Supporting Information

A Strongly Absorbing Class of Non-Natural Labels for Probing Protein Electrostatics and Solvation with FTIR and 2D IR Spectroscopies

Ann Marie Woys, Sudipta S. Mukherjee, David R. Skoff, Sean D. Moran, Martin T. Zanni

General Synthetic Procedures. $CpRe(CO)_3$ (**2**) was prepared as published[1] and subsequently reacted to give $[\eta^5\text{-}(C_5H_4)CH_2CH_2CH_2CH_2I]Re(CO)_3$, $[\eta_5\text{-}(C_5H_4)CH_2CH_2CH_2I]Re(CO)_3$, and $[\eta^5\text{-}(C^5H^4)CH^2OH]Re(CO)_3$ as described by literature methods.[2-4] Glassware and needles were dried in an oven prior to use. All reactions were performed under nitrogen. CH_2Cl_2 and ACN were purified over alumina.[5] DMF vacuum distilled from CaH_2. Acetone was used as received. All other reagents were used as received. TLC was performed with Silica Gel 60 F254 plates. 40-63 µm SiliaFlash® P60 was used for flash column chromatography. Silica gel columns were prepared with Silicycle 40-63 Å silica (230-400 mesh) according to the method described by Still, except that the packing was by the slurry method.[6]

The ^1H and ^{13}C NMR spectra for $[\eta^5\text{-}(C_5H_4)CH_2I]Re(CO)_3$ were collected on a Bruker AC+ 300 spectrometer in the UW-Madison Chemistry Department. All other ^1H and ^{13}C spectra were acquired with an Avance-500 spectrometer at NMRFAM. The HSQC spectra were used for determining peak assignments and collected on Varian Inova-600 and Inova-500 spectrometers in the UW-Madison Chemistry Department. All spectra were referenced with TMS at 0 ppm. EI mass spectrometry was performed with a Waters (Micromass) AutoSpec® in the UW-Madison Chemistry Department. MALDI-TOF spectra were collected on an Applied Biosystems/MDS SCIEX 4800 MALDI TOF/TOF. FTIR spectra were acquired with a Thermo Scientific Nicolet iS10 FT-IR Spectrometer.

$[\eta^5\text{-}(C_5H_4)CH_2CH_2CH_2CH_2SSO_2CH_3]Re(CO)_3$ (1a). To $[\eta^5\text{-}(C_5H_4)CH_2CH_2CH_2CH_2I]Re(CO)_3$ (**4**, 185 mg, 0.357 mmol) was added $NaSSO_2CH_3$ (192 mg, 1.43 mmol) in four molar excess and DMF (8 mL). The reaction was left to stir overnight, and then diluted with water. Ethyl acetate extracts were washed with water and dried to give 174 mg (0.346 mmol) of product $[\eta^5\text{-}(C_5H_4)CH_2CH_2CH_2CH_2SSO_2CH_3]Re(CO)_3$ (**1a**) in 97% yield. IR (DMSO): ν_{CO} 1915.17 cm^{-1} as, 2014.17 cm^{-1} s. IR (1:1 CCl_4:CH_2Cl_2): ν_{CO} 1926.36 cm^{-1} as, 2022.36 cm^{-1} s. ^1H NMR: (499.8 MHz, $CDCl_3$) δ 5.195 (Cp, m, 4H), 3.262 (-CH_3, s, 3H), 3.116 (-CH_2S-, t, J = 7.3 Hz, 2H), 2.399 (CpCH_2-, m, 2H), 1.775 (C*H*$_2$CH$_2$S-, p, J = 7.6 Hz, 2H), 1.572 (CpCH$_2$C*H*$_2$-, m, 2H). ^{13}C NMR: (125.7 MHz, $CDCl_3$) δ 194.7 (*C*O), 110.6 (*ipso*-Cp), 84.0 (Cp), 83.2 (Cp), 50.9 (-*C*H$_3$), 36.1 (-*C*H$_2$S-), 30.7 (CpCH$_2$*C*H$_2$-), 29.4 (-*C*H$_2$CH$_2$S-), 27.7 (Cp*C*H$_2$-). EI mass $[M+Na]^+$ meas. 522.9766 m/z (calc. 522.9783 m/z).

$[\eta^5\text{-}(C_5H_4)CH_2CH_2CH_2SSO_2CH_3]Re(CO)_3$ (1b). To $[\eta^5\text{-}(C_5H_4)CH_2CH_2CH_2I]Re(CO)_3$ (**5**, 69.0 mg, 0.137 mmol) was added $NaSSO_2CH_3$ (101 mg, 0.753 mmol) in four molar excess and DMF (8 mL). The reaction was left to stir overnight. The product was purified by diluting the reaction mixture in water and extracting with ethyl acetate. The extracts were washed with water and dried to give 24.7 mg (0.0492 mmol) of product $[\eta^5\text{-}(C_5H_4)CH_2CH_2CH_2SSO_2CH_3]Re(CO)_3$ (**1b**) in 36% yield. IR (DMSO): ν_{CO} 1917.16 cm^{-1} as, 2015.57 cm^{-1} s. IR (1:1 CCl_4:CH_2Cl_2): ν_{CO} 1931.00 cm^{-1} as, 2024.48 cm^{-1} s. ^1H NMR: (499.8 MHz, $CDCl_3$) δ 5.211 (Cp, m, 4H), 3.272 (-CH_3, s, 3H), 3.141 (-CH_2S-, t, J = 7.2 Hz, 2H), 2.508 (CpCH_2-, m, J(H,H) = 7.8 Hz, 2H), 1.775 (CH$_2$C*H*$_2$-CH$_2$-, p, J = 7.5 Hz, 2H). ^{13}C NMR: (125.7 MHz, $CDCl_3$) δ 194.4 (*C*O), 109.1 (*ipso*-Cp), 84.1 (Cp), 83.4 (Cp), 51.0 (-*C*H$_3$), 35.7 (-*C*H$_2$S-), 31.7 (CpCH$_2$*C*H$_2$-), 26.998 (Cp*C*H$_2$-). EI mass $[M+Na]^+$ meas. 508.9605 m/z (calc. 508.9627 m/z).

