Supplemental Data

Cross-Strand Split Tetra-Cys Motifs as Structure Sensors in a β-Sheet Protein

Beena Krishnan and Lila M. Gierasch

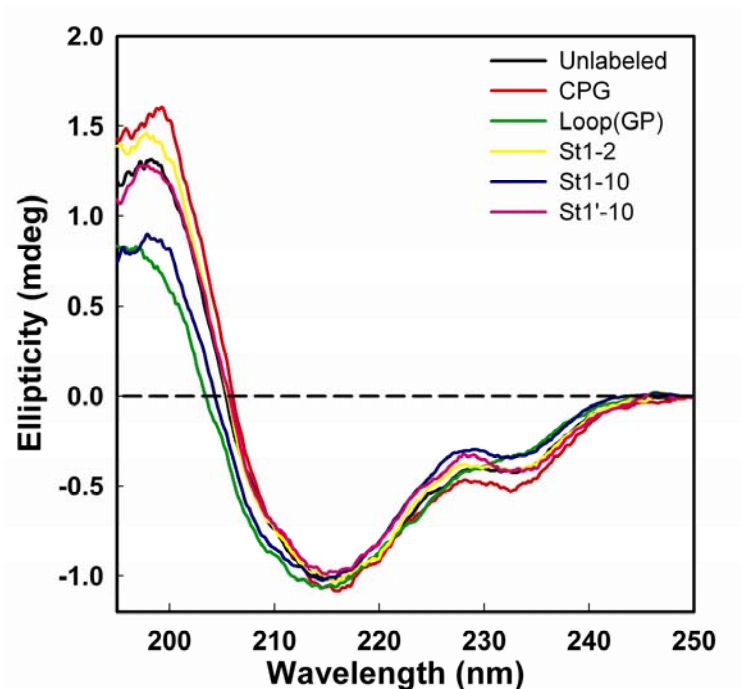


Figure S1. Far-UV Circular Dichroism Spectra of the Unlabeled CPG Protein (black line) and the FlAsH-Labeled Tetra-Cys Proteins (colored lines)
CD data were acquired at 25 °C in 10 mM sodium phosphate buffer (pH 7.4) on a JASCO J715 spectropolarimeter. The presence of a shoulder at 232 nm is a characteristic of well-folded CRABP I (Clark et al., 1998), and was observed for all labeled proteins.

