Supplementary Information

A second-generation expression system for tyrosine sulfated proteins and its application in crop protection

Experimental Details

Construction of second-generation sY protein expression vectors

pULTRA-sY and pEVOL-sY were cloned by replacing the *Methanocaldococcus janaschii* (*Mj*) tyrosyl-tRNA synthetase (*Mj*-TyRS) in pULTRA[1] and pEVOL[2] with STyrRS, the *Mj* aminoacyl-tRNA synthetase (aaRS) evolved to specifically charge *Mj*-$tRNA_{CUA}$ with sY. *Mj*-$tRNA_{CUA}$ suppresses UAG codons.[3] To clone pULTRA-sY, STyrRS was amplified from pCDF-sY[4] using primers aaRStopULTRAPCR1 and aaRStopULTRAPCR2, and subsequently inserted into pULTRA using the NotI cloning site. To clone pEVOL-sY, two pEVOL backbone fragments were amplified from pEVOL using primer pairs pEVOL-BB-F/pEVOL-BB-R and pEVOL-BB2-F/pEVOL-BB2-R, respectively. Two STyrRS fragments were amplified from pULTRA using primer pairs aaRS-F/aaRS-R and aaRS2-F/aaRS2-R, respectively. pEVOL was constructed by assembling the pEVOL and STyrRS fragments using the Gibson method.[5]

Cloning, expression, and analyses of GFP

The pGLO expression vector was used to express wild-type GFP (pGLO-GFP-WT) or GFP with one or three UAG codons. pGLO-GFP-1UAG was cloned via site-directed mutagenesis PCR from pGLO-GFP-WT using primers pGLO-GFP-1TAG-F and pGLO-GFP-1TAG-R. pGLO-GFP-3UAG was cloned via site-directed mutagenesis from pGLO-GFP-1UAG using primers pGLO-GFP-3TAG-F and pGLO-GFP-3TAG-R.

To compare sY incorporation in GFP in different strains and systems, SS320 (Lucigen) or C321.ΔA.exp[6] *E. coli* cells were cotransformed with 1) a plasmid expressing *Mj* $tRNA_{CUA}$ and *Mj*TyRS (pULTRA-Y) or STyrRS (pULTRA-sY or pEVOL-sY), and 2) a plasmid expressing GFP with one or three UAG codons (pGLO-GFP-1UAG or pGLO-GFP-3UAG). Transformed cells were grown in M9T medium (1X M9 salts, 10 g/L tryptone, 5 g/L NaCl, 1 mM $MgSO_4$, and 80 □M biotin) with or without 20 mM of sY, synthesized as described previously[7]. Once cultures reached mid-log phase (OD_{600} of 0.6-0.8), the expression of GFP and aaRS were co-induced with 0.02% L-Arabinose (for pEVOL-sY) or 1 mM IPTG and 0.02% L-Arabinose (for pULTRA-Y and pULTRA-sY). The cultures were incubated for 18 hours at 30°C. 100 □L of each cell culture were transferred into 96-well black clear-bottom plates (Corning), and GFP fluorescence (excitation/emission, 395/509 nm) was measured using an Infinite 200 PRO plate reader (Tecan). Fluorescence signals were normalized by dividing by OD_{600}. Autofluorescence of cells not expressing GFP and aaRS was subtracted, and the resulting fluorescence for each sample was divided by fluorescence of *E. coli* cells expressing wild-type GFP to obtain the relative fluorescence ratio.

Cloning, Expression and Purification of RaxX60

Full length RaxX60 from *Xoo* PXO99 was expressed as a MBP-3C-RaxX-His fusion protein in *E. coli* C321.ΔA.exp[6]. For this purpose RaxX60 genomic DNA was amplified using the two primers RaxX-FL_EcoRI-3C_forward and RaxX-His-stop-HindIII-reverse and cloned into pMAL-c4X (NEB) (Table S4). A sequence verified clone was used as template for amplification using the two primers pMAL-start_E_NcoI and RaxX-His-stop-HindIII-reverse and the product inserted into pBAD_A_Myc (Invitrogen). The resulting plasmid was sequence verified and named pBAD/MBP-3C-RaxX-His. The UAG codon was introduced at the corresponding position of Y41 in RaxX60 using point mutagenesis on pBAD/MBP-3C-RaxX-His with the primer pair RaxX_amber_TAG_F and RaxX_amber_TAG_R. The resulting plasmid was sequence verified and named pBAD/MBP-3C-RaxX-His (Amber). *E. coli* C321.ΔA.exp was cotransformed with pULTRA-sY and pBAD/MBP-3C-RaxX-His (Amber) to generate sulfated RaxX60 or transformed with pBAD/MBP-3C-RaxX-His to generate unsulfated RaxX60. Transformed bacteria were grown in 0.5 L of M9T media in 2L baffled Erlenmeyer flasks. For expression of sulfated RaxX, sY was added to a final concentration of 5 mM. Cultures were grown at 37°C with shaking at 300 rpm. Expression was induced

