Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Electronic Supplementary Information

Extended timescale 2D IR probes of proteins: *p*-cyanoselenophenylalanine

S. Ramos, K. J. Scott, R. E. Horness, A. L. Le Sueur, M. C. Thielges

Department of Chemistry, Indiana University, 800 E. Kirkwood, Bloomington, Indiana, 47405, United States

I. Experimental Methods

Synthesis of p-cyano-seleno-phenylalanine (CNSePhe). Synthesis of *p*-cyano-seleno-phenylalanine (CNSePhe) was performed as described by Ganther.[1] In an ice bath, 10 mmol of *p*-amino-L-phenylalanine (H_2NPhe) was dissolved in 4 mL of 6 M HCl. A chilled, fresh solution of 6 M sodium nitrite was added dropwise until the solution tested positive for excess nitrile on potassium iodide starch paper. The solution was then adjusted to pH 4 with 5 M sodium acetate. Finally, a 2.5 M solution of potassium selenocyanate (KSeCN) was added to the solution dropwise to a slight excess of KSeCN. Upon the cessation of N_2 gas evolution, the reaction mixture was removed from the ice bath and allowed to stir at room temperature, protected from light, for ~19 hours. Care was taken with all subsequent use of the CNSePhe to reduce its exposure to light and oxygen, which promoted degradation to an insoluble red solid. The CNSePhe was purified by HPLC using a water/acetonitrile gradient. Calculated mass 270.2; experimental 270.9.

UV/visible spectroscopy of CNSePhe. Solutions of CNSePhe at 100 mM in pH <1 H_2O or at 20-40 mM in 65:35 methanol:water were characterized by UV-vis spectroscopy (Agilent Cary 300 spectrometer). The UV/vis spectrum displays a characteristic band at 236 nm and a small shoulder stretching from 250-300 nm (Fig. S2), as previously reported.[1]

FT IR Spectroscopy. The samples were loaded between two 1 mm thick calcium fluoride (CaF_2) windows separated by a 38.1 μm Teflon spacer. Spectra were collected with an Agilent Cary 670 FTIR spectrometer equipped with a liquid N_2-cooled MCT detector. All spectra were collected after a 15-minute purge with dry N_2 gas. Sample and background spectra were averaged over 10,000 scans at a resolution of 4 cm^{-1}, and all interferograms were processed using the 4-term Blackman-Harris apodization function with a zero-filling factor of 8. Background spectra were collected of pure solvent. Slowly varying baselines were removed by fitting and then subtracting a polynomial from the spectral regions excluding the CN band (Matlab R2016b). To determine the IR extinction coefficients in 65:35 methanol:water, the IR absorption bands were measured for solutions at several probe concentrations, and a 127 μm spacer was used. For the CN of CNSePhe, the solution concentrations were determined using an extinction coefficient of 7.18 $mM^{-1}cm^{-1}$ at 236 nm in 65:35 methanol:water.[1]

Pump-probe Spectroscopy. A Ti:Sapphire oscillator/regenerative amplifier (Spectra Physics) generating ~75 fs pulses at 800 nm and 1 kHz is used to pump a home-built optical parametric amplifier (OPA), which allows for the generation of tunable, ~30 μJ, ~170 fs mid-IR pulses. Mid-IR pulse duration was determined by cross correlation using the strong non-resonant signal generated by CCl_4. Temporal chirp is measured by frequency-resolved transient grating and corrected to ≤ 0.02 fs/cm^{-1} by placement of Ge and/or CaF_2 flats in the beam. The mid-IR beam was split into a pump (8.4 μJ) and probe (0.7 μJ) beam, which were focused to ~300 um spot sizes (99% diameter) and spatially overlapped at the sample. A half-waveplate followed by a holographic wire-grid polarizer was placed in the probe beam to orient its polarization by magic angle (54.7°) with respect to the pump beam in order to eliminate contributions to the signal decay from rotational reorientation. A second wire-grid polarizer set at 54.7° was placed after the sample to selectively detect only signal components polarized at the magic angle.[2] The delay

