**Supplementary Data**

## Bland-Altman statistical analysis of Sypro Orange fluorescence and DSC thermal stability data.

Rather than plotting the values of the thermal transmissions against each other (Figure 2D), the difference between the two methods (DSC minus Sypro Orange) for each thermal transition is plotted against the average value of each thermal transition, a so-called Bland-Altman plot [1] (Figure S1).

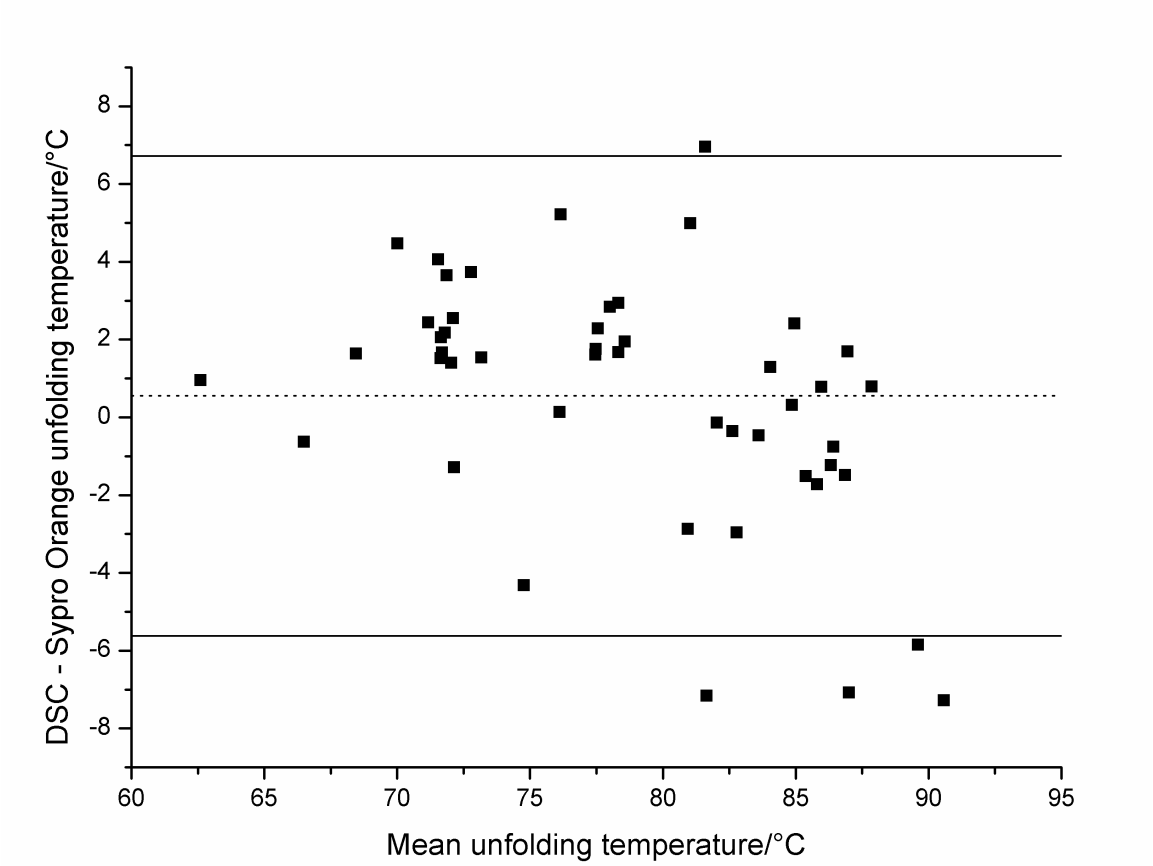


**Figure S1. Fluorescence intensity at different protein concentrations**
For each unfolding transition, the difference between the value obtained using DSC minus the value obtained in the fluorescence experiment is plotted against the average of the two values. Dashed line: mean difference. Solid line: two standard deviations above and below the mean.

From the Bland-Altman plot it can be seen that out of 48 datapoints, 43 are within two standard deviations of the mean, or 89.5%. This is short of the expected 95% for a normal distribution, but 72.9% are within one standard deviation, and 100% are within three standard deviations, suggesting the distribution is most likely normal. The mean difference $\bar{d}$ was 0.55°C with a standard deviation $s$ of 3.15°C and a 95% confidence

interval from -0.36°C to +1.47°C. The mean difference suggests the fluorescence method slightly underestimates the actual melting temperatures, however, even taking into account the confidence intervals, the difference is smaller than the minimum temperature difference required for two proteins to be regarded as having significantly different melting temperatures. For this particular dataset, the "limits of agreement"[1], which are defined as $\bar{d} \pm 1.96s$, range from -5.62°C to +6.73°C. This range is quite significant but may not be all due to discrepancies between the two thermal stability methods tested: instrumental measurement errors and fitting approximations in extracting melting temperatures from both techniques may contribute to this. It is also possible that the two methods measure subtly different aspects of the unfolding process. Overall, therefore, while the mean difference itself suggests the overall agreement between the two techniques is very good, for individual transitions there can be quite significant variation between the two. This suggests that the fluorescence method of thermal stability measurement is a useful tool for early drug candidate characterization, but may not be accurate enough to be used in characterizing the one or two promising drug candidates that are destined for larger animal or human trials.

**Reference**

1. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Int J Nurs Stud 2010;47(8):931-936.