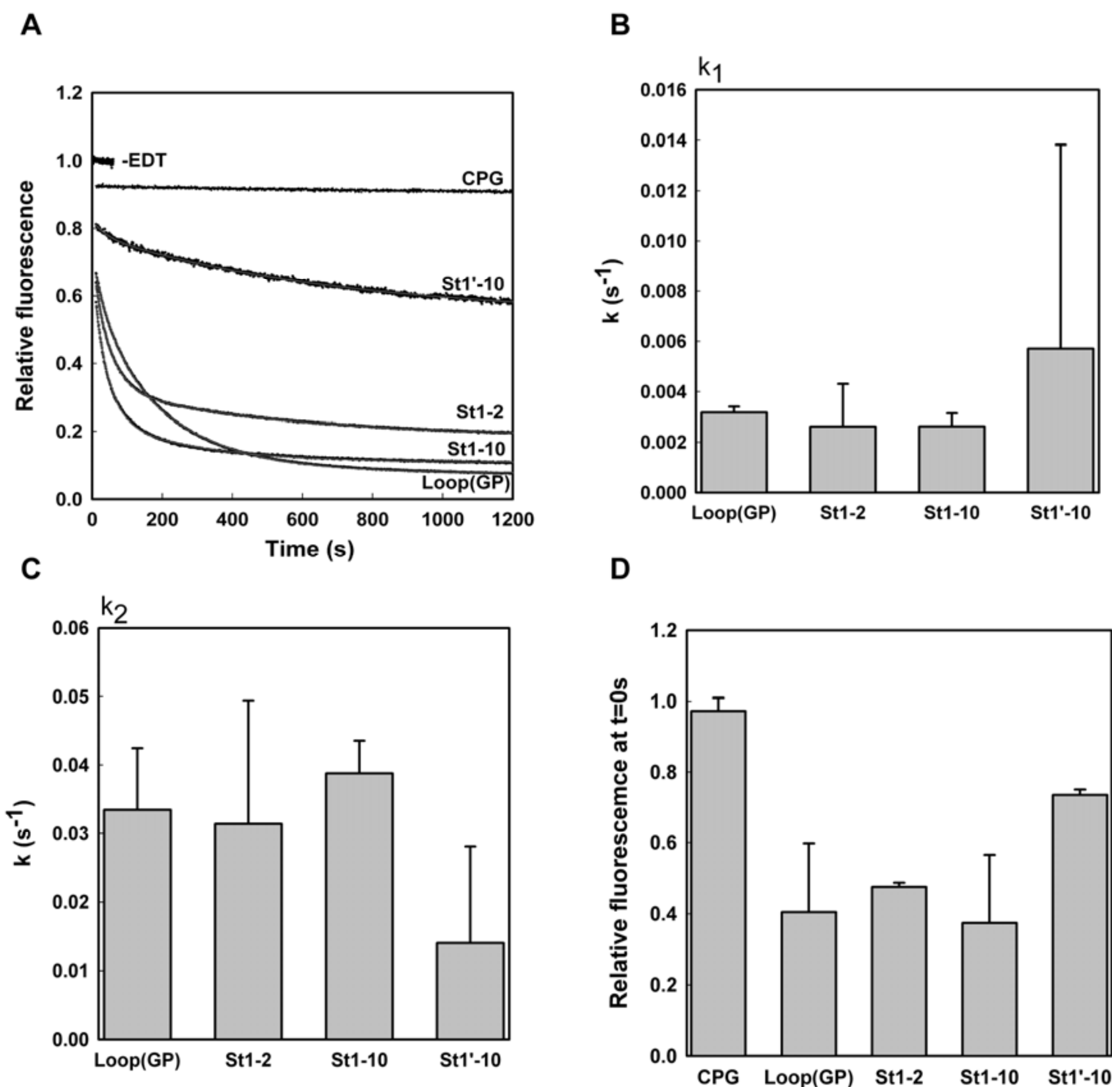


Figure S2. EDT-Mediated Displacement of FlAsH from Labeled Proteins
The kinetics of the EDT displacement of FlAsH from labeled proteins fit well to a bi-exponential equation, indicating a two-step dye displacement process (A). The rate-limiting step may involve the release of one of the di-cysteine motifs, which then accelerates dislodging of the dye molecule in a second faster step. Interestingly, the rate constants of the first and second kinetic phases were observed to be very similar in magnitude for all the protein constructs (B and C). We conclude that the observed differences in apparent affinity of FlAsH-protein complexes arise from the initial displacement of the FlAsH molecules as detected at the earliest observation time point in these manual mix experiments (burst phase amplitudes) or from the fluorescence extrapolated to t=0 (D).
(A) Representative kinetic traces for FlAsH displacement from the labeled proteins in the presence of 0.1 mM EDT. The biexponential fits to the data are represented by solid lines. (
(B and C) The two rate constants (k_1 and k_2) obtained from three independent titrations for each protein.
(D) The fluorescence signals for each kinetic curve extrapolated to t = 0, expressed relative to the fluorescence in the absence of EDT.

