SUPPLEMENTARY INFORMATION

Rational Design of a Highly Reactive Dicysteine Peptide Tag For Fluorogenic Protein Labelling

Miroslava Strmiskova, Kelvin Tsao, Jeffrey W. Keillor*

Department of Chemistry and Biomolecular Sciences

University of Ottawa,

10 Marie-Curie, Ottawa, ON K1N 6N5 Canada

Email: jkeillor@uottawa.ca

Table of Contents

Cloning

Table S1. Sequences of oligonucleotides used in this study. Mutations are underlined.

Oligonucleotide	**Sequence (5' → 3')**
dC10-A17X_fw	GCAGAGAAGCTVRNGCTAGAGCTGGAGGAAAGTAGGGAATTCC
dC10-A17X_bw	CCTCCAGCTCTAGCNYBAGCTTCTCTGCATGCAGCTTCTCTAGC
dC10-A16HR-A17K_fw	GCAGAGAACRYAAAGCTAGAGCTGGAGGAAAGTAGGGAATTCC
dC10-A16HR-A17K_bw	CCTCCAGCTCTAGCTTTRYGTTCTCTGCATGCAGCTTCTCTAGC
dC10-A16K-A17K_fw	GCAGAGAAAAGAAAGCTAGAGCTGGAGGAAAGTAGGGAATTCC
dC10-A16K-A17K_bw	CCTCCAGCTCTAGCTTTCTTTTCTCTGCATGCAGCTTCTCTAGC
dC10-A3HR_fw	GAGGGAAGGCTGAGCCRYGCTGAGTGCGCTGCTAGAGAAG
dC10-A3HR_bw	CAGCGCACTCAGCRYGGCTCAGCCTTCCCTCGATCCCGAG
dC10-A3K_fw	GAGGGAAGGCTGAGCAAAGCTGAGTGCGCTGCTAGAGAAG
dC10-A3K_bw	CAGCGCACTCAGCTTTGCTCAGCCTTCCCTCGATCCCGAG
dC10-A3R_fw	GAGGGAAGGCTGAGCAGAGCTGAGTGCGCTGCTAGAGAAGCTG
dC10-A3R_bw	CAGCGCACTCAGCTCTGCTCAGCCTTCCCTCGATCCCGAG
dC10-A3H_fw	GAGGGAAGGCTGAGCCATGCTGAGTGCGCTGCTAGAGAAGCTG
dC10-A3H_bw	CAGCGCACTCAGCATGGCTCAGCCTTCCCTCGATCCCGAG
H2B-dC10*_gBlock	CATGCTGTGTCCGAGGGCACTAAGGCAGTTACCAAGTACACTAG CTCTAAGGATCCAGGATCTTCACTGAGCAAGGCTGAGTGCGCTG CTAGAGAAGCTGCATGCAGAGAAAGAAAGGCTAGAGCTGGAGG AAAGTAATCTAGAGGGCCCTATTCTATAGTGTCACCTAAATGCT AGAGCTCGCTG

