Formal reduction potentials of difluorotyrosine and trifluorotyrosine protein residues: Defining the thermodynamics of multistep radical transfer

Kanchana R. Ravichandran,† Allan B. Zong, § Alexander T. Taguchi,† Daniel G. Nocera,[¶] JoAnne Stubbe,^{†,‡} and Cecilia Tommos^{§,*}

† Department of Chemistry and ‡ Department of Biology, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, United States. § Department of Biochemistry and Biophysics, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania 19104, United States. ¶ Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, MA 02138, United States.

Corresponding author: tommos@upenn.edu

1. Materials. 2,6-difluorophenol, 2,3-difluorophenol, 2,3,6-trifluorophenol, 2,3,5-trifluorophenylboronic acid, kanamycin (Kan), chloramphenicol (Cm), 2XYT medium powder, L-arabinose and isopropyl-β-Dthiogalactoside (IPTG), CDP, ATP, HEPES, MgSO4 and EDTA were purchased from Sigma Aldrich and ultra pure urea from Nacalai Tesque Inc., Kyoto, Japan. Wilmad-LabGlass provided the X-band EPR tubes (706-PQ). The 1 mm (21-Q-1), 2 mm (21-Q-2) and 10 mm (29B-Q-10) path length quartz cuvettes were purchased from Starna Cells.

2. Construction of the pE-SUMO- α_3TAG_{32} plasmid. α_3TAG_{32} was amplified from the previously described pET32b- α_3 TAG₃₂ plasmid (Ravichandran et al. *Biochemistry* 2013, 52, 8907–8915) using the forward primer 5′-ccgcgaacagattggaggtGGATCCCGTGTG-3′ and the reverse primer 5′-cgaattcggatcctc tagtcatTTACAGTTTTTTAATTTCTTC-3ʹ. The nucleotides shown in lower case overlap with the pE-SUMO vector (Integrated DNA technologies) and the nucleotides shown in upper case overlap with α_3 TAG₃₂. The gene was inserted into the pE-SUMO vector using the Gibson isothermal assembly kit (New England Biolabs, Inc.) following the manufacturer's protocol (Gibson et al. *Nat. Methods.* **2009**, *6*, 343–345.). The construct was confirmed by sequencing at the MIT Biopolymers Laboratory.

Figure S1. SDS-PAGE analysis of His₆-Smt3- α_3F_nY *expression at optimized conditions.* SDS-PAGE gels (15%) display in lane (1) molecular weight markers, lane (2) pre-induction sample and in lane (3) expression of His₆-Smt3- α_3F_nY in the presence of (panel A) 3,5-F₂Y, (B) 2,3,5-F₃Y, (C) 2,3-F₂Y and (D) 2,3,6-F₃Y. The calculated molecular weight of truncated His₆-Smt3- α_3 (residues 1–31) is 15971 Da. The calculated molecular weights of full-length His_6 -Smt3- α_3F_nY are 19949 Da (F₂Y at position 32) and 19967 Da $(F_3Y$ at position 32) respectively. The red arrow indicates the full-length His_6 -Smt3- α_3 F_nY protein in each panel.

*3. Expression of the His6-Smt3-*α*3FnY fusion proteins, small-scale optimization.* pE-SUMO-α3TAG32 was cotransformed with a plasmid encoding the fluorotyrosine tRNA and aminoacyl-tRNA synthetase (FnY-RS) genes (pEVOL-FnY-RS-E3; Minnihan et al. *J. Am. Chem. Soc.* **2011**, *133*, 15942–15945) into BL21(DE3) chemically competent cells and grown overnight (37 °C) on LB plates supplemented with Kan (50 μ g/mL) and Cm (35 μ g/mL). F_nY incorporation and protein expression were optimized for each α_3F_nY variant. In a series of small-scale test expressions (100 mL 2XYT cultures), the following variables were tested for their effects on the expression of His_6 -Smt3- α_3F_nY in *E. coli* BL21(DE3): F_nY concentration, L-arabinose and IPTG concentrations, and induction times. Expression of $\alpha_3(3,5)F_2Y$ and $\alpha_3(2,3,5)F_3Y$ were optimal with 1 mM F_nY , whereas the concentrations of 2,3-F₂Y and 2,3,6-F₃Y were increased to 1.5 mM for expression of $\alpha_3(2,3)F_2Y$ and $\alpha_3(2,3,6)F_3Y$, respectively. Expression of His₆-