between the pump and probe pulses was varied using computer-controlled delay stages (Aerotech) and time zero was determined using the strong pump-probe signal observed from an indium arsenide (InAs) wafer. The pump beam was chopped at 500 Hz. The probe beam was frequency-dispersed onto a 32-pixel liquid nitrogen-cooled MCT detector array (Infrared Associates). The pump-induced change in the transmission of the probe beam was measured as a function of pump-probe delay. A large, slowly decaying contribution from water heating was present in the data for the aqueous CNSePhe solution. To correct for this background signal, the difference spectrum for the longest pump-probe delay (200 ps), when the contribution from the CNSePhe was assumed be negligible, was subtracted from the spectra at every other pump-probe delay. The pump-probe spectra at each frequency were then fit to a biexponential decay, with one component accounting for the ultrafast nonresonant response near time zero and a second for the time-dependent change in the CN probe transmission (MATLAB 2016).

2D IR Vibrational Echo Spectroscopy. 2D IR vibrational echo spectra were collected as previously reported using the mid-IR laser system described above for pump-probe spectroscopy.[3] The mid-IR output (set to 18 μJ total energy using a crossed-polarizer attenuator) from the OPA was split along four separate paths: beams 1, 2, 3, and a local oscillator (LO). Three of these paths (beams 1, 2, and LO) are temporally controlled using computer-controlled delay stages (Aerotech). The timing between the pulses of beams 1, 2, and 3, as well as the temporal width of the pulses, was determined by cross-correlation measured with the non-resonant signal produced in carbon tetrachloride (CCl_4). A reference beam was split from the LO and detected on a single-element liquid nitrogen-cooled MCT detector to account for any shot-to-shot laser fluctuations. In a 2D IR vibrational echo experiment, the application of three pulses (beams 1, 2 and 3) onto the sample lead to emission of a third order signal in the phase-matched ($-k_1 + k_2 + k_3$) direction. The third-order signal is heterodyne-detected by overlapping with the LO to provide amplification and phase information. The combined signal and LO are dispersed by a spectrograph (Horiba iHR320, 300 grooves/mm grating) onto a 32-element MCT detector. A single 2D spectrum is generated by scanning the time between the first two pulses (τ) with the time period between the second and third pulses (T_w) held constant. The heterodyned third-order signal generated at time $\leq \tau$ after the application of the third pulse produces an interferogram along τ, which is Fourier transformed to produce the ω_1 axis. Frequency-resolved detection with the MCT array generates the ω_3 axis of the 2D spectrum. All 2D IR spectra were phase-corrected as described by Park et al.[3] To correct for a large background contribution to the 2D spectra from water, a 2D spectrum acquired at T_w of 140 ps, when the signals from the CN are largely decayed, was subtracted from the 2D spectra at earlier T_w times.

Expression and Purification of CNSePhe11 SH3^{Sho1}. Attempt was made at incorporation of CNSeF via amber codon suppression into SH3^{Sho1}, a Src-homology 3 domain from yeast protein Sho1 involved in the osmotic stress response. The expression plasmid encoded SH3^{Sho1} with a C-terminal thrombin cleavage recognition sequence, followed by a hexahistidine tag (His6), and the codon TAG introduced in the gene sequence at residue 11, as has been described previously.[4] The pUltraCNF plasmid which encodes the orthogonal tRNA synthetase and tRNA for amber codon incorporation of *CN*Phe was generously provided by Peter Schultz (Scripps Research Institute).[5] The expression vector for SH3^{Sho1} with the amber codon mutation was co-

transformed with pUltraCNF into chemically competent BL21(DE3) *E. coli*. A single colony was used to inoculate 4 mL of LB supplemented with 100 μg/mL ampicillin and 60 μg/mL streptomycin. After approximately 12 hours shaking at 37 °C, 10 μL of culture were used to inoculate 25 mL of LB, which was allowed to shake at 37 °C for 12 to 14 hours. One mL of this culture was used to inoculate 250 mL of TB. The culture was allowed to grow to an optical density at 600 nm (OD_{600}) of 0.1 at 37 °C, at which time the culture was protected from light, and filter-sterilized DL-CNSePhe in aqueous solution was added to a final concentration of approximately 2 mM. At this time, cell growth slowed. Once the cells began to recover and the OD_{600} reached 0.5, protein production was induced for 24 hours by addition of 0.2 mM isopropyl β-D-1-thiogalactopyranoside (IPTG).