Kinetic Studies

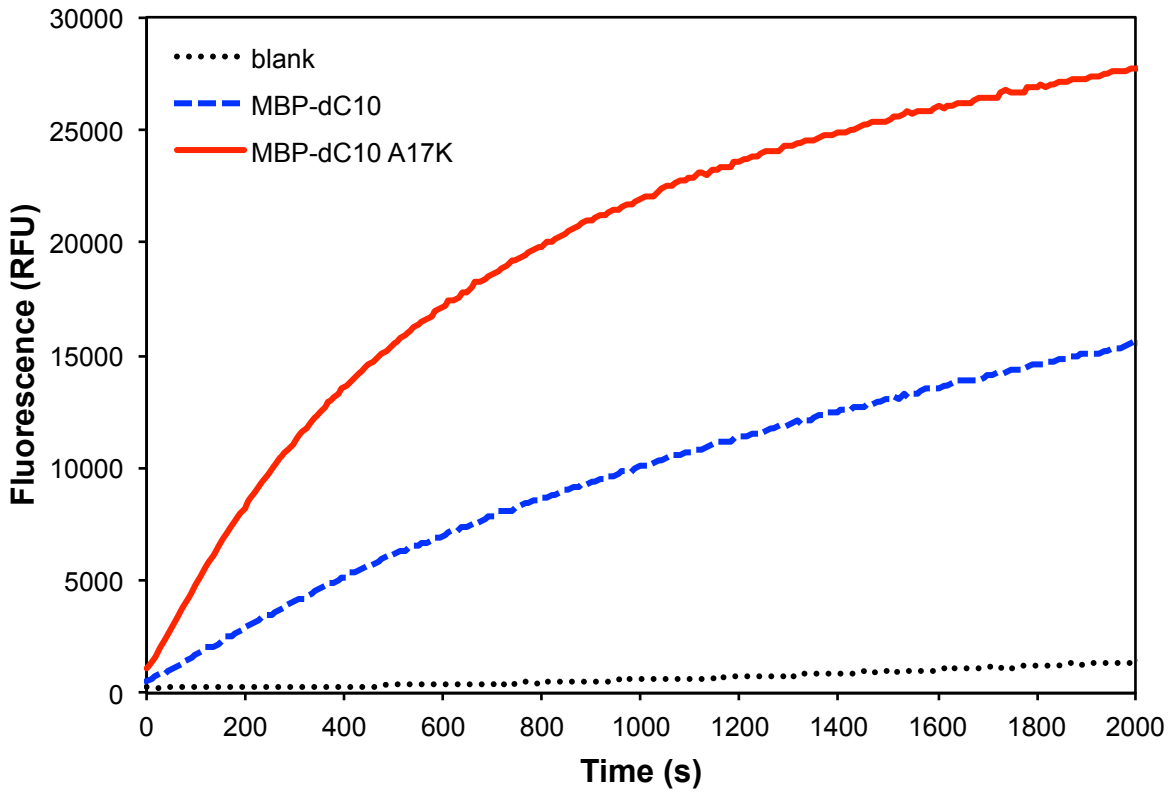

Figure S1. Fluorogenic addition reactions of fluorogen 1. Comparison of background reaction of fluorogen **1** (*dotted line*) with its reaction with target protein MBP-dC10α (*dashed blue line*) and single point mutant MBP-dC10α A17K (*solid red line*). Briefly, in a final reaction volume of 200 µL, 50 µM MBP-dC10α (or variants) and 50 µM **1** were mixed in 50 mM HEPES buffer (pH 7.5) in the presence of 1 mM TCEP. Fluorescence increase was followed at 515 nm, upon excitation at 330 nm.

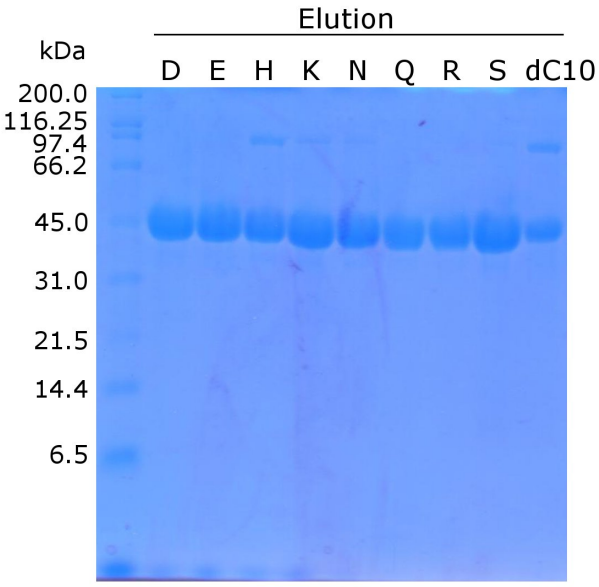

Figure S2. Purity of MBP-dC10α A17X mutants. Elution fractions from amylose columns with pure MBP-dC10α variants were separated on a Tris-tricine SDS-PAGE gel and stained with Coomassie blue. Broad range molecular weight marker (Bio-Rad) was used to estimate as a reference for molecular weight. dC10α variants were denoted as follows: D = A17D, E = A17E, H = A17H, N = A17N, Q = A17Q, R = A17R, S = A17S.

