Supplementary information to

A General Approach to Detect Folding Intermediates from Steady-State and Time-Resolved Fluorescence of Single Tryptophan-Containing Proteins

Sergey P. Laptenok,[¶1] Nina V. Visser,[†1] Ruchira Engel,[‡] Adrie H. Westphal,[‡] Arie van Hoek,[†] Carlo P.M. van Mierlo,[‡] Ivo H.M. van Stokkum,[¶] Herbert van Amerongen[†*], Antonie J.W.G. Visser[‡]

[¶]Department of Physics and Astronomy, Faculty of Exact Sciences, Vrije Universiteit Amsterdam, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands,
Laboratories of [†]Biophysics and [‡]Biochemistry, Microspectroscopy Centre, Wageningen University, P.O. Box 8128, 6700 ET Wageningen, The Netherlands.

Global analysis of steady-state fluorescence spectra with the help of a spectral model

At each GuHCl concentration, the observed emission spectrum $\psi(\lambda)$ is described as a linear combination of spectra arising from native protein, folding intermediate and unfolded protein:

$$\psi(\lambda) = c_n\varepsilon_n + c_i\varepsilon_i + c_u\varepsilon_u$$

Steady state fluorescence spectra obtained of protein at 0 M and 4 M GuHCl are used as reference spectra that characterize native and unfolded molecules, respectively. The steady state fluorescence spectrum of the folding intermediate was modeled as a skewed Gaussian in the energy domain ($\bar{\nu} = 1/\lambda$), and is described by three parameters: peak location $\bar{\nu}_{\max}$, width $\Delta\bar{\nu}$ and skewness b (*1, 2*):

$$\varepsilon_i(\bar{\nu}) = \bar{\nu}^5 \exp(-\ln(2)\{\ln(1+2b(\bar{\nu}-\bar{\nu}_{\max})/\Delta\bar{\nu})/b\}^2)$$

where the parameter $\bar{\nu}_{max}$ is the Franck-Condon wavenumber of maximum emission. The FWHM is given by $\Delta\bar{\nu}_{1/2} = \Delta\bar{\nu}\sinh(b)/b$. Note that with skewness parameter b equal to zero the expression simplifies to a Gaussian.

Now all spectra are globally analyzed as a linear combination of spectra arising from native protein, folding intermediate and unfolded protein. The three parameters that describe the shape of the spectrum of the folding intermediate and the relative concentrations of each folding species are the parameters that are allowed to float freely during the global fit. The concentration parameters were constrained to be nonnegative. In this way the product $c_i\varepsilon_i$ can be estimated, and thus we can determine the shape of c_i as a function of GuHCl concentration. In order to estimate the concentration c_i relative to the other concentrations, we use the constraint that the sum $c_n + c_i + c_u$ should be close to one at all GuHCl concentrations. This is estimated by means of linear regression. All estimated concentrations and spectra are depicted in Fig. S1, as well as this sum $c_n + c_i + c_u$.

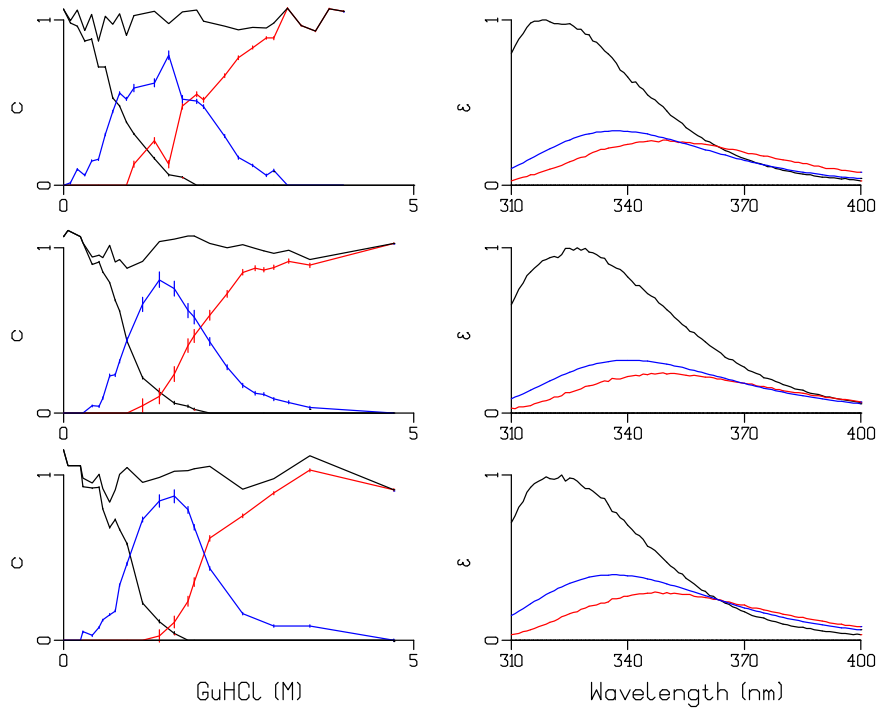

Fig. S1. (right) Fluorescence spectra of WFF (top), WWF (middle) and WFW (bottom). Key: black, native; blue, intermediate; red, unfolded. (left). Relative concentrations of the different states versus [GuHCl] estimated from global analysis. The sum $c_n + c_i + c_u$ is shown as the top black curve.

References

1. van Stokkum, I. H. M., Larsen, D. S., and van Grondelle, R. (2004) Global and target analysis of time-resolved spectra, *Biochim. Biophys. Acta 1657*, 82-104.
2. van Stokkum, I. H. M., Linsdell, H., Hadden, J. M., Haris, P. I., Chapman, D., and Bloemendal, M. (1995) Temperature-induced changes in protein structures studied by Fourier-transform infrared-spectroscopy and global analysis, *Biochemistry 34*, 10508-10518.