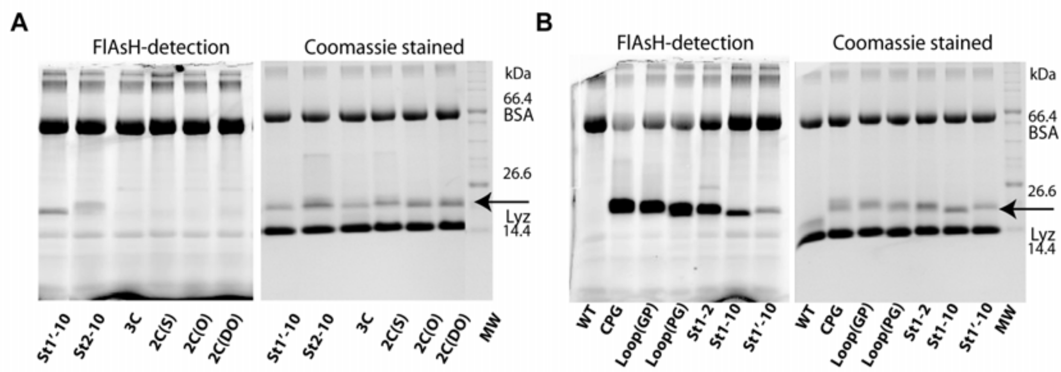


Figure S3. FlAsH Labeling Under Stringent Conditions
Control proteins with zero to three cysteines (as illustrated in Figure 4A) were labeled in vitro in glutathione redox buffer containing BSA and lysozyme (Lyz) to mimic cellular conditions (A), as described in Experimental Procedures.
(B) Labeling of the purified tetra-Cys proteins in the presence of glutathione and a protein mixture of BSA and Lyz.
The left and the right panels in (A) and (B) correspond to detection of FlAsH fluorescence by phosphorimaging and Coomassie staining of the same gel, respectively. Only the proteins with four cysteines were labeled under stringent labeling conditions, demonstrating the selectivity of FlAsH binding only to the tetra-Cys proteins. The arrow indicates the position of CRABP I.

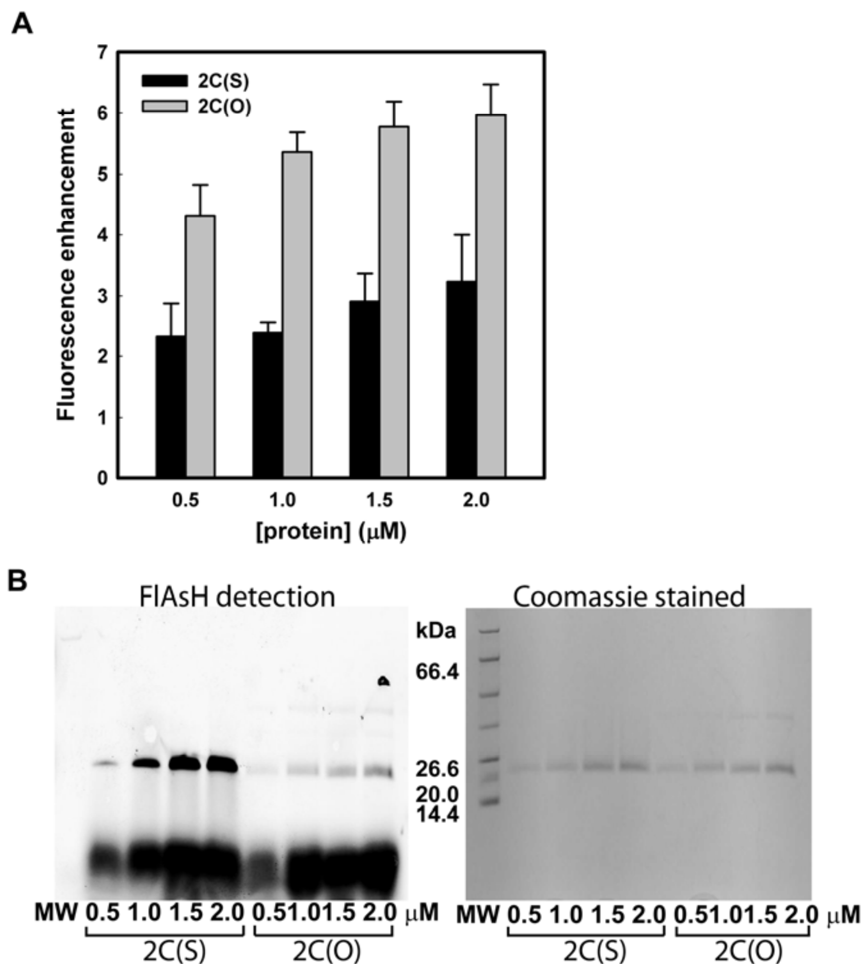


Figure S4. FlAsH binding to Di-Cysteine Constructs with a Pair of Cysteines Either on the Same Strand (2C(S)) or Across Strands (2C(O))