at an $O.D._{600}$ of 0.7-0.9 by the addition of 1 mM IPTG and 0.25% (w/vol) L-Arabinose for 5 h. MBP-3C-RaxX-His was purified from intracellular total protein extracts in extraction buffer A (25mM Tris pH 8, 500 mM NaCl, 40 mM Imidazol, fresh 1mM PMSF, 0.1 µl/ml Nuclease mix (GE Healthcare), 2.5 mM $MgCl_2$, 1 mM DTT) over a 5 mL Ni-NTA column on an FPLC (Akta, GE Healthcare) and eluted with a constant linear gradient of buffer A containing 500 mM Imidazol. Fractions containing MBP-3C-RaxX-His were concentrated using Amicon Ultra-15 centrifugal filter units (10 kDa) and resuspended in buffer B (20 mM Hepes pH 8) to a final concentration of 1-5 mg/ml. The MBP-tag was removed by overnight digestion with 3C protease (Thermo Scientific) at a concentration of 1:200 (w/w) at 4°C with head over mixing. The MBP was separated from full length RaxX-His over a 1 mL SP-XL column, a strong cation exchanger, on an FPLC (Akta, GE Healthcare) with a constant gradient of buffer B containing 1 M NaCl. Fractions containing MBP-3C-RaxX-His were concentrated using Amicon Ultra-15 centrifugal filter units (3 kDa). In a third step RaxX60 was further purified by size exclusion chromatography on a Superdex 75 10/300 in buffer C (20mM Tris pH 8, 50 mM NaCl). During some experiments over 50% of the sample was lost at the final third purification step. It is recommended to avoid this last purification step if not required. Unsulfated RaxX was expressed and purified in the same way using pBAD/MBP-3C-RaxX-His in the absence of sY in the media. The yield for sulfated and unsulfated highly purified RaxX-His was up to 4 mg protein per 1L of culture.

Gene expression assay in rice

Rice defense gene expression assays were performed as described previously.[8]

Statistical analysis

Statistical analyses were performed using JMP software (ASAS Institute Inc., Cary, NC, USA).

Proteomic Analysis

RaxX peptides were digested with 1/10 (w/w) of trypsin at 37°C overnight in a buffer containing 25 mM Tris pH 8, 50 mM NaCl, and 5 mM DTT. Peptides were purified over C18 bench top spin column (Harvard Apparatus).

Shotgun proteomics

Samples were analyzed on an Agilent 1290 liquid chromatography system coupled to an Agilent 6550 iFunnel Q-TOF mass spectrometer (Agilent Technologies; Santa Clara, CA). Peptide samples (4 µg) were loaded onto an Ascentis Express ES-C18 column (10 cm length x 2.1 mm, 2.7 µm particle size) (Sigma Aldrich; St. Louis, MO) operating at 60 ºC at a flow rate of 400 µL/min. A 13.5-minute method with the following separation gradient was used: 95% Buffer A (0.1% formic acid) and 5% Buffer B (99.9% acetonitrile, 0.1% formic acid) were held for 1 minute, then Buffer B was increased to 35% over 5.5 minutes, followed by an increase to 80% Buffer B over 1 minute where it was held at 600 µL/min for 3.5 minutes. Buffer B was ramped back down to 5% over 0.5 minutes and the system was re-equilibrated for 2 minutes. Peptides were introduced into the mass spectrometer from the LC by using a Dual Jet Stream Electrospray Ionization source (Agilent Technologies) operating in positive-ion mode. Source parameters used include: gas temperature (250°C), drying gas (14 L/min), nebulizer (35 psig), sheath gas temp (250°C), sheath gas flow (11 L/min), VCap (5,000 V), fragmentor (180 V), and OCT 1 RF Vpp (750 V). The data were acquired with Agilent MassHunter Workstation Software, LC/MS Data Acquisition B.05.01 (Build 5.01.5125.2) and peak lists were extracted via the MassHunter Qualitative Analysis MGF file export option. Protein and peptide identifications from MS/MS data were accomplished with Mascot (version 2.3.02; Matrix Science), filtered, and validated using Scaffold (version 4.4.5; Proteome Software).