Helical Propensity

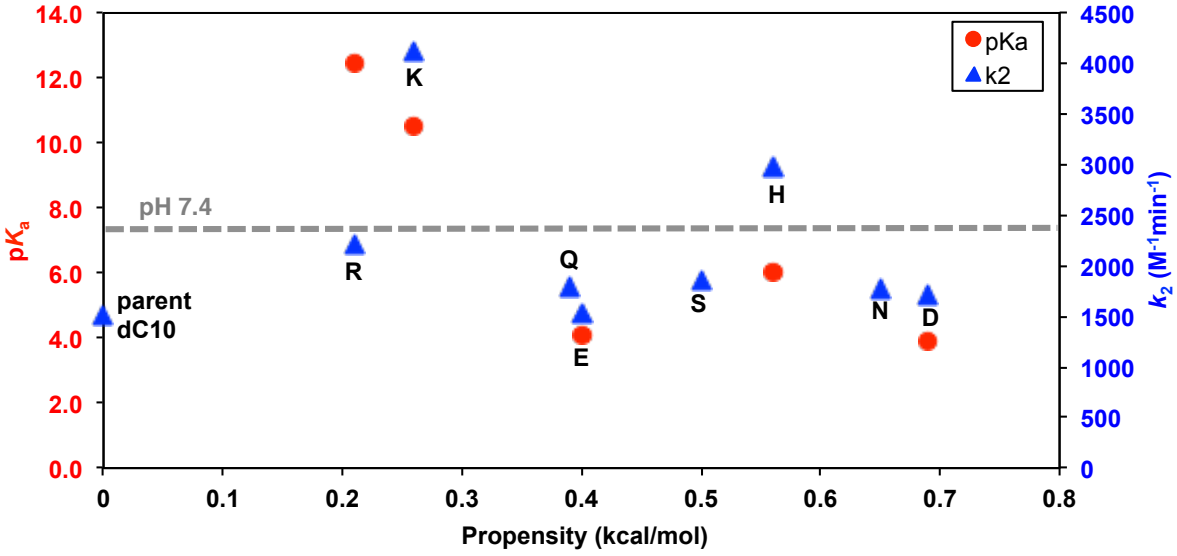

Figure S3. Predicted helical propensity and pK_a of mutant residues used in MBP-dC10α mutant library I. Propensity values represent the destabilization energy of a helix of a given residue relative to alanine, as reported by Pace *et al*.[1] For the single mutant library A17X, the pK_a of each residue X[2,3,4] is shown (*red circles*, left vertical axis) along with the measured second order rate constants shown in Figure 3 (*blue triangles*, right vertical axis). The dashed line represents the physiological pH of 7.4, intended to indicate the expected protonation state of each residue.

Table S2. Second order rate constants for the fluorogenic addition of MBP-dC10α mutants and fluorogen **1**.

MBP-dC10 Mutant	k_2 **($M^{-1}min^{-1}$)**
dC10 parent	1563 ± 84
S2H (20 °C)	647 ± 24
A7H (20 °C)	897 ± 8
R9H (20 °C)	826 ± 142
S2H-A17H (20 °C)	2628 ± 293
A17D	1711 ± 123
A17E	1531 ± 296
A17N	1781 ± 62
A17Q	1786 ± 189
A17S	1855 ± 59
A17H	2980 ± 123
A17K	4125 ± 26
A17R	2208 ± 146
A16H-A17K	6834 ± 669
A16K-A17K	9747 ± 129
A16R-A17K	12052 ± 889
A3H-A16H-A17K	7036 ± 1371
A3K-A16H-A17K	8963 ± 1213
A3R-A16H-A17K	6431 ± 584
A3H-A16K-A17K	7772 ± 1726
A3K-A16K-A17K	12183 ± 1104
A3R-A16K-A17K	12781 ± 347
A3H-A16R-A17K	8540 ± 975
A3K-A16R-A17K (dC10*)	15008 ± 153
A3R-A16R-A17K	9688 ± 1849