Smt3- α_3F_nY was not affected by variations in the concentration of L-arabinose (0.05–0.1%) or IPTG $(0.5-1 \text{ mM})$. Thus, concentrations of 0.05% (w/v) and 0.5 mM were used for L-arabinose and IPTG respectively. Finally, expression of His_6 -Smt3- α_3F_nY was most sensitive to the timing of L-arabinose and IPTG addition. Optimized expression of His₆-Smt3- $\alpha_3F_nY_s$ required early addition of L-arabinose $(OD₆₀₀$ of 0.1–0.2) and late addition of IPTG (OD₆₀₀ of 0.9) to the growth medium. Figure S1 shows the SDS-PAGE expression gel for each His_{6} -Smt3- $\alpha_{3}F_{n}Y$ under these optimized conditions.

4. Expression of the His₆-Smt3- α_3F_nY fusion proteins, preparative scale. A 5 mL LB culture containing Kan (50 μ g/mL) and Cm (35 μ g/mL) was started from a single colony and grown for ~12 h at 37 °C. This culture was used to inoculate (100-fold dilution) 100 mL of 2XYT medium containing Kan and Cm. After 12 h at 37 °C, the culture was diluted 100-fold into 4 x 2 L of 2XYT medium containing the antibiotics and 1–1.5 mM F_nY (500 mM stock solution prepared in water, NH₄OH solubilized). At an OD₆₀₀ of 0.1–0.2, 0.05% (w/v) L-arabinose was added to induce F_nY-RS . IPTG (0.5 mM) was added at an OD₆₀₀ of 0.9 to induce the expression of His₆-Smt3- α_3F_nY . Cells were harvested after an additional 5 h of growth at 37 °C. A typical yield of 2.5 g cell paste/L culture was obtained.

5. $\alpha_3 F_n Y$ protein purification protocol. Cell paste (10–20 g) was re-suspended (5 mL/g cell paste) in buffer A (20 mM Tris pH 7.9, 500 mM NaCl, 1 mM phenylmethylsulfonyl fluoride) and lysed via a single passage through the French pressure cell (14,000 psi). Cell debris was removed by centrifugation (20,000 g, 15 min, 4 °C) and the supernatant was loaded onto 30 mL of Ni-NTA agarose (Qiagen) equilibrated with buffer A. The column was washed with 20 column volumes (CVs) of buffer A containing 20 mM imidazole. Bound His_{6} -Smt3- $\alpha_{3}F_{n}Y$ was eluted with buffer A containing 250 mM imidazole. Protein containing fractions were identified by the Bradford assay, exchanged into buffer A and concentrated to \sim 5 mg/mL using an Amicon ultrafiltration cell with a YM10 membrane. The His₆-Smt3 tag was removed by incubating the fusion protein with SUMO protease (1:500 protease:Smt3- α_3F_nY) and 5 mM DTT for 1 h at 25 °C. The digestion mixture was passed through a nickel column (30 mL of Ni-NTA agarose) to remove the SUMO protease, $His₆$ -Smt3 and any undigested material. The column was washed with 2–3 CVs of buffer A containing 20 mM imidazole. The α_3F_nY containing fractions were identified by SDS-PAGE, pooled, and purified by reversed-phase HPLC. All proteins except $\alpha_3(2,3)F_2Y$ were isolated on a semi-preparative C18 column (Grace VYDAC 218TP, particle size 10 mm, column size 10×250 mm; sample injection volume 5–10 mL) using a linear water/acetonitrile/0.1% (w/v) trifluoroacetic acid gradient (35–48% acetonitrile over 30 min, flow rate 5 mL/min). $\alpha_3(2,3)F_2Y$ was isolated on an analytical C18 column (Grace VYDAC 218TP, particle size 5 mm, column size 4.6×250) mm; injection volume 0.5 mL; gradient 20–50% acetonitrile over 50 min, flow rate 1.0 mL/min). As shown below, sample purity was evaluated using the analytical C18 column (20–70% acetonitrile over 100 min, flow rate 1 mL/min). Molecular weights were verified by matrix-assisted desorption/ionization-time of flight (MALDI-TOF) mass spectrometry using a Bruker Microflex 3.1. The purified proteins were stored at room temperature as lyophilized powders.