Targeted proteomics

Samples were analyzed using an Agilent 1290 liquid chromatography system coupled to an Agilent 6460 mass spectrometer (Agilent Technologies). Peptide samples (0.5 µg) were loaded and separated on an Ascentis Express Peptide C18 column (15 cm length x 2.1 mm ID, 2.0 µm particle size; Sigma Aldrich) operating at 60 ºC in normal flow at 400 µL/min. A 43-minute method with the following gradient was used: 98% Buffer A (0.1% formic acid) and 2% Buffer B (99.9% acetonitrile, 0.1% formic acid) were held for 12 seconds, then Buffer B was increased to 35% in 35.5 minutes, followed by an increase to 90% Buffer B in 18 seconds where it was held for 2 minutes. Buffer B was brought down to 5% over 30 seconds, then further decreased to 2% B over 2.5 minutes where it was held for 2 minutes to re-equilibrate the column to the starting conditions. Peptides were introduced into the mass spectrometer from the LC, using an Jet Stream Electrospray Ionization source (Agilent Technologies) operating in positive-ion mode. Source parameters used include: gas temperature (250°C), gas flow (13 L/min), nebulizer (35 psi), sheath gas temp (250°C), sheath gas flow (11 L/min), and VCap (3,500 V). The data were acquired with Agilent MassHunter Workstation Software, LC/MS Data Acquisition B.06.00 (Build 6.0.6025.3 SP3). Targeted

proteomic methods were developed and data were analyzed with Skyline software (MacCoss Lab, University of Washington). Data can be accessed at http://tinyurl.com/2nd-gen-sY-RaxX (E-mail: jbeireviewer@gmail.com and Password: jbeireview1).

Calculations of Relative sY status

It is well described that sulfate groups are abundantly lost from tyrosine during electrospray ionization (ESI) in the positive ion mode.[9,10] In order to adequately measure relative sulfation status, we first estimated the loss of SO_4^{2-} from the tryptic peptide covering RaxX Y41 [K.HVGGGDsYPPPGANPK.H] during ESI. Because of their distinct physiochemical properties, unsulfated [K.HVGGGDYPPPGANPK.H] and sulfated [K.HVGGGDsYPPPGANPK.H] tryptic peptides were eluted at markedly different retention times (RT) of 6.7 and 7.8 min, respectively, without any overlap. This led to the observation that RaxX60-Y derived unsulfated $[K.HVGGGDYPPPGANPK.H]_{RT=6.7}$ peptides eluted at RT of 6.7 min and RaxX60-sY derived originally sulfated peptides that lost their SO_4^{2-} during ESI $[K.HVGGGDYPPPGANPK.H]_{RT=7.8}$ eluted at RT of 7.8 min. The relative neutral loss of SO_4^{2-} from [K.HVGGGDsYPPPGANPK.H] $\%SO_4^{2-}loss$ was estimated by dividing the normalized intensity of $[K.HVGGGDYPPPGANPK.H]^{3+}{}_{RT=7.8}$ derived from RaxX60-sY by the normalized intensity (I) of $[K.HVGGGDYPPPGANPK.H]^{3+}{}_{RT=6.7}$ derived from RaxX60-Y. Equal loading (0.5 µg) and division by the intensity of an unmodified tryptic peptide $[K.GAASLQRPAGAK.G]^{2+}{}_{RT=4.7}$ of the same protein preparation achieved normalization (equation 1).

$$\%SO_4^{2-}loss = \frac{I_{RaxX60-sY}\,[\text{K.HVGGGDYPPPGANPK.H}]^{3+}_{\text{RT=7.8}}}{I_{RaxX60-sY}I[\text{K.GAASLQRPAGAK.G}]^{2+}_{\text{RT=4.7}}} \Big/ \frac{I_{RaxX60-Y}\,[\text{K.HVGGGDYPPPGANPK.H}]^{3+}_{\text{RT=6.7}}}{I_{RaxX60-Y}\,[\text{K.GAASLQRPAGAK.G}]^{2+}_{\text{RT=4.7}}} \quad \text{[Equation 1]}$$

Based on two different protein preparations and three technical replicate runs each we estimated an approximate loss of SO_4^{2-} from [K.HVGGGDsYPPPGANPK.H] of 11% ($\%SO_4^{2-}loss = 11$). This enabled us to calculate the relative sulfation status ($\%sY$) of RaxX-sY as shown in equation 2 and Table S2.