Helical Wheel Projection of dC10 Sequence

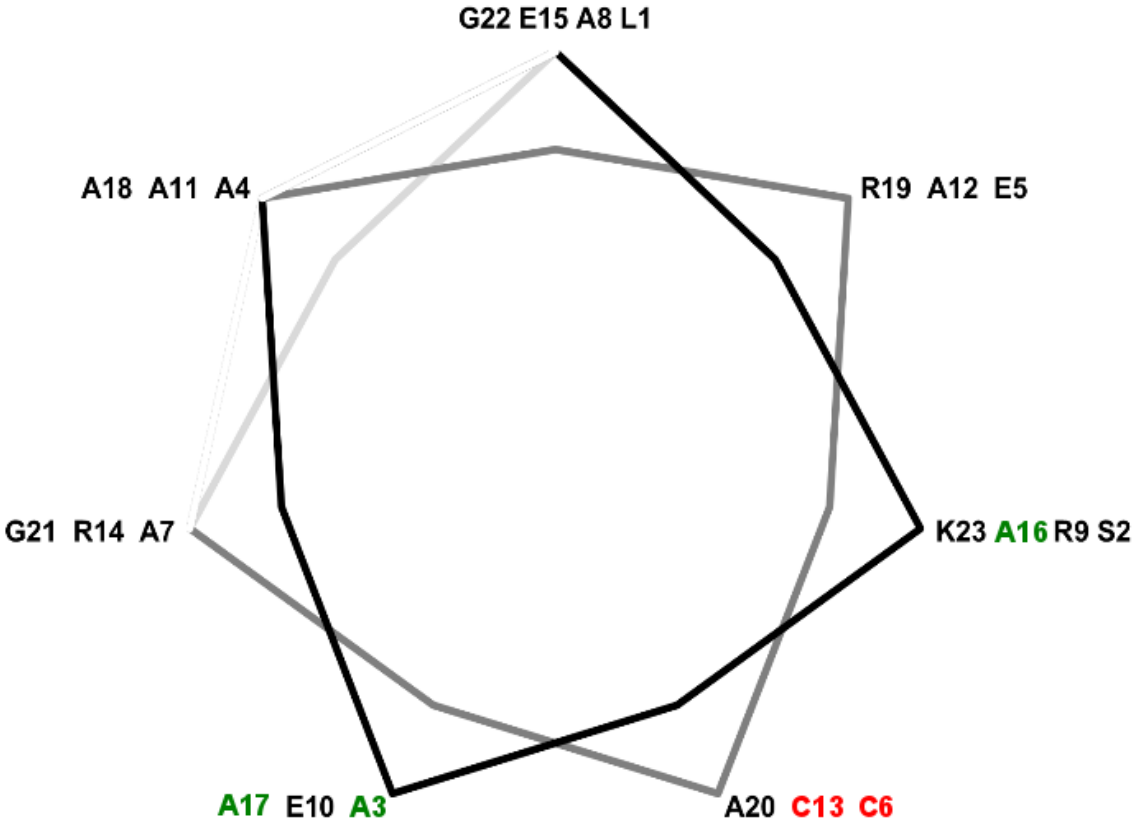

Figure S4. Helical wheel projection of dC10α helix. The first turn contains residues L1, S2, A3 and A4 (*black line*), the second turn contains residues E5, C6, A7 and A8 (*dark grey line*), etc. Reactive cysteines C6 and C13 are highlighted (*red*) as well as mutation sites A3, A16 and A17 (*green*).