6. Spectroscopic measurements

Absorption spectroscopy. The pH titrations were performed using an Evolution 300 UV-Vis spectrophotometer equipped with a Smart Peltier thermostatted cell holder. Samples were prepared by gradually dissolving lyophilized protein in a 10 mm path length quartz cuvette containing 0.9 mL 20 mM sodium acetate, 20 mM potassium phosphate, 20 mM sodium borate (APB buffer), pH 10.3 to a 278 nm absorption ~ 0.6 –0.7. A portion of the protein solution (667 µl) was removed from the cuvette and diluted in 20 mM APB, pH 1.6 buffer to a final volume of 2.0 mL. The remaining protein solution was diluted in 20 mM APB, pH 10 buffer to a final volume of 700 μ L. This generated two samples with a 278 nm absorption ~ 0.2 and a final pH of ~ 5 and 10.3, respectively. The low pH buffer was pre-adjusted to generate a pH \sim 5 upon addition of the protein stock. The pH studies were performed by manual equalvolume (0.7 mL) titration using the following settings: spectral range 240 – 500 nm, scan rate 120 nm min⁻¹, bandwidth 2 nm, step size 0.2 nm, and temperature 23.5 °C. The pH was measured before and

after each data recording. The apparent Y(O⁻/OH) pK_a (pK_{app}) was obtained by fitting the pH-induced change in $(278 - 340 \text{ nm}$ baseline) \triangle Abs to a single p K_a .

CD spectroscopy. Data were collected on an Aviv 202 CD spectrometer equipped with a thermostatted cell holder and an automated titrator system for concentration and pH-dependence studies. Protein stability measurements were conducted by dissolving lyophilized protein in 20 mM sodium acetate, 20 mM potassium phosphate (AP buffer), pH 5.0 or 5.5, to a raw 222 nm ellipticity (θ_{222}) in the –215 \pm 45 millidegree range (1 mm path). The protein stock solution was added to 20 mM AP buffer containing 0 and 9.5 M urea, respectively. The 9.5 M stock solutions were freshly made using ultra pure urea. The pH was matched to 5.02 ± 0.02 or 5.51 ± 0.02 in the protein/AP buffer and protein/AP buffer/urea solutions. The protein dilution was 16.5-fold generating a final θ_{222} reading of -130 ± 30 millidegree (10) mm path) at zero molar denaturant. The urea denaturation experiments were performed by automated equal-volume (2.0 mL) titration. Settings: titration steps 0.1 M, equilibrium time 30 s (0–5 M urea) and 60 s (5–8 M urea), wavelength 222 nm, bandwidth 3 nm, averaging time 60 s, and temperature 25 °C.

The remaining pH 5.0 and 5.5 stock samples were combined, additional protein powder added if required, and used to measure the degree of α -helix as a function of pH. The pH titrations were conducted by diluting equal portions of protein stock into APB buffer generating two samples with pH \sim 4.9 and \sim 12, respectively. The pH titrations were performed by automated equal-volume (2.0 mL) titration. Settings: starting pH \sim 4.9, ending pH 10, interval 0.15 pH units, pH dead band 0.05, wavelength 222 nm, bandwidth 3 nm, averaging time 60 s, and temperature 25 °C. Far-UV CD control spectra were collected before and after each titration series.

To determine the absolute α -helical content, CD spectra were collected from $\alpha_3 W$ (reference protein), $\alpha_3(2,3)F_2Y$, $\alpha_3(2,3,5)F_3Y$ and $\alpha_3(2,3,6)F_3Y$ dissolved in 40 mM sodium acetate, pH 5.1. The CD samples were matched to an θ_{222} of -85 ± 10 millidegree (2 mm path). Settings: spectral range 200–260 nm, step size 0.25 nm, time average 1.5 s, number of scans 2, and temperature 25 °C.