After expression, isolation of cells and all further purification steps were performed in the dark. Cells were isolated by centrifugation, resuspended in 50 mM sodium phosphate, pH 8.0 with 300 mM NaCl and 10 mM imidazole, and lysed by treatment with 1 mg/mL lysozyme and sonication (Qsonica Q500, 6 bursts of 10 seconds at 60% amplitude with 30 seconds of rest on ice between each). The cell lysate was treated with benzonase (Santa Cruz Biotechnology) and clarified by centrifugation. The clarified supernatant was combined with Ni^{2+}-nitriloacetate (NiNTA) affinity media (Gold Bio) and the slurry was rocked gently on ice for 1 hour. The media was then washed with approximately 7 volumes of 50 mM sodium phosphate, pH 8.0 with 300 mM NaCl and 20 mM imidazole, followed by 8 volumes of the same buffer with 250 mM imidazole to elute bound protein. The eluent was concentrated to a final volume of 1 mL in a 3 kDa MWCO centrifugal filter (Millipore) and subsequently dialyzed into phosphate-buffered saline (PBS: 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4) before treatment with thrombin (Novagen) at 1 unit/mg of protein. The cleavage reaction was allowed to proceed for 16 hours at room temperature before inhibition with 1 mM phenylmethylsulfonyl fluoride (PMSF). Any remaining undigested protein was removed through passage over NiNTA media. Finally, the protein was purified by size exclusion chromatography using a S100HR Sephacryl column (GE Life Sciences) in PBS.

A trypsin digest of the purified SH3 domain was performed at a protein concentration of 0.1 mg/mL of solution with 1:30 mass ratio of SH3:trypsin (Thermo Scientific) and 25% acentonitrile. After overnight digest at 37 °C, the sample was desalted with a C18 ZipTip (Millipore), mixed with α-cyano-4-hydroxycinnamic acid (CCA) matrix, and analyzed by MALDI-TOF-MS on a Bruker Autoflex III instrument (Billerica, MA).

II. Supplemental Figures

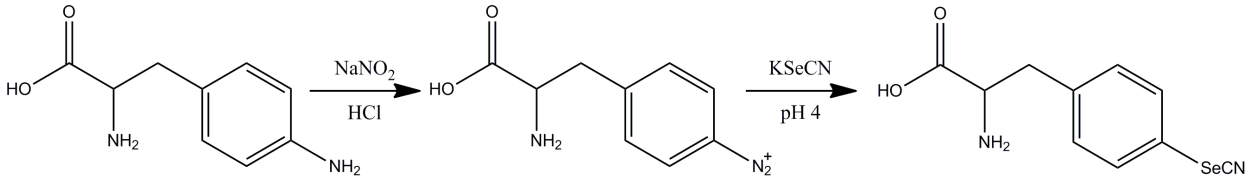

Figure S1: Synthetic scheme for CNSePhe.