Predicted Helicity

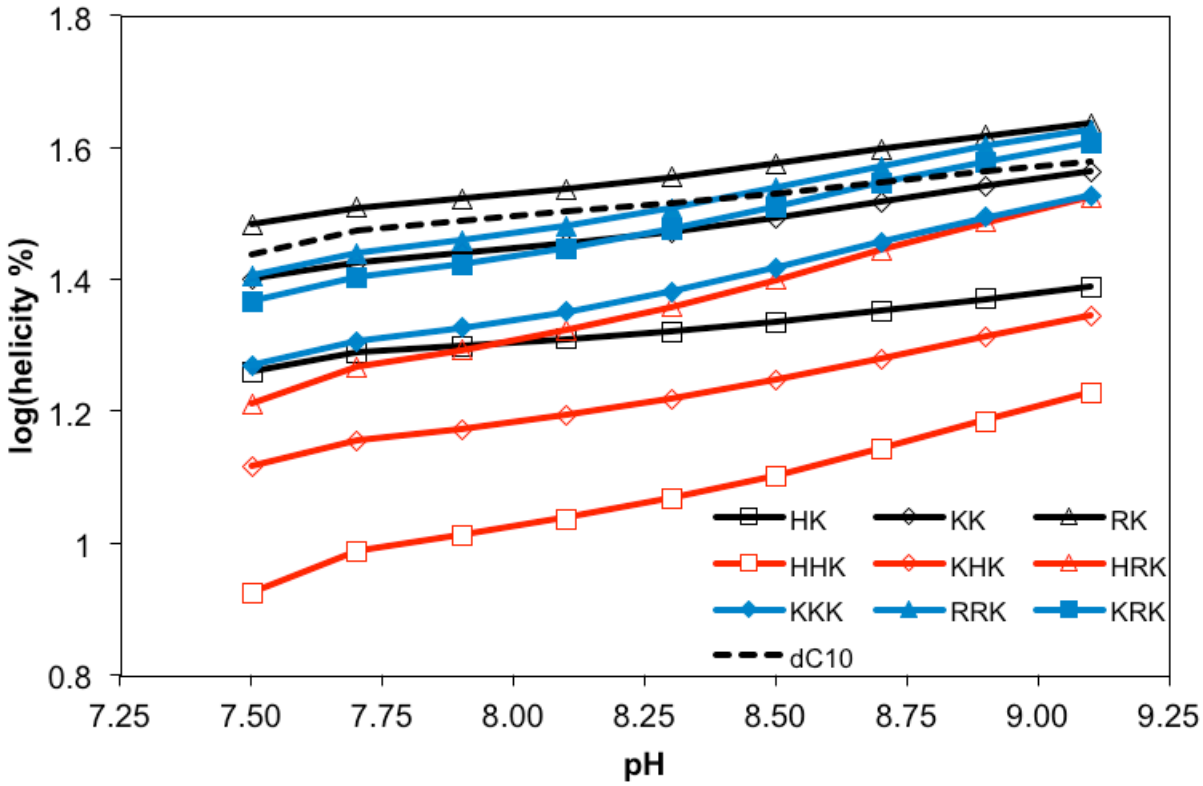

Figure S5. Predicted helicity of dC10α peptides. Helicity of dC10α (or variant) peptide sequences was predicted using the AGADIR algorithm (http://agadir.crg.es/)[5,6,7] at 301 K, ionic strength 0.1 M. Slopes range from 0.07 to 0.18.

Effect of Thiol pK_a on Reaction Rate

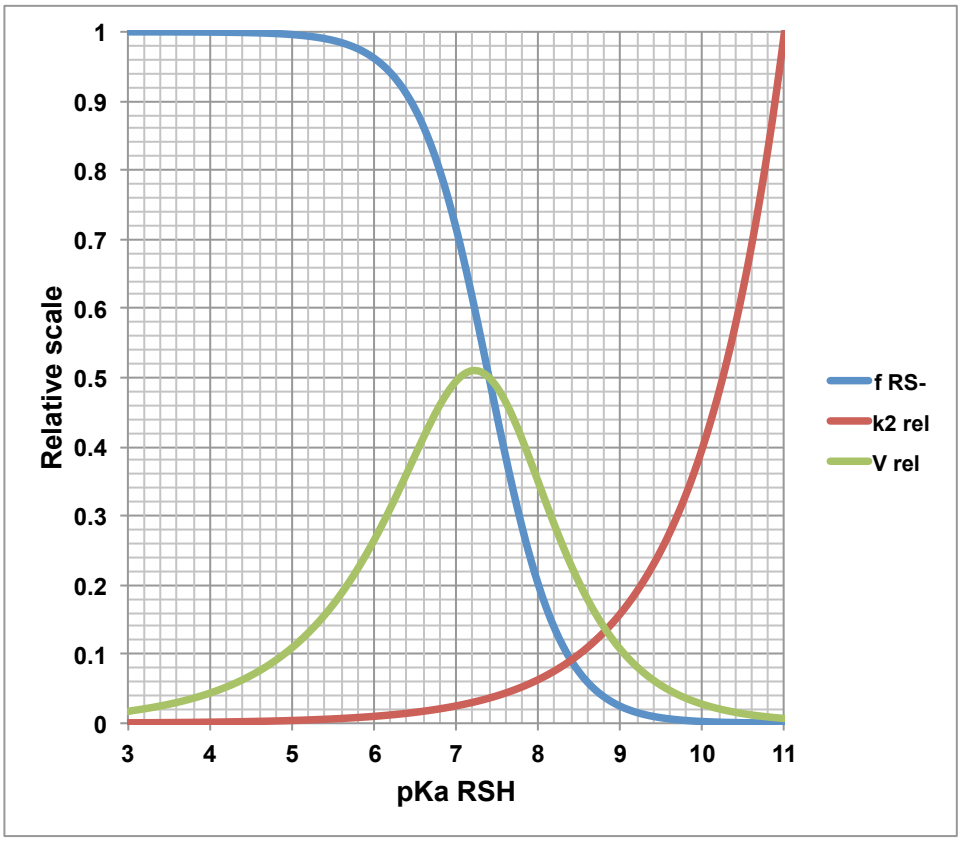

Figure S6. Modelled effect of thiol pK_a on reaction rate. The blue line shows the effect of changing thiol pK_a on the proportion of thiolate, the reactive species,[8,9,10] that is present at pH 7.4. The red line shows the modelled effect on thiolate nucleophilicity, reflected in the second order rate constant k_2, assuming a Brønsted β value of 0.4.[8,9] The line shown in green is the predicted relative rate, calculated as a product of rate constant and reactant concentration. Note the increase in predicted reaction rate, as thiol pK_a is decreased from that of the parent Cys residue.