Fluorescence spectroscopy. The NanoOrange protein quantitation method (Molecular Probes) was employed to determine the protein concentration in the CD samples used to measure the absolute α -helical content. A 0–4 mg protein/mL standard curve was prepared by dissolving $\alpha_3 Y$ (ε_{277} 1490 M⁻¹ cm⁻¹; molecular weight 7520 g mol⁻¹) in 40 mM sodium acetate, pH 5.1. Fluorescence spectra were collected on a Horiba Jobin Yvon Spex Fluorolog spectrofluorometer equipped with a thermostatted cell. Settings: $\lambda_{\rm ex}$ 485 nm, $\lambda_{\rm em}$ range 585–600 nm, slit width excitation 2 nm, slit width emission 2 nm, step size 0.5 nm, time average 0.5 s, number of scans 3, and temperature 25 °C.

X-band EPR spectroscopy. $Y_{731}F-\alpha$ 2 (25 mM), CDP (1 mM) and ATP (3 mM) in 50 mM HEPES pH 7.6, 15 mM MgSO₄, 1 mM EDTA were incubated in a Lauda RM6 circulating water bath set at 15 °C, 25 °C or 37 °C. Each reaction was initiated with $2.3-F_2Y_{122}$ •-β2 (25 mM, 0.85 $F_2Y_2/82$). The samples were transferred to X-band EPR tubes maintained in the water bath, and quenched in liquid isopentane (–140 °C) within 20 s. Settings: microwave frequency 9.45 GHz, power 30 mW, modulation amplitude 1.50 G, modulation frequency 100 kHz, time constant 40.96 ms, conversion time 20.48 ms, and temperature 77 K.

Figure S2. Analytical HPLC and MALDI-TOF evaluation of purified α*3FnY.* Analytical C18 reversedphase chromatograms of purified (A) $\alpha_3(2,3)F_2Y$, (B) $\alpha_3(2,3,5)F_3Y$, and (C) $\alpha_3(2,3,6)F_3Y$. The purified freeze-dried proteins were dissolved in 20 mM sodium acetate, pH 5.5, 100 μ L loaded, and eluted with a linear 20–70% acetonitrile/100 min gradient. Panel (D) shows MALDI-TOF traces of purified $\alpha_3(2,3)F_2Y$ (orange) and α_3Y (grey). Each trace displays a single major peak, whose maxima are separated by 37 \pm 1 Da. This is consistent with the exchange of \overline{Y}_{32} to F_2Y_{32} (calculated $\Delta m/z = 36$ Da). MALDI-TOF traces of purified (E) $\alpha_3(2,3,5)F_3Y$ (green) and (F) $\alpha_3(2,3,6)F_3Y$ (red) are also shown relative to the $\alpha_3 Y$ (grey) trace. For the $\alpha_3 F_3 Y_s$, the peak separation increases to (E) 54 \pm 1 Da and (F) 53 \pm 1 Da, which is consistent with the exchange of \overline{Y}_{32} to F_3Y_{32} (calculated $\Delta m/z = 54$ Da). Equivalent data for purified $\alpha_3(3,5)F_2Y$ are shown in Ravichandran et al. *Biochemistry* 2013, 52, 8907–8915.

Figure S3. Urea-induced denaturation of $\alpha_3(2,3,6)F_3Y$ *at (A) pH 5.0 and (B) pH 5.5.* **The chemical** melts were conducted by automated titration, as described in the Materials and Methods section, and the quality of the data is typical for the $\alpha_3 X$ proteins. Data points are displayed as red circles and the grey lines represent nonlinear curve fits to determine the global stability of the protein in the absence of denaturant (Santoro & Bolen *Biochemistry* **1988**, *27,* 8063–8068). The global stability values of α3Y and the $\alpha_3 F_n Y$ proteins are summarized in Table S1 below.