$$\%sY = (100 - \%SO_4^{2-}loss) + \%SO_4^{2-}loss * \frac{I\,[\text{K.HVGGGDYPPPGANPK.H}]^{3+}_{\text{RT=7.8}}}{I\,[\text{K.HVGGGDYPPPGANPK.H}]^{3+}_{\text{RT=6.7}} + I\,[\text{K.HVGGGDYPPPGANPK.H}]^{3+}_{\text{RT=7.8}}} \quad \text{[Equation 2]}$$

Circular Dichroism

The melting curves of RaxX60-Y and RaxX60-sY covered a temperature gradient from 20 °C to 85 °C, and were recorded with a JASCO J-815 CD spectrometer. The experimental concentrations were 0.1 mg/mL protein, 5 mM NaCl and 2 mM Tris (pH 8). The results represent averages of two accumulations each. All spectra were taken at a wavelength range from 195 nm to 250 nm, applying 0.2 nm data pitches to enhance the overall experimental detail. The spectra shown in Figure S3 represent the cross sections of both melting curve plots at 20 °C. In order to predict the secondary structure contents of RaxX60-Y and RaxX60-sY on the basis of the CD spectra, the 20 °C data were processed using the programs CRDATA, CDSSTR, SELCON3 and CONTIN/LL of CDPro, while applying different sets of parameters (Table S3; http://lamar.colostate.edu/~sreeram/CDPro/).[11,12] In particular, we chose the protein reference datasets 1, 7 and 10 as a basis for a comparative analysis, because they cover the largest wavelength range (1), the largest dataset of soluble proteins (7), or the overall largest dataset (10), respectively.

Post-treatment experiment

Kitaake, and a Kitaake transgenic line expressing Myc-XA21 under the maize ubiquitin promoter (UX) were used for the post-treatment assay.[13] Rice seeds were surface-sterilized using 15% bleach, rinsed with water and germinated in distilled water at 28°C for 1 week. Well-grown seedlings were transplanted into trays filled with A-OK Starter Plugs (Grodan) and watered twice a week with Hoagland's solution. For the post-treatment assay, 4-week-old plants were inoculated using the scissors clipping method, and treated with 1 µM RaxX60-Y, RaxX60-sY or mock solution (with 0.02% Tween20) at 2 days post-infection (dpi). The lesions were measured at 10 to 14 dpi depending on lesion progression.

Commercial peptides

Sulfated (sY) and unsulfated (Y) versions of RaxX39 (KGRPEPLDQRLWKHVGGGDYPPPGANPKHDPPPRNPGHH), RaxX24 (LWKHVGGGDYPPPGANPKHDPPPR), RaxX21 (HVGGGDYPPPGANPKHDPPPR), RaxX18 (LWKHVGGGDYPPPGANPK) were ordered from Pacific Immunology. In one case the supplier was unable to provide high quality peptides for RaxX39-sY. Mass-spectrometry analysis by the supplier indicated a K^+ adduct, which was indirectly supported by in house SRM analysis. In addition, we ordered RaxX21-sY peptides also from one additional commercial vendor. However this vendor was not able to provide us with peptides that met commercial quality standards. The peptides were assessed for purity by HPLC and MS analysis by the vendor. All peptides were resuspended in ddH_2O. No other solvents were required to dissolve the peptide.

Sequence of GFP-1UAG

MA*SKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGY
VQERTISFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYITADKQKNGIKANFKIRHNIEDGSVQLADHYQQNTPI
GDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK

Sequence of GFP-3UAG

MA*SKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGY
VQERTISFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSH*V*ITADKQKNGIKANFKIRHNIEDGSVQLADHYQQNTPI
GDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK

"*" indicates an amino acid specified by the UAG amber codon

Supplementary Figure legends:

Supplementary Figure 1. Vector Maps of pULTRA-sY, pEVOL-sY, and pGLO-GFP-3UAG

Plasmid maps of (A) pULTRA-sY, encoding an inducible STyrRS (the sY-specific *Mj* aaRS) controlled by the *tacI* promoter, an optimized *Mj*-$tRNA_{CUA}$, an antibiotic resistance marker for spectinomycin (SmR), and origin (CloDF) for compatibility with common plasmids, (B) pEVOL-sY, an older generation plasmid for sY incorporation, and (C) pGLO-GFP-3UAG, the plasmid for expressing GFP with 3 UAG codons at previously reported permissive sites.[1]

Supplementary Figure 2. High purity isolation of recombinantly produced RaxX60-Y and RaxX60-sY

8-20% SDS-PAGE gradient gel of different amounts of three-step purified RaxX60-Y and RaxX60-sY from *E.coli* stained with SimplyBlue safe stain. Both proteins migrate at around 12kDa slightly above their predicted molecular weight of 7.2kDa.