Labelling Stoichiometry

Test protein MBP bearing the 'KRK' variant of the re-engineered dC10α tag (aka MBP-dC10*) was reacted with varied concentrations of fluorogen **3**, and fluorescence intensity of the labelled protein was plotted as a function of molar equivalents of **3**.

Specifically, solutions of 50 μM MBP-dC10*, 0-87.5 μM **3**, 10% DMSO, 1 mM TCEP, and 50 mM HEPES (pH 7.4) were incubated for up to 14 h at 37 °C. Emission spectra were recorded of each sample 14 h (to ensure complete reaction of all samples) using an excitation laser of 438 nm (Figure S7). The fluorescence intensity at 485 nm was then plotted against the number of molar equivalents of fluorogen 3 (Figure S8).

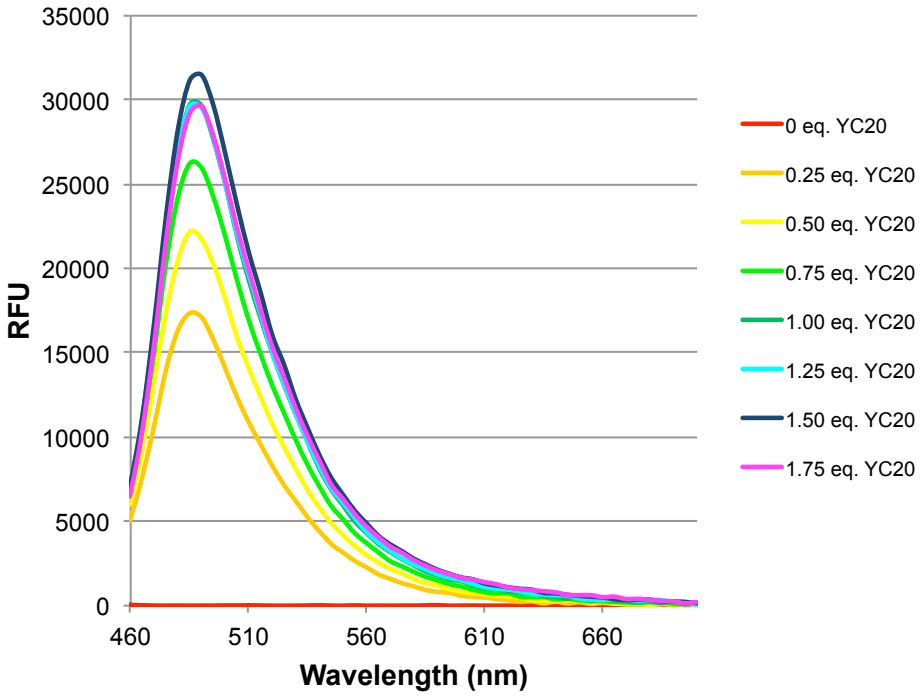

Figure S7. Emission spectra (λ_{exc} = 438 nm) recorded of 50 μM MBP-dC10* after reaction with varied equivalents of fluorogen **3**, for 14 h at 37 ºC.