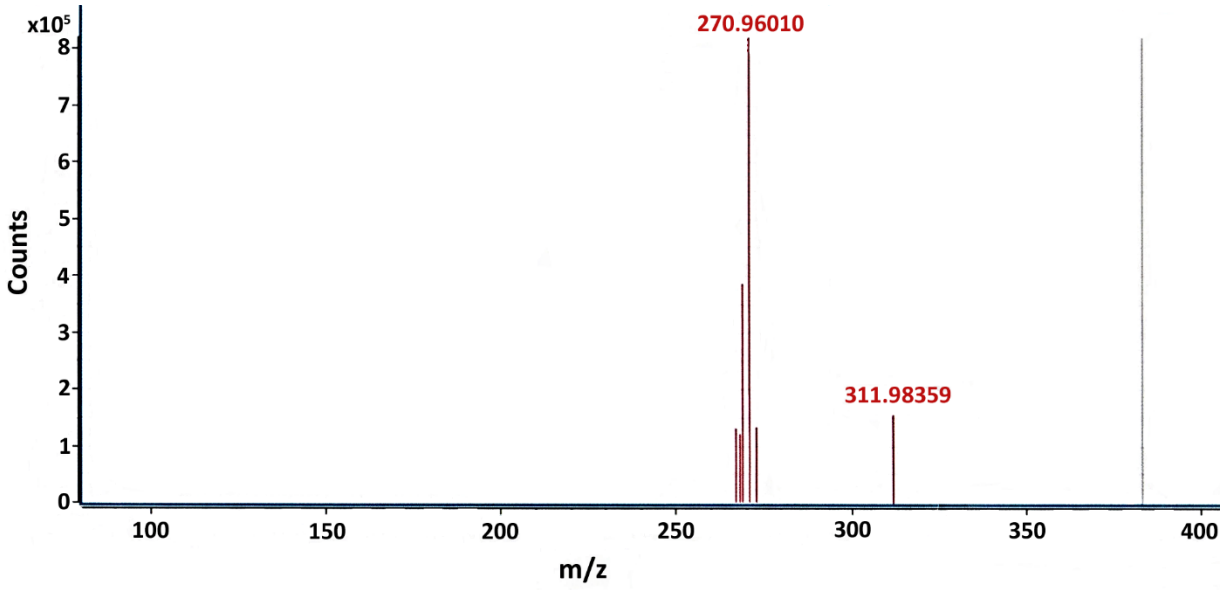

Figure S2: Mass spectrum of CNSePhe (m/z experimental: 270.9, calculated 270.2).

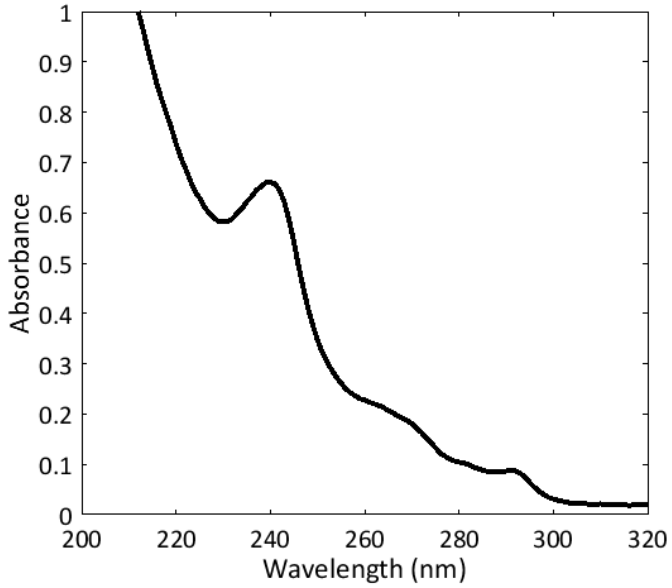

Figure S3: UV/visible absorption spectrum of 0.10 mM CNSePhe in 30% HCl.