Supplementary Figure 3. RaxX60-Y and RaxX60-sY are largely disordered and potentially form a single β-sheet

Circular dichroism (CD) spectra of 0.1mg/ml RaxX60-Y and RaxX60-sY at 20^{o}C.

Identified Proteins	Accession Number	Molecular Weight	Normalized total spectral count	
			RaxX60-Y	**RaxX60-sY**
RaxX60-His_6	RS05990_PXO	7 kDa	76 (~97%)	71 (~92%)
Cyclic di-GMP phosphodiesterase YfgF	YFGF_ECOLI	86 kDa	2	0
50S ribosomal protein L2	RL2_ECOLI	30 kDa	0	6

Supplementary Table 1. High purity isolation of RaxX60-Y and RaxX60-sY

Normalized total spectral count of tryptic peptides identified by HPLC-ESI-MS/MS analysis of three-step purified recombinantly produced RaxX60-Y and RaxX60-sY. Proteins are identified by at least one peptide with a Mascot Ion score ≥ 25.0. Values in brackets indicate approximate relative purity of RaxX60-Y and RaxX60-sY, respectively.

RaxX variant	Purity	Relative sulfation status [%sY]
RaxX21-Y[a]	>80%[c]	N.A.
RaxX24-Y[a]	>80%[c]	N.A.
RaxX39-Y[a]	>95%[c]	N.A.
RaxX60-Y[b]	>90%[d]	N.A.
RaxX21-sY[a]	>80%[c]	95%
RaxX24-sY[a]	>80%[c]	97%
RaxX39-sY[a]	>95%[c]	97%
RaxX60-sY[b]	>90%[d]	>99.5%

Supplementary Table 2. High quality production of sulfated RaxX60-sY in the second-generation sY *E.coli* expression system

Summary of the approximated relative sulfation status as measured by SRM-MS analysis of different RaxX variants either produced by standard commercial chemical synthesis ([a]) or recombinant expression in *E. coli* ([b]). Purity is given as provided by the commercial supplier ([c]) or measured by MS/MS analysis ([d]) in Supplementary table 1. "N.A." indicates not applicable.

		RaxX60-Y				RaxX60-sY			
		α-helix [%]	β-strand [%]	turns [%]	unordered [%]	α-helix [%]	β-strand [%]	turns [%]	unordered [%]
CDPro	overall	1.4 – 7.0	18 – 34	11 – 25	39 – 69	3.7 – 10.7	10 – 36	11 – 29	31 – 73
	1	1.4 – 5.8	24 – 34	23 – 25	41 – 48	3.7 – 8.7	23 – 36	24 – 29	31 – 43
	7	2.6 – 4.3	18 – 21	11 – 14	61 – 69	5.4 – 9.0	10 – 18	11 – 14	65 – 73
	10	3.0 – 7.0	29 – 34	23 – 25	39 – 42	5.5 – 10.7	25 – 31	24 – 26	37 – 40
	CDSSTR	1.4 – 3.3	19 – 34	13 – 25	40 – 64	3.7 – 6.2	17 – 33	11 – 26	37 – 65
	CONTIN/LL	4.3 – 7.0	21 – 31	13 – 23	40 – 62	5.1 – 6.7	16 – 36	14 – 29	31 – 65
	SELCON3	3.7 – 5.8	18 – 34	11 – 24	39 – 69	8.7 – 10.7	10 – 25	11 – 26	40 – 73

Supplementary Table 3. RaxX60-Y and RaxX60-sY are largely disordered and potentially form a single β-sheet