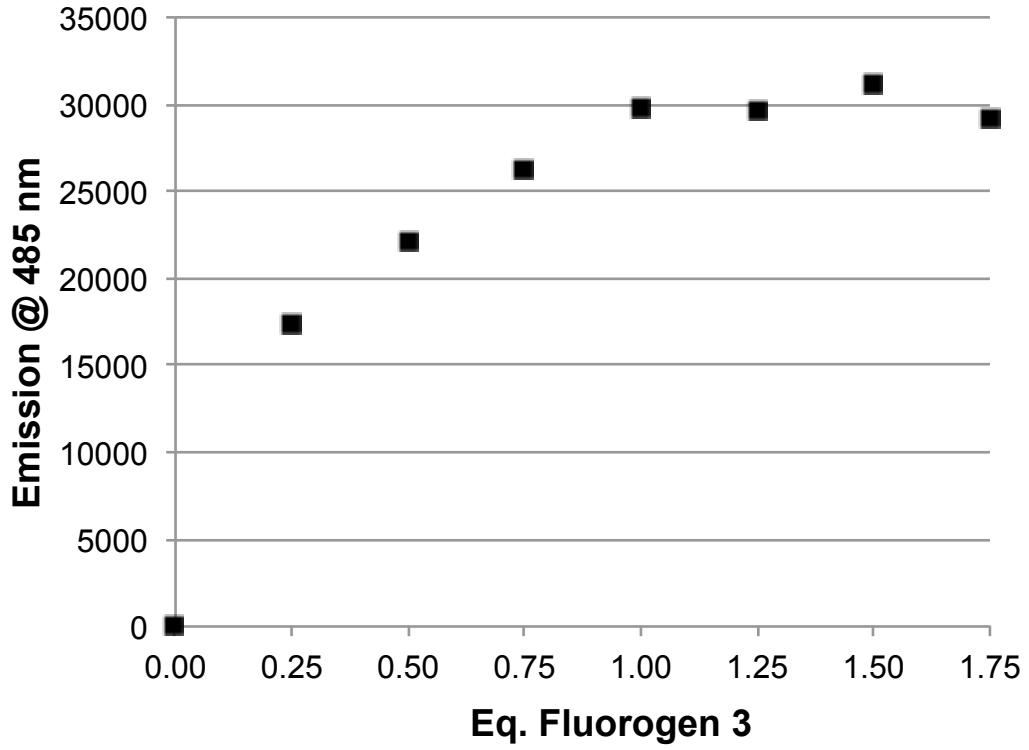

Figure S8. Fluorescence intensity of 50 μM MBP-dC10* as a function of the number of molar equivalents of fluorogen **3**, after labelling for 14 h at 37 ºC (see **Figure S7**, λ_{em} = 485 nm). Note the roughly linear increase in fluorescence, up to the addition of one equivalent of labelling agent, as expected.

Labelling Selectivity

Purified MBP-dC10* (final concentration 1 mg/mL) was added to the soluble fraction of *E. coli* lysate (final concentration 10 mg/mL) in 50 mM HEPES, pH 7.4. To this protein mixture was added TCEP (1 mM final concentration) and fluorogen **3** in DMSO to a final concentration of 10 µM **3** and 10% DMSO. This sample was incubated at 37 ºC for 45 min and then mixed in a 1:1 ratio with Tris-tricine loading buffer without reducing agent (100 mM Tris pH 6.8, 24% (v/v) glycerol, 3.5% (w/v) SDS, 0.01% (w/v) Coomassie brilliant blue) and loaded onto a SDS-PAGE gel. After migration, band fluorescence was evaluated using a GelDoc equipped with a blue excitation laser and a 530/30 nm emission filter. After exposure, gels were stained with Coomassie brilliant blue.

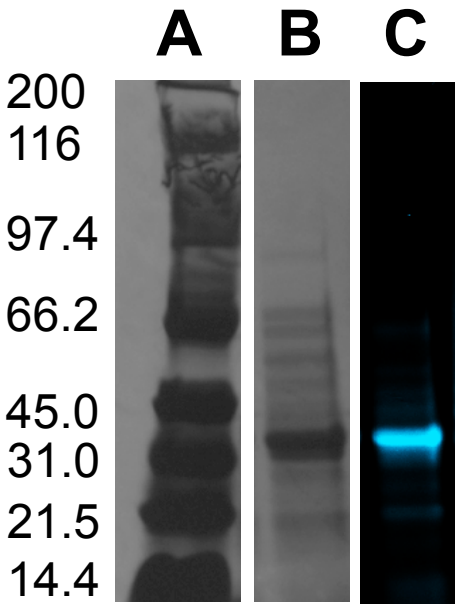

Figure S9. SDS-PAGE gel of MBP-dC10* labelled with 10 µM fluorogen **3** in bacterial lysate. A) Molecular weight markers. B) Coomassie staining. C) Fluorescence imaging, showing only one unambiguously labelled protein, corresponding to the expected mass of MBP-dC10*.

Intracellular Labelling

In addition to the labelling shown in **Figure 8**, a labelling experiment was attempted over a shorter incubation time, using the more reactive tag sequence dC10*.

The procedure for this labelling experiment was the same as reported in the Material and Methods section 2.5, except that cells were incubated for only 30 min before imaging.