(A) Fluorescence enhancement with respect to the free dye fluorescence upon labeling varying protein concentrations (0.5 – 2.0 μM) with a two-fold molar excess of FlAsH at each concentration.
(B) The same samples analyzed on a 10% tricine SDS-PAGE. The left and the right panels in (B) represent the fluorescence detection by phosphorimaging and the Coomassie staining of the same gel, respectively.

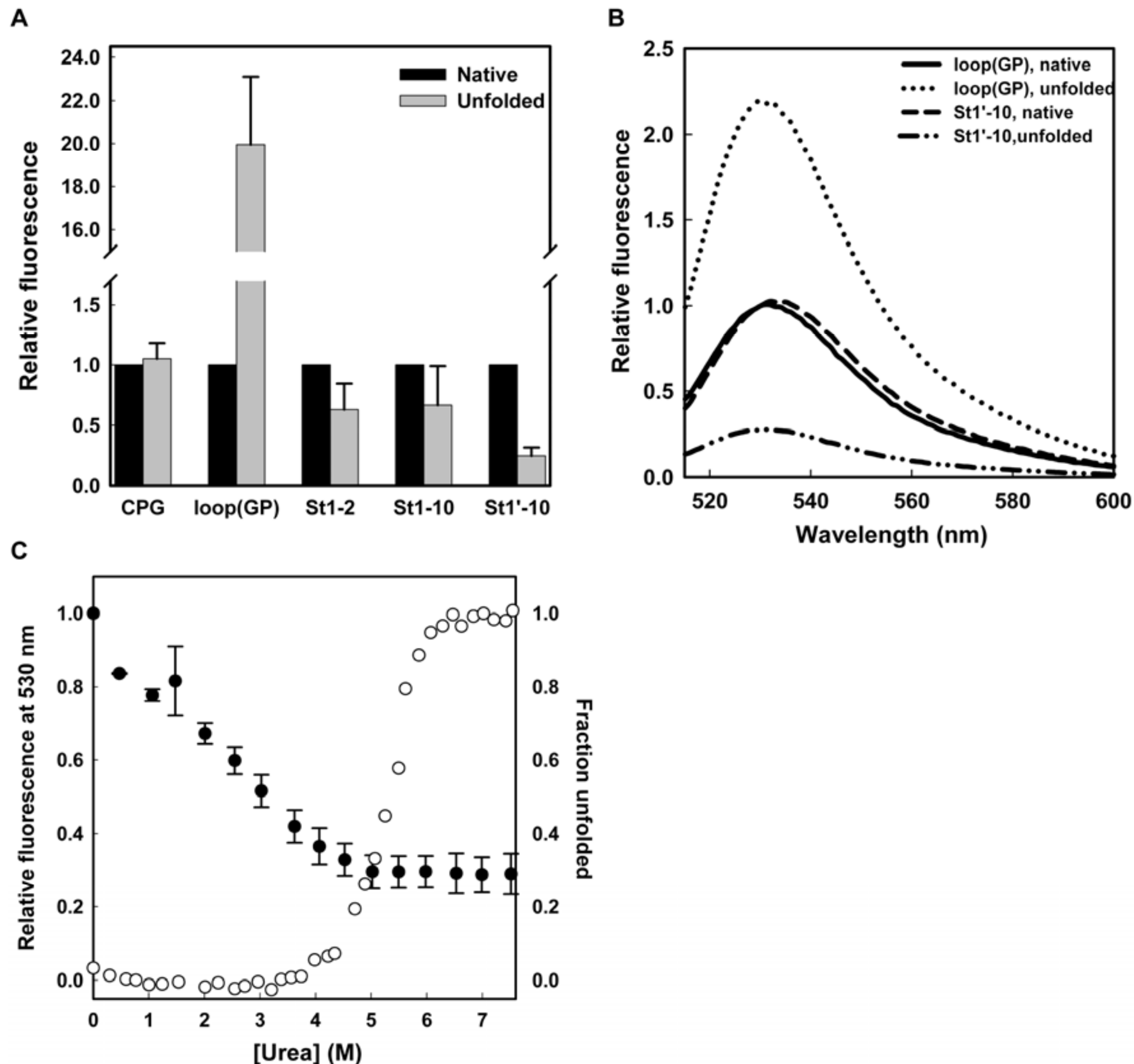


Figure S5. Comparative Analysis of FlAsH Binding
(A) FlAsH fluorescence of the purified labeled tetra-Cys proteins in the native and urea-unfolded state, measured in the presence 1 mM EDT. The fluorescence intensity of the labeled unfolded protein is given relative to the fluorescence signal obtained for the labeled native CPG protein. The absolute fluorescence intensities of the FlAsH-bound split tetra-Cys motif-containing and the loop(GP) continuous motif-containing proteins in the native state were 2 – 30 fold lower than that of the CPG protein.
(B) Fluorescence observed for FlAsH bound to the native (N) and urea-unfolded (U) states of the loop(GP) and the St1'-10 proteins in cell lysate. The relative fluorescence signal was obtained by normalizing the data using the fluorescence intensity at 530 nm for the labeled protein in the native state. In the case of the loop(GP) protein, the labeled unfolded protein (dotted line) is