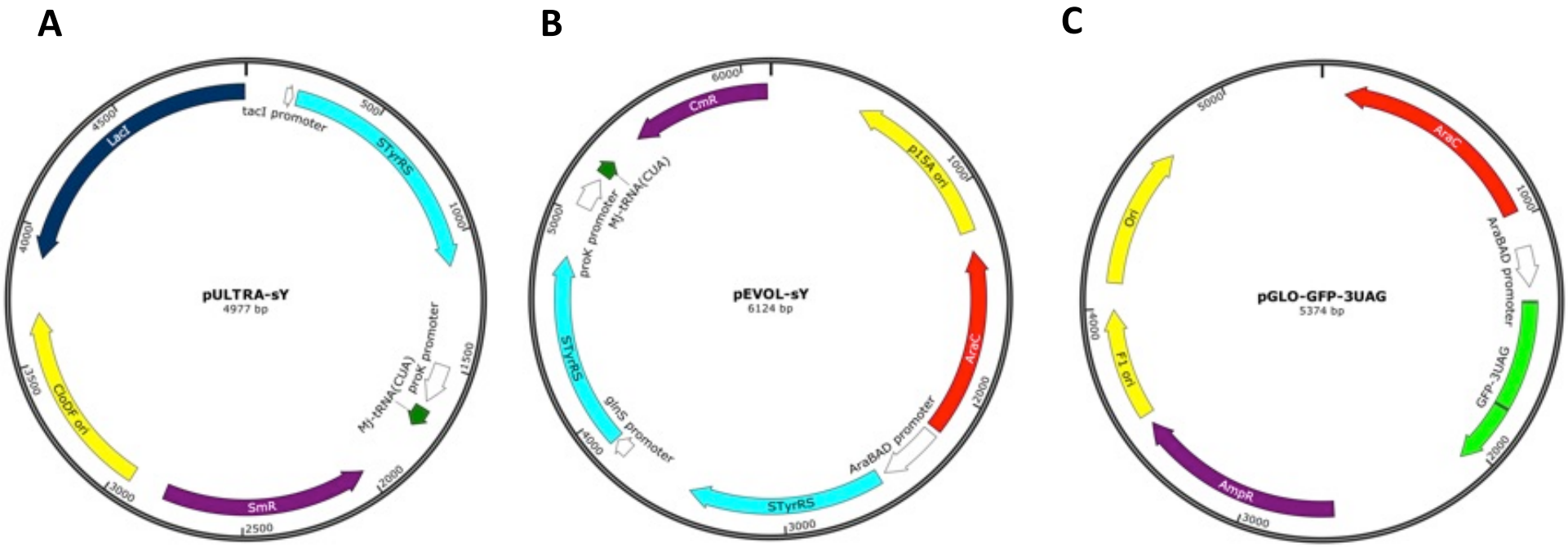
Summary of RaxX60-Y and RaxX60-sY relative secondary structure contents as determined by circular dichroism (CD) spectroscopy. The top row summarizes the overall output ranges generated from the data. The following rows split up the data into output subsets specifying the relationship between the outputs and the CDPro reference protein datasets (1, 7 or 10) or programs (CDSSTR, CONTIN/LL or SELCON3) used to generate them, respectively.

Name	Sequence
aaRStopULTRAPCR1	catcgcggccgcATGGACGAATTTGAAATGATAAAGAGA
aaRStopULTRAPCR2	catcgcggccgcTTATAATCTCTTTCTAATTGGCTCTAAAATC
pEVOL-BB-F	GTCGACCATCATCATCATCA
pEVOL-BB-R	AGATCTAATTCCTCCTGTTA
pEVOL-BB2-F	gagattataaCTGCAGTTTCAAACGCTAAA
pEVOL-BB2-R	attcgtccatATGGGATTCCTCAAAGCGTA
aaRS-F	taacaggaggaattagatctATGGACGAATTTGAAATGAT
aaRS-R	tgatgatgatgatggtcgacTTATAATCTCTTTCTAATTGGC
aaRS2-F	ggaatcccatATGGACGAATTTGAAATGAT
aaRS2-R	gaaactgcagTTATAATCTCTTTCTAATTGGCTC
pGLO-GFP-1TAG-F	catatggcttagAGCAAAGGAGAAGAACTTTT
pGLO-GFP-1TAG-R	tcctttgctctaAGCCATATGTATATCTCCTT
pGLO-GFP-3TAG-F	tcacactaggtatagATCACGGCAGACAAACAAAA
pGLO-GFP-3TAG-R	gtgatctatacctaGTGTGAGTTATAGTTGTACTC
RaxX-FL_EcoRI-3C_forward	AAATGGAATTCGGTCTGGAAGTTCTGTTTCAGGGCCCGAACCACTCGAAAAAATCG
RaxX-His-stop-HindIII-reverse	AAATGTAAGCTTTCAGTGATGGTGGTGATGGTGATGGTGCCCGGGGTTGCG
pMAL-start_E_NcoI	AAATGTCCATGGAAATCGAAGAAGGTAAACTG
RaxX_amber_TAG_F	GGCGGTGGGGACTAGCCCCCGCCGGGCGCGAAT
RaxX_amber_TAG_R	ATTCGCGCCCGGCGGGGGCTAGTCCCCACCGCC

