycine residues of the native state (black) and denatured ensemble (red) are visible in the HSQC spectrum. Spectra were recorded every 2.5°C and were analysed while keeping the contour levels constant (see Fig. 5 and 6, main text). The temperatures are reported in Celsius (°C).

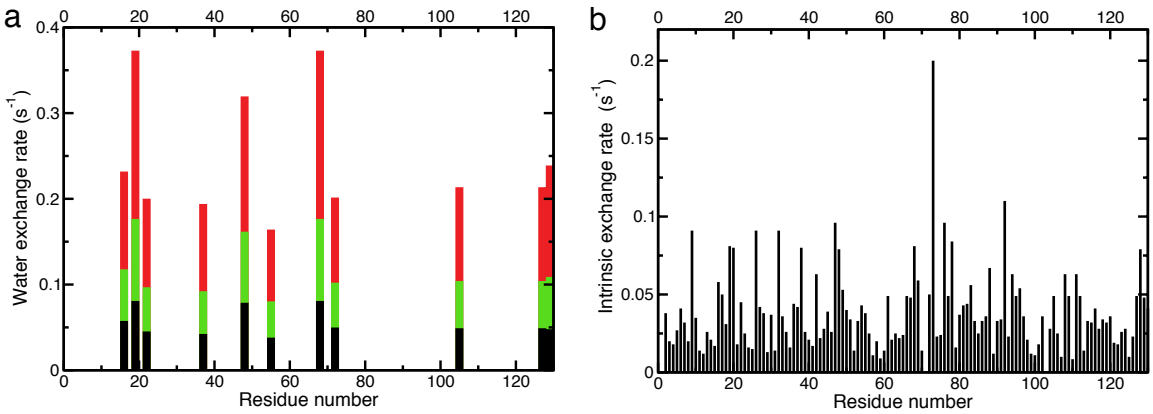


**Figure S8: Predicted water exchange rates of amide protons.** (a): Water exchange rates of backbone glycine amides calculated for human lysozyme at pH 1.2 and 30°C (black), 40°C (green) or 50°C (red), according to http://hx2.med.upenn.edu/ [3](b): Same but for the whole sequence of human lysozyme, at 30°C. For both panels, data relative to the WT protein only are shown as the single point mutations I56T and I59T have only a very small and local effect on the exchange rates (over 2 neighbouring residues).