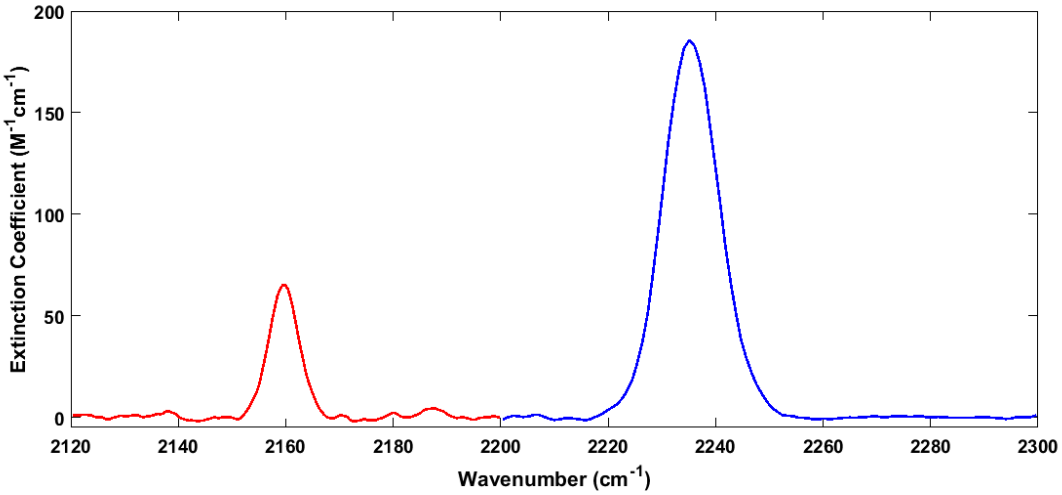

Figure S4: FT IR extinction spectra of CNSePhe and cyanophenylalanine (CNPhe).

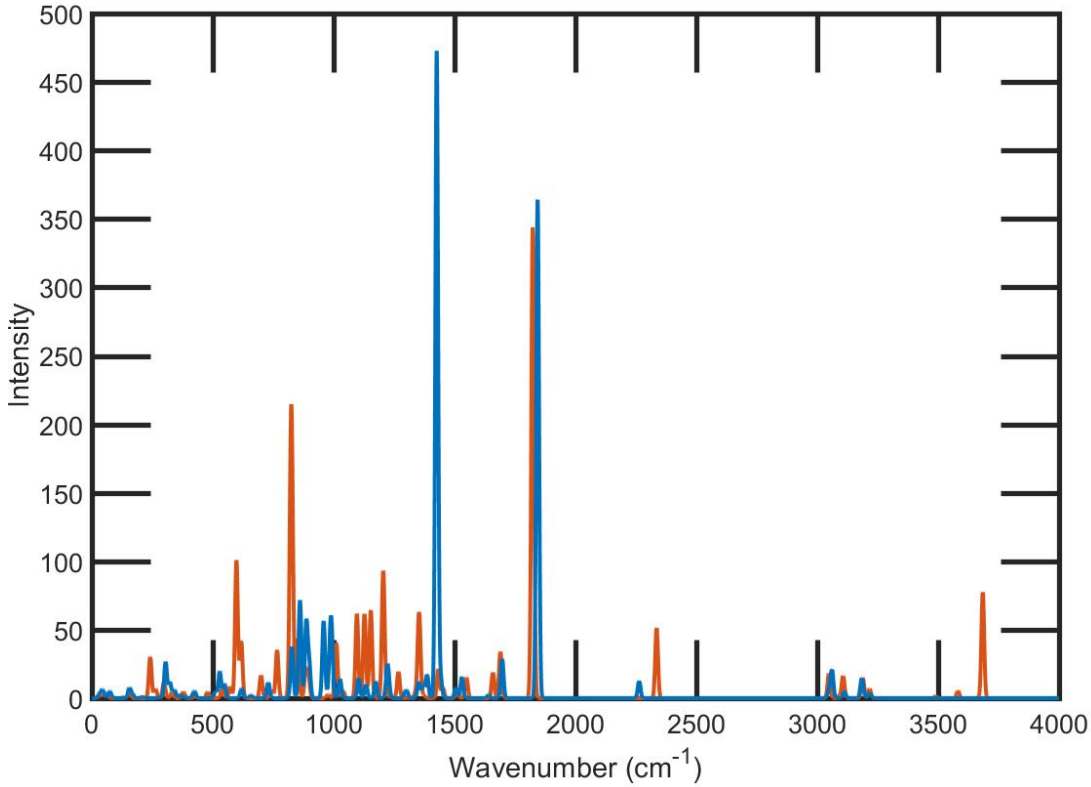

Figure S5: Calculated spectra of CNPhe (orange) and CNSePhe (blue) from density functional theory calculations performed using the Spartan ’14 software package with the B3LYP functional and the 6-31+G* basis set.