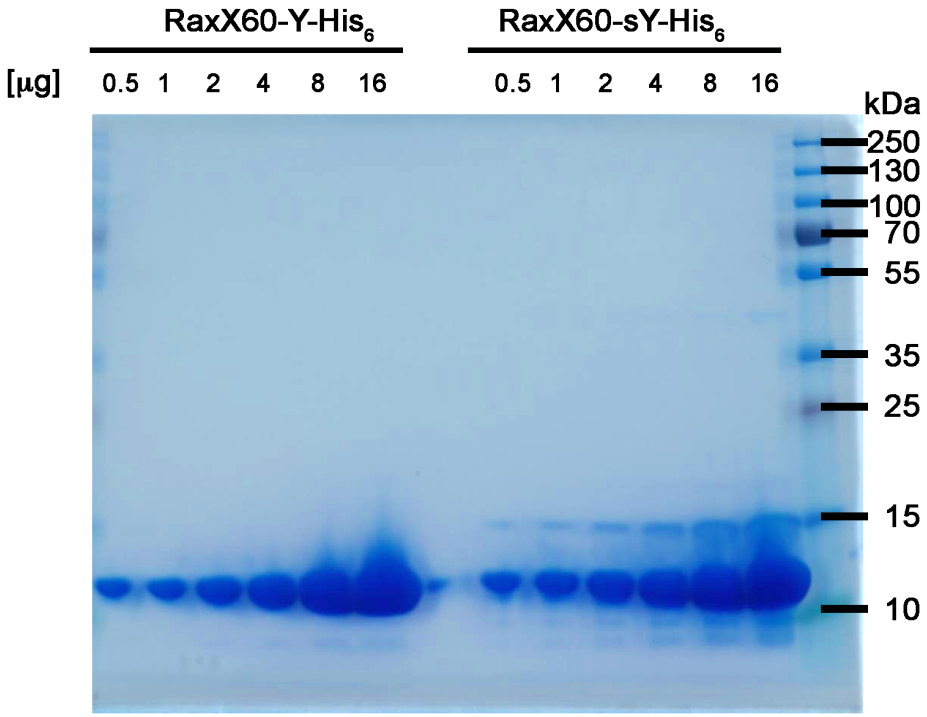
Supplementary Table 4. Primers used in this study