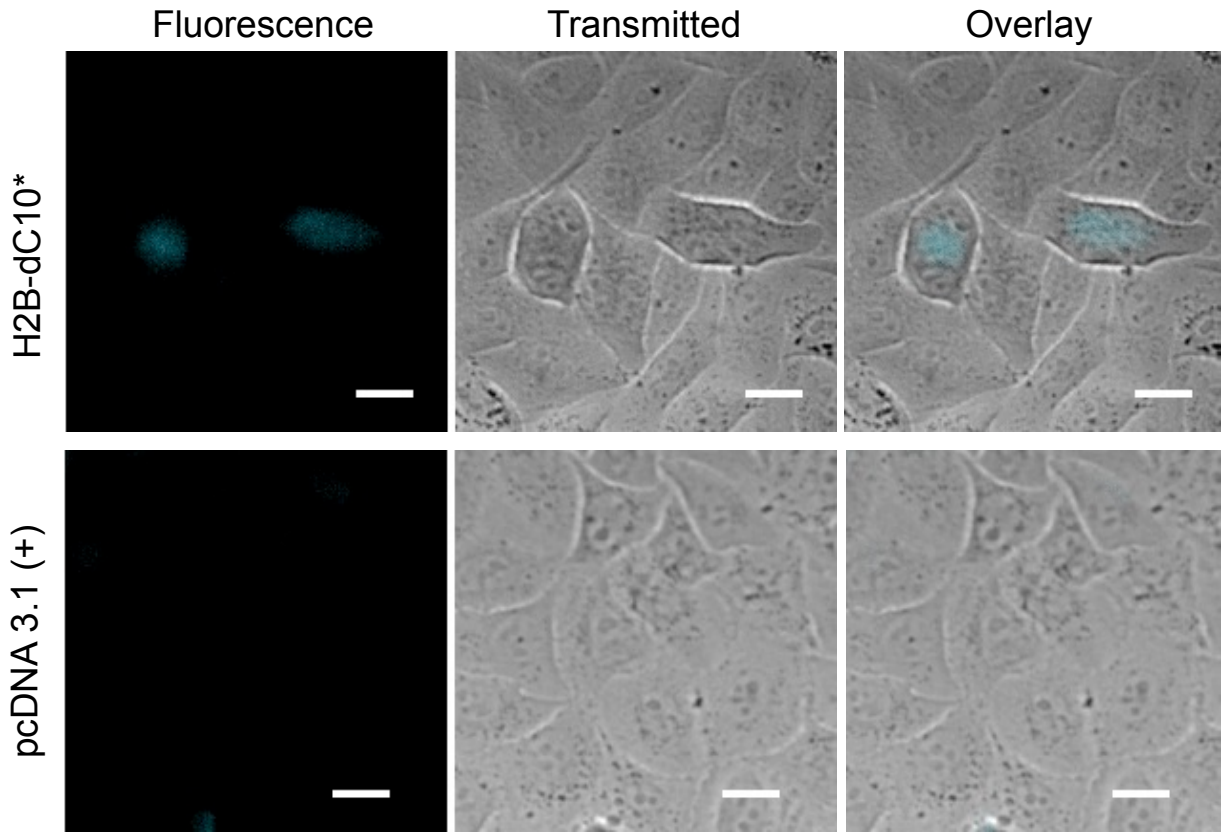

Figure S10. Intracellular labelling of histone H2B. The histone H2B-dC10* protein was labelled with 10 µM of fluorogen **3** over only 30 min in living HeLa cells. Fluorescence was visible in HeLa cells expressing the target protein (excitation laser: 438 nm; emission filter: 485/25 nm) whereas cells transfected with empty plasmid pcDNA3.1(+) show negligible background fluorescence. White scale bars are 20 µm.

References

[1] C. N. Pace and J. M. Scholtz, *Biophys. J.*, 1998, **75**, 422–427.

[2] R. Parthasarathy, S. Subramanian and E. T. Boder, *Bioconjugate Chem.*, 2007, **18**, 469–476.

[3] W. M. Haynes, Ed., *Handbook of Chemistry and Physics*, CRC Press, 2016.

[4] M. W. Popp, J. M. Antos, G. M. Grotenbreg, E. Spooner and H. L. Ploegh, *Nat. Chem. Biol.,* 2007, **3**, 707–708.

[5] V. Muñoz and L. Serrano, *Nat. Struct. Mol. Biol.,* 1994, **1**, 399–409.

[6] V. Muñoz and L. Serrano, *J. Mol. Biol.,* 1995, **245**, 275–296.

[7] V. Muñoz and L. Serrano, *J. Mol. Biol.,* 1995, **245**, 297–308.

[8] T. Sekine, K. A. Kato, K. Takamori, M. Machida and Y. Kanaoka, *Biochim. Biophys. Acta*, 1974, **354**, 139–147.

[9] R. A. Bednar, *Biochemistry*, 1990, **29**, 3684–3690.

[10] Y. Chen, K. Tsao, É. De Francesco and J. W. Keillor, *J. Org. Chem.*, 2015, **80**, 12182–12192.