hyperfluorescent relative to the labeled native protein (solid line), as previously reported (Ignatova and Gierasch, 2004). In contrast, the St1'-10 protein exhibited a higher FlAsH fluorescence in the native state (dashed line) than in the unfolded state (dash-dotted line).

(C) In vitro urea titration of the pure unlabeled St1'-10 protein monitoring the tryptophan fluorescence at 350 nm (open circles), represented as the fraction unfolded (right axis), and FlAsH-labeled St1'-10 protein monitoring the FlAsH fluorescence at 530 nm (filled circles), represented as fluorescence relative to that observed in 0 M urea (left axis). The urea titration of the FlAsH-labeled protein is described in the Experimental Procedures. The error bar for the labeled protein samples were obtained as the standard error from two measurements. Together with the gel filtration results presented in Figure 7A, these data show that reduction in dye binding as the protein unfolds is an indication of loss of structure, but the apparent C_m is not directly comparable to that from a traditional urea melt. Here the process monitored is dependent on EDT concentration, as it reflects a dye displacement reaction.

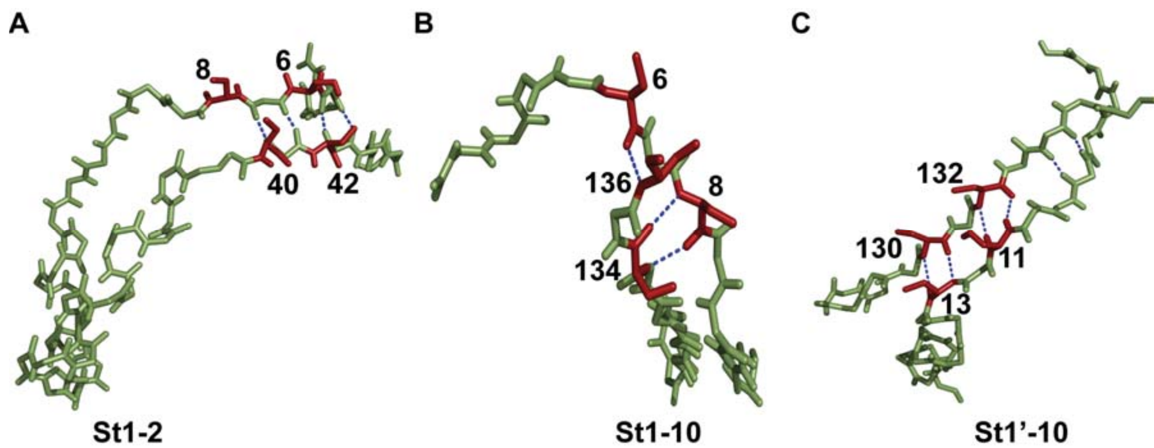


Figure S6. Details of the Structures of the Split Tetra-Cys Motifs Modeled into the CRABP I Structure, and Showing the Backbone Hydrogen-Bonding Networks