References

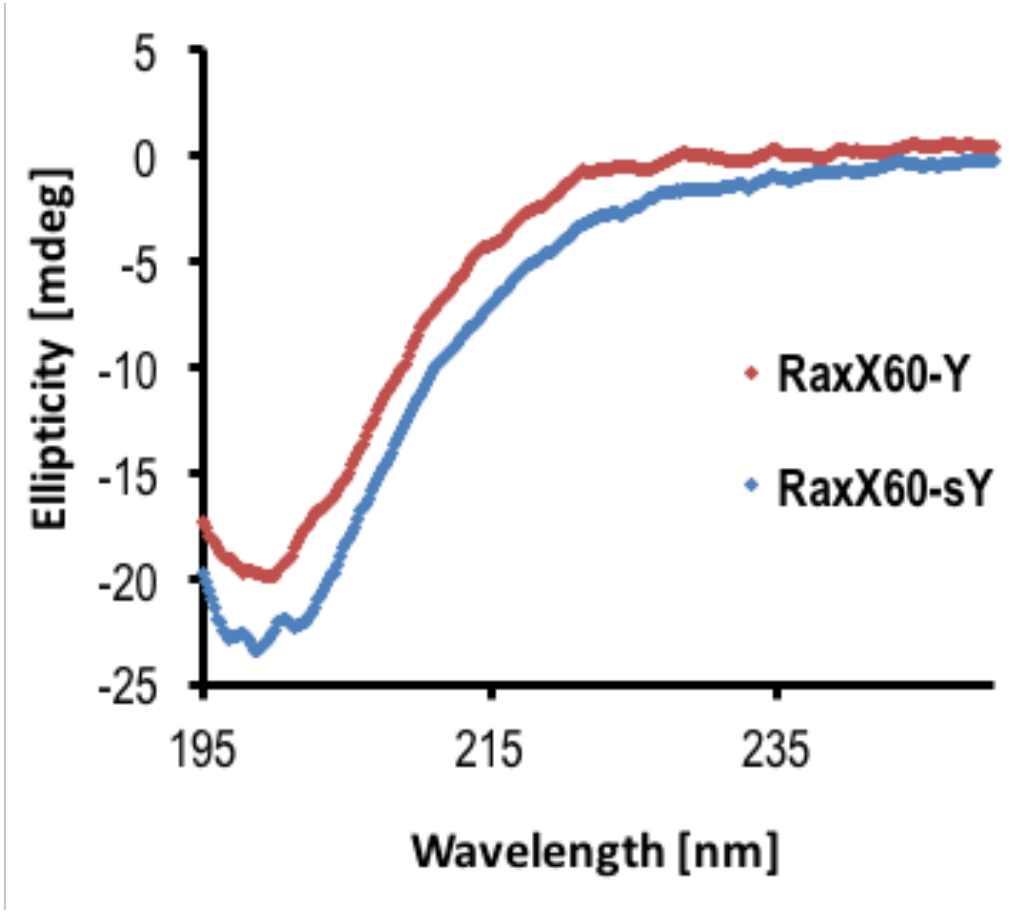
1 A. Chatterjee, S. B. Sun, J. L. Furman, H. Xiao and P. G. Schultz, *Biochemistry*, 2013, **52**, 1828–1837.
2 T. S. Young, I. Ahmad, J. A. Yin and P. G. Schultz, *Journal of Molecular Biology*, 2010, **395**, 361–374.
3 C. C. Liu and P. G. Schultz, *Annual Review of Biochemistry*, 2010, **79**, 413–444.
4 C. C. Liu and P. G. Schultz, *Nat Biotech*, 2006, **24**, 1436–1440.
5 D. G. Gibson, L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison and H. O. Smith, *Nat Meth*, 2009, **6**, 343–345.
6 M. J. Lajoie, A. J. Rovner, D. B. Goodman, H.-R. Aerni, A. D. Haimovich, G. Kuznetsov, J. A. Mercer, H. H. Wang, P. A. Carr, J. A. Mosberg, N. Rohland, P. G. Schultz, J. M. Jacobson, J. Rinehart, G. M. Church and F. J. Isaacs, *Science*, 2013, **342**, 357–360.
7 C. C. Liu, S. E. Cellitti, B. H. Geierstanger and P. G. Schultz, *Nat. Protocols*, 2009, **4**, 1784–1789.
8 R. N. Pruitt, B. Schwessinger, A. Joe, N. Thomas, F. Liu, M. Albert, M. R. Robinson, L. J. G. Chan, D. D. Luu, H. Chen, O. Bahar, A. Daudi, D. D. Vleesschauwer, D. Caddell, W. Zhang, X. Zhao, X. Li, J. L. Heazlewood, D. Ruan, D. Majumder, M. Chern, H. Kalbacher, S. Midha, P. B. Patil, R. V. Sonti, C. J. Petzold, C. C. Liu, J. S. Brodbelt, G. Felix and P. C. Ronald, *Science Advances*, 2015, **1**, e1500245.
9 F. Monigatti, B. Hekking and H. Steen, *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 2006, **1764**, 1904–1913.
10 C. Seibert and T. P. Sakmar, *Biopolymers*, 2008, **90**, 459–477.
11 N. Sreerama and R. W. Woody, *Analytical Biochemistry*, 2000, **287**, 252–260.
12 N. Sreerama and R. W. Woody, *Protein Science*, 2004, **13**, 100–112.
13 C.-J. Park, S.-W. Lee, M. Chern, R. Sharma, P. E. Canlas, M.-Y. Song, J.-S. Jeon and P. C. Ronald, *Plant Science*, 2010, **179**, 466–471.



Supplementary Figure 1. Vector Maps of pULTRA-sY, pEVOL-sY, and pGLO-GFP-3UAG
Plasmid maps of (A) pULTRA-sY, encoding an inducible STyrRS (the sY-specific *Mj* aaRS) controlled by the *tacI* promoter, an optimized *Mj*-$tRNA_{CUA}$, an antibiotic resistance marker for spectinomycin (SmR), and origin (CloDF) for compatibility with common plasmids, (B) pEVOL-sY, an older generation plasmid for sY incorporation, and (C) pGLO-GFP-3UAG, the plasmid for expressing GFP with 3 UAG codons at previously reported permissive sites.[1]



Supplementary Figure 2. High purity isolation of recombinantly produced RaxX60-Y and RaxX60-sY 8-20% SDS-PAGE gradient gel of different amounts of three-step purified RaxX60-Y and RaxX60-sY from *E.coli* stained with SimplyBlue safe stain. Both proteins migrate at around 12kDa slightly above their predicted molecular weight of 7.2kDa.



Supplementary Figure 3. Circular dichroism (CD) spectra of 0.1mg/ml RaxX60-Y and RaxX60-sY at 20°C. RaxX60-Y and RaxX60-sY are largely disordered and potentially form a single β-sheet.