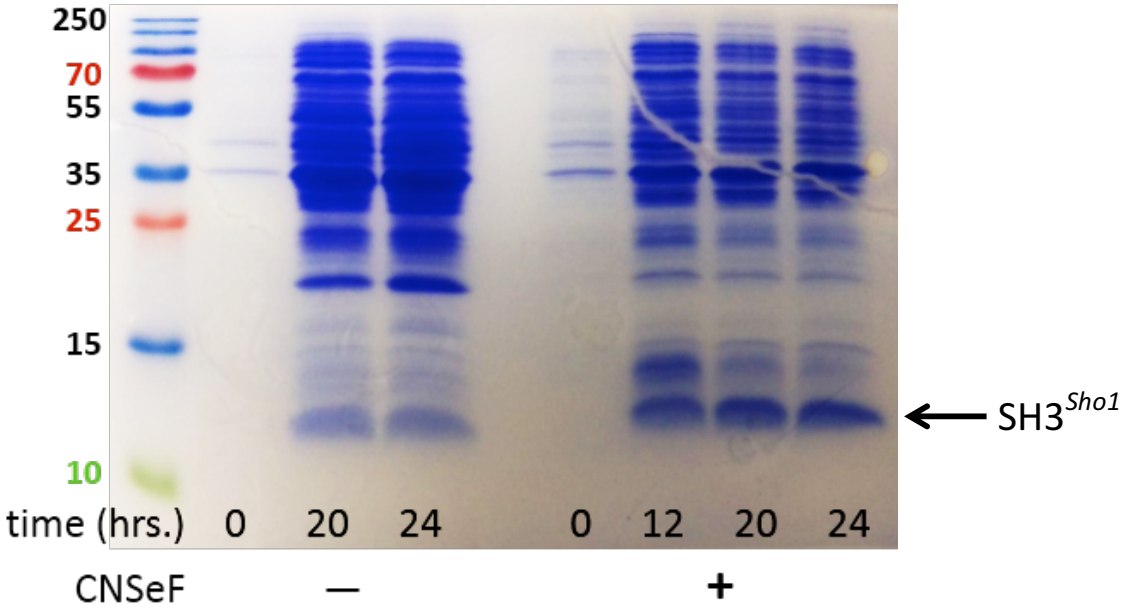

Figure S6: SDS-PAGE analysis of CNSePhe11-labeled SH3^{Sho1} expression.