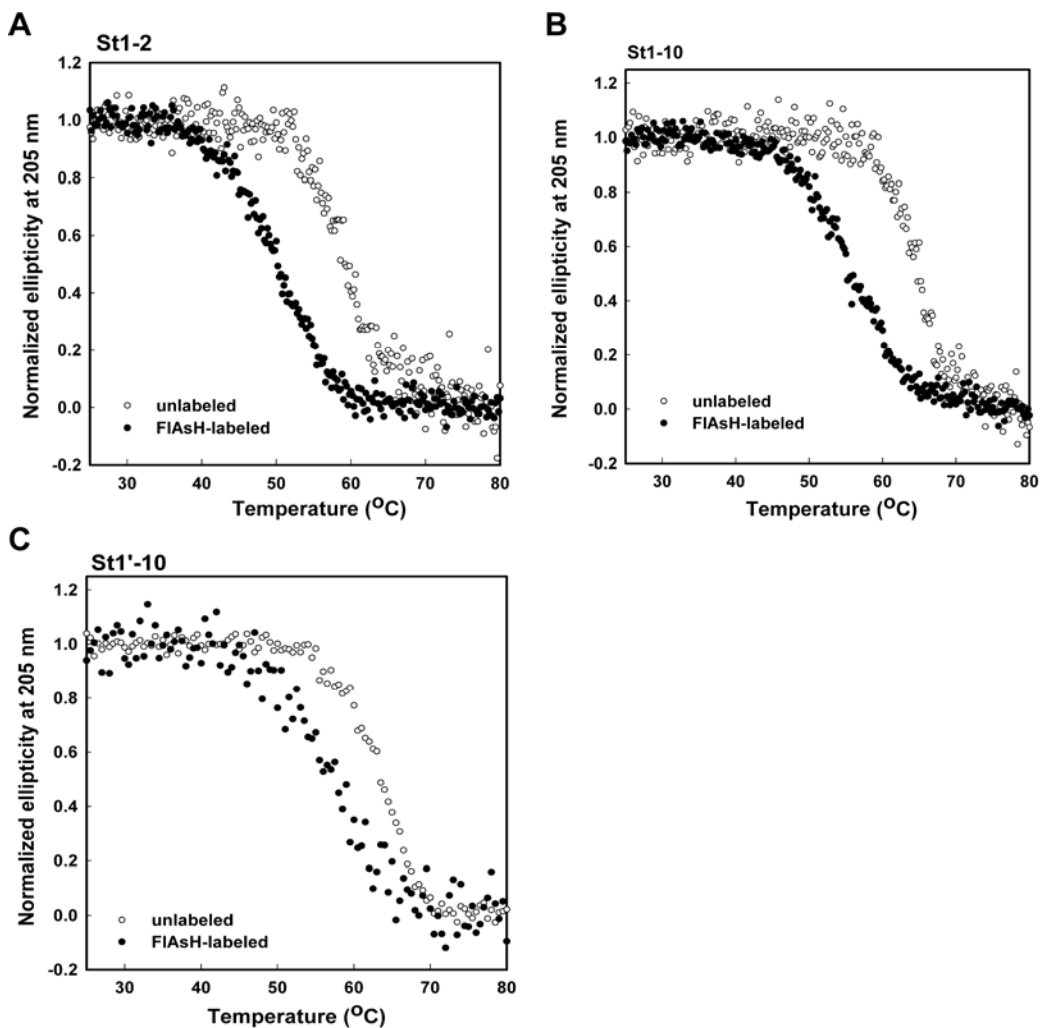


Figure S7. Effect of FlAsH-Binding on the Stability of the Split Tetra-Cys Proteins
(A–C) Unlabeled (open circles) and FlAsH-labeled (filled circles) split tetra-Cys proteins were thermally unfolded (heating rate 1 degree C/min) and secondary structure monitored by CD at 205 nm. FlAsH labeling of the proteins was carried out in the absence of EDT. Although the extent of labeling was not 100%, the observed shifts in melting curves show that dye binding results in lower thermal stability of all three variants.