Protein	$[Q]_{222} \times 10^3$ (deg cm ² dmol ⁻¹)	Helical residues ^{<i>d</i>}	Protein stability (kcal mol ⁻¹)
$\alpha_3 Y$	-20.0 ± 0.8 (pH 5.6) ^b	52 of 65 (pH 5.6) ^{c}	-3.77 ± 0.08 (pH 5.0 - 8.5) ^{b,d,e}
$\alpha_3(2,3)F_2Y$	-21.0 ± 0.9 (pH 5.0)	53 ± 2 (pH 5.0)	not determined
$\alpha_3(3,5)F_2Y$	-20.9 ± 0.6 (pH 5.6) ^b	53 ± 2 (pH 5.6)	-3.67 ± 0.02 (pH 5.0) ^b
			-3.68 ± 0.02 (pH 5.5) ^b
$\alpha_3(2,3,5)F_3Y$	-20.5 ± 0.2 (pH 5.0)	52 ± 1 (pH 5.0)	-3.83 ± 0.20 (pH 5.0)
			-3.85 ± 0.16 (pH 5.5)
$\alpha_3(2,3,6)F_3Y$	-21.1 ± 0.7 (pH 5.0)	53 ± 2 (pH 5.0)	-4.20 ± 0.06 (pH 5.0)
			-4.22 ± 0.06 (pH 5.5)

Table S1. CD and structural properties of the $\alpha_3 F_n Y$ *proteins*

(a) Determined using the reference CD spectrum (Figure 3A, main manuscript) of the structurally characterized α3W protein; PDB ID 1LQ7, Dai et al. *J. Am. Chem. Soc.* **2002**, *124,* 10952–10953; (*b*) Ravichandran et al. *Biochemistry* **2013**, *52,* 8907–8915; (*c*) Determined from the solution NMR structure of α3Y, PDB ID 2MI7, Ref. *e*; (*d*) Martinéz-Rivera et al. *J. Am. Chem. Soc.* **2011**, *133,* 17786–17795. (*e*) Glover et al. *J. Am. Chem. Soc.* **2014**, *136,* 14039–14051.

Figure S4. Square-wave voltammetry (SWV) quasi-reversible maxima of α*3Y and 2-mercaptophenol*^α*3C (2MP-*α*3C).* The plots were obtained by analyzing data originally presented in Berry et al. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 9739–9743 and in Tommos et al. *Biochemistry* **2013**, *52*, 1409–1418.

Figure S5. Representative ^α*3Y protein film square-wave voltammograms.*

Figure S6. Representative $\alpha_3(2,3)F_2Y$ *protein film square-wave voltammograms.*

7. Data processing of ^α*3X protein film square-wave voltammograms.* Figures S5 and S6 illustrate the processing routines used for voltammetry data analysis. Briefly, raw net, forward and reverse voltammograms were smoothened lightly (using a Loess function with $a \sim 10$ mV smoothening window) and the non-Faradaic baseline current was cropped. The resulting waveforms (shown in grey) were fitted with a cubic baseline (dotted blue lines) and an asymmetric Gaussian (red lines) using PeakFit (systatsoftware.com). The background-corrected voltammograms shown in Figure 4 in the main manuscript were obtained by subtracting the calculated baselines from the smoothened voltammograms. PeakFit generates the potential and the amplitude at the peak maximum along with the width at half height from the calculated lineshapes.

Figure S7. Representative α*3Y protein film net square-wave voltammograms.* The backgroundcorrected voltammograms were obtained at pH 6.9 (green), 8.4 (red) and 9.9 (blue). SWV settings: Equilibration time 5 s, step potential 0.15 mV, SW pulse amplitude 25 mV, SW frequency 240 Hz.

Table S2. Nernst equations and parameters used to model the Pourbaix diagrams in Figure 5

Protein	$\mathbf{p}K_{\text{app}}$	$E^{\circ}(\text{pH 0})$ (V)	E° Y ₃₂ (O•/O ⁻)(V)	E° (pH 7.0) (V)
$\alpha_3(3,5)F_2Y$	8.0	1.346	0.901	0.958
α_3 Y	11.3	1.372	0.749	0.986
$\alpha_3(2,3,5)F_3Y$	7.2	1.412	1.008	1.032
$\alpha_3(2,3)F_2Y$	8.6	1.447	0.965	1.057
$\alpha_3(2,3,6)F_3Y$	7.9	1.509	1.070	1.122

The Pourbaix diagrams shown in Figure 5 in the main manuscript were modeled by including two protonation states for the reduced species (Y₃₂–OH and Y₃₂–O[–]) and one for the oxidized species (Y₃₂–O•) (Clarke, W. M. Oxidation-reduction potentials of organic systems, 1960, The Williams & Wilkins Company, Waverly Press Inc., Baltimore, MD). The $Y_{32}-O\bullet^+$ state in not included since the p K_a of aqueous Y–O•⁺ is below pH 0 (\sim -2; Dixon, W. T.; Murphy D. *J. Chem. Soc. Faraday Trans. II* **1976**, 72, 1221–1230).