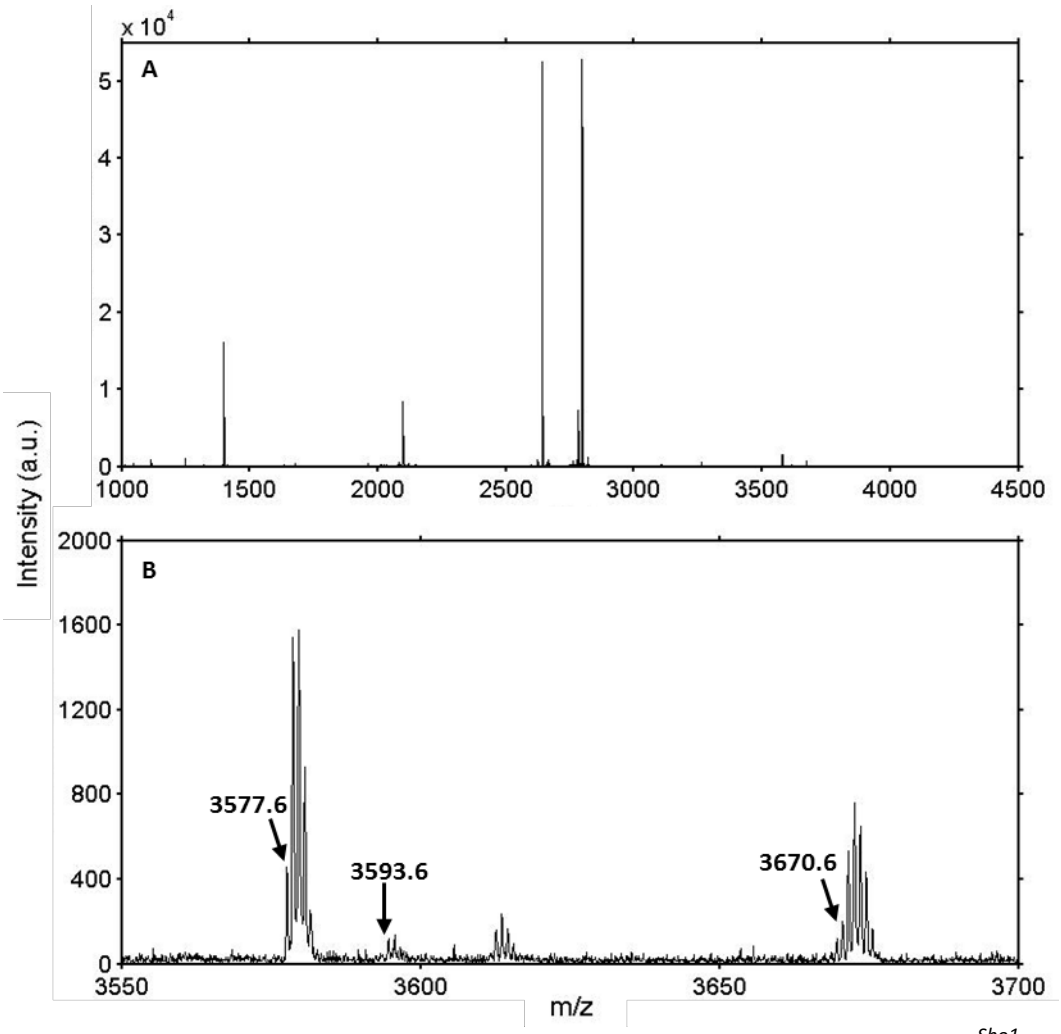

Figure S7: Mass spectral data from expression of CNSePhe11-labeled SH3^{Sho1} via amber codon suppression. A) The complete mass spectrum showing prominent peaks at m/z 2796.4 and 2640.3 corresponding to tryptic digest fragments containing residues 50-74 and 51-74, respectively. B) A mass spectrum focusing on the tryptic digest fragment containing residues 14-44, where CNSePhe11 corresponds to residue 16. The very small peak at m/z 3593.6 indicates the presence of wild type SH3^{Sho1} containing Tyr11, while the peak at m/z 3577.6 indicates incorporation of Phe rather than CNSePhe. The slightly broader envelope of peaks at m/z 3670.6 is consistent with incorporation of the heavy element Se, but is about 12 Da lower than expected for the tryptic fragment of CNSePhe11 SH3^{Sho1} (m/z 3682).

III. References

1. H. E. Ganther, *Bioorg. Med. Chem.*, 2001, **9**, 1459-1466.
2. H. S. Tan, I. R. Piletic and M. D. Fayer, *J. Opt. Soc. Am. B*, 2005, **22**, 2009-2017.
3. S. Park, K. Kwak and M. D. Fayer, *Laser Phys. Lett.*, 2007, **4**, 704-718.
4. R. E. Horness, E. J. Basom and M. C. Thielges, *Anal. Methods*, 2015, **7**, 7234-7241.
5. A. Chatterjee, S. B. Sun, J. L. Furman, H. Xiao and P. G. Schultz, *Biochemistry*, 2013, **52**, 1828-1837.