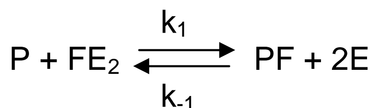
Table S1. CRABP I Mutants Used in the Current Study

Protein name	Mutations
CRABP I WT	R131Q (Template for all the mutants)
CPG	SCCPGCCG appended at the C-terminus
Loop (GP)	^{102}CCGPCC^{107} (in the Ω loop)
St1-2	T6C, K8C, H40C, E42C
St1-10	T6C, K8C, V134C, E136C
St1'-10	S11C, E13C, C129A, T130C, I132C
St2-10	H40C, E42C, C129A, T130C, I132C
3C	S11C, C129A, T130C, I132C
2C(S)	C129A, T130C, I132C
2C(O)	E13C, C129A, T130C
2C(DO)	S11C, C129A, T130C

S, same strand; O, opposite strands; DO, diagonally opposite.

Rate Equation Analysis for FlAsH-Binding to Tetra-Cys Motif

FlAsH binding to a tetra-Cys motif present on a protein or a peptide in the absence of free EDT can be represented by the following rate expression:



$$P + FE_2 \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} PF + 2E$$

where P is unlabeled protein; FE_2 is FlAsH bound to two EDT molecules; PF is protein labeled with FlAsH; E is EDT; k_1 is the forward rate constant, and k_{-1} is the reverse rate constant.

The rate of formation of the labeled protein, d[PF]/ dt, can be written as

$d[PF] / dt = k_1 * [P]_f * [FE_2]_f - k_{-1} * [PF] * [E]^2$

where the subscript f refers to the available free concentration of the corresponding species.

At steady state, d[PF] / dt = 0, and

$k_1 * [P]_f * [FE_2]_f = k_{-1} * [PF] * [E]^2$.

At a very high molar ratio of total protein ($[P]_t$) to the dye ($[F]_t$), *i.e.,* $[P]_t >>> [FE_2]_t$, FlAsH binding can be considered to be a pseudo uni-molecular reaction, and the following assumptions can be made: $[P]_t >>> [PF]$ and $[P]_f \sim [P]_t$.

Also under such a labeling condition, the amount of released EDT ([E]) will be insignificant and $[PF] * [E]^2$ can be considered to be negligible; thereby implying that $k_1 * [P]_f * [FE_2]_f = 0$.

Substituting $[P]_t$ for $[P]_f$ and $[FE_2]_t$ - [PF] for $[FE_2]_f$ in the above expression results in $[PF] = [FE_2]_t$, *i.e.*, the final yield of the labeled protein in such labeling conditions equals the amount of the dye used.

Under any other labeling condition, such as when $[P]_t \sim [FE_2]_t$ or at a larger excess of $[FE_2]_t$, the extent of the reaction will depend on the affinity of fluorophore for the binding motif. This is an important concern in the case of the

split tetra-cys proteins, since, unlike FlAsH-binding to a tetra-Cys motif in a peptide or to a protein such as CPG, wherein the reaction is practically irreversible with a negligible reverse reaction due to $k_{-1} <<< k_1$, the reverse rate constant, k_{-1}, for the split motifs is expected to be significant.
Thus, it is advantageous to carry out FlAsH labeling of the split tetra-Cys motif carrying proteins under a condition of large excess of protein with a very low dye concentration, such that the bulk of the dye is expected to be in complex with the protein.

Supplemental Reference

Clark, P.L., Weston, B.F., and Gierasch, L.M. (1998). Probing the folding pathway of a β-clam protein with single-tryptophan constructs. Fold. Des. *3,* 401-412.