Either of the following two equations applies:

$$
E^{\circ}(\text{pH}) = E^{\circ}(\text{pH } 0) + m \times \log(10^{-\text{pH}} + 10^{-\text{pKred}})
$$

$$
E^{\circ}(\text{pH}) = E^{\circ}(\text{Y}_{32}(\text{O} \bullet/\text{O}^{-})) + m \times \log(1 + 10^{-\text{pH}}/10^{-\text{pKred}})
$$

where $E^{\circ}(\text{pH 0})$ is the extrapolated potential at pH 0, *m* is 2.303*RT/nF* which equals 0.0552 for $\alpha_3 Y$ and 0.0557 for the four $\alpha_3 F_n Y$ proteins, and p K_{red} represents the p K_{app} of the reduced Y₃₂ and F_nY₃₂ residues. Table S2 also displays *E*°' at pH 7.0 as a general reference.

Figure S8. Current thermodynamic landscape of the RT pathway in E. coli class Ia RNR. The overall pathway is proposed to be > 200 mV uphill and driven forward by rapid irreversible chemistry that occurs in the active site of $\alpha 2$.¹ The $\Delta E^{\circ}(Y_{356} - Y_{731}) = -100 \text{ mV}$, pH 7.6), $\Delta E^{\circ}(2,3,5-F_3Y_{122}) = -Y_{356}$; 20 ± 10 mV, pH 7.6)¹ and $\Delta E^{\circ}(Y_{122} - 3.5 - F_2Y_{356}) = 70 \pm 5$ mV, pH 8.2)¹ values were previously reported, whereas $\Delta E(2,3-\mathrm{F}_2\mathrm{Y}_{122} - \mathrm{Y}_{356}^{\bullet}; 47 \pm 11 \text{ mV}, \text{pH } 7.6)$ was measured by equilibration studies with 2,3-F₂Y₁₂₂•-β2 and Y₇₃₁F- α 2, as shown below in Figure S9.

Figure S9. Temperature dependence of Y₃₅₆• formation in the reaction of $F_nY \cdot B2$ (n = 2 or 3), Y₇₃₁Fa2, CDP and ATP. The reaction mixture contained the RT pathway blocked $Y_{731}F-\alpha^2$ mutant, CDP (substrate), ATP (effector) and either $2,3,5-\frac{F_3Y_{122}}{P_2P_122}$ (green) or $2,3-\frac{F_2Y_{122}}{P_2P_122}$ (orange). Samples were incubated for 20 s in a water bath set at the desired temperature and quenched in liquid isopentane (–140 °C) for analysis by X-band EPR spectroscopy. From each composite EPR spectrum, residual 2,3,5-F₃Y₁₂₂•-β2 or 2,3-F₂Y₁₂₂•-β2 was subtracted as previously described¹ to quantitate the percentage of Y₃₅₆• shown. The equilibrium constant (K_{eq}) for the 2,3F₂Y₁₂₂• \Leftrightarrow Y₃₅₆• reaction was calculated from the ratio of $[Y_{356} \cdot]/[2,3F_2Y \cdot]$ and a $\Delta E^{\circ}(2,3-F_2Y_{122} \cdot -Y_{356} \cdot)$ of 47 ± 11 mV (pH 7.6, 25 °C) was determined using the equation $\Delta E^{\circ} = RT \ln K_{eq} / nF$. $\Delta E^{\circ}(2,3,5-F_3Y_{122} - Y_{356})$ was 20 ± 10 mV (pH 7.6, 25) $^{\circ}$ C).¹

(1) Ravichandran, K. R.; Taguchi, A. T.; Wei, Y.; Tommos, C.; Nocera, D. G.; Stubbe, J. *J. Am. Chem. Soc.*, **2016**, *138*, 13706. (2) Yokoyama, K.; Smith, A. A.; Corzilius, B.; Griffin, R. G.; Stubbe, J. *J. Am. Chem. Soc.* **2011**, *133*, 18420.