$[\eta^5\text{-}(C_5H_4)CH_2I]Re(CO)_3$ (7). A solution of $[\eta^5\text{-}(C_5H_4)CH_2OH]Re(CO)_3$ (**6**, 396 mg, 1.13 mmol) and NaI (163 mg, 1.21 mmol) in ACN (15 mL) was stirred under a stream of nitrogen as TMS-Cl (137 µL, 1.08 mmol) was added dropwise. The resulting solution was stirred at room temperature for 20 minutes. Ether (15 mL) was added to the reaction mixture to extract the product, which was successively washed with water (15 mL), thiosulfate (10%, 15 mL), and brine (15 mL). The ether extracts were dried with $MgSO_4$. Celite filtration and

removal of the solvent afforded 455 mg (0.957 mmol, 84% yield) of $[\eta^5\text{-}(C_5H_4)CH_2I]Re(CO)_3$. ^{1}H NMR: (300.1 MHz, $CDCl_3$) δ 5.513 (Cp, t, J = 2.2 Hz, 2H), 5.289 (Cp, t, J = 2.2 Hz, 2H), 4.1444 (-C*H*$_2$-, s, 2H). ^{13}C NMR: (75.4 MHz, $CDCl_3$) δ 193.7 (*C*O), 104.0 (*ipso*-Cp), 85.6 (Cp), 84.4 (Cp), 29.9 (-*C*H$_2$S-).

$[\eta^5\text{-}(C_5H_4)CH_2SSO_2CH_3]Re(CO)_3$ (1c). $[\eta^5\text{-}(C_5H_4)CH_2I]Re(CO)_3$ (**7**, 454.9 mg, 0.957) and $NaSSO_2CH_3$ (642 mg, 4.79 mmol) were dissolved in DMF (8 mL) and stirred overnight. An aqueous workup afforded $[\eta^5\text{-}(C_5H_4)CH_2SSO_2CH_3]Re(CO)_3$ (**1c**, 79.0 mg, 0.172 mmol) in 18% yield. IR (DMSO): ν_{CO} 1914.88 cm^{-1} as, 2014.04 cm^{-1} s. IR (1:1 CCl_4:CH_2Cl_2): ν_{CO} 1930.72 cm^{-1} as, 2024.06 cm^{-1} s. ^{1}H NMR: (499.8 MHz, $CDCl_3$) δ 5.483 (Cp, t, J = 2.2 Hz, 2H), 5.262 (Cp, t, J = 2.2 Hz, 2H), 4.064 (CpC*H*$_2$S-, s, 2H), 3.228 (-CH_3, s, 3H). ^{13}C NMR: (125.7 MHz, $CDCl_3$) δ 193.3 (*C*O), 101.6 (*ipso*-Cp), 85.7 (Cp), 84.7 (Cp), 51.5 (-*C*H$_3$), 33.9 (-*C*H$_2$S-). EI mass $[M+Na]^+$ meas. 480.9325 m/z (calc. 480.9314 m/z).

Mass Spectrometry of Labeled Proteins. Labeled proteins were characterized by MALDI mass spectrometry.

α–Synuclein Mutant. V71C-ReL4 $[M+H]^+$ meas. 14,884.0 m/z (calc. 14,887.1 m/z).
Ubiquitin Mutants. (K6C-ReL1 calc. 9035.3, meas. 9042.3 $[M+H^+]$; K6C-ReL3 calc. 9063.3, meas. 9068.2 $[M+H^+]$; K6C-ReL3 calc. 9077.4, meas. 9077.2 $[M+H^+]$; K63C-ReL1 calc. 8919.2, meas. 8920.2 $[M+H^+]$; K63C-ReL3 calc. 8948.3, meas. 9488.9 $[M+H^+]$; K63C-ReL4 calc. 8961.3, meas. 8961.8 $[M+H^+]$).

Rendering. The method for drawing Fig. 4 from the main text is described as follows. The ubiquitin structure (PDB ID: 1ubq) was imported into UCSF Chimera.[7] With this program, the labeled residues were mutated to cysteine (for the K6C sample, an aspartate residue was added to the end of the sequence). To the cysteine residue, the label, $CpRe(CO)_3$ XRD structure,[8] was attached through the sulfur atom. For each of the mutants, UCSF Chimera rendered a hydrophobic surface, based each residue's hydrophobicity as assessed by Kyte and Doolittle.[9] The coloring of the surfaces ranges from red (hydrophobic) to blue (hydrophilic). The circles in Fig. 4b-c are centered about the cysteine α–C to β–C bond axis. The circle radii are the estimated reach of the label from the sulfur atom in cysteine to the carbonyl ligands.

Additional 2D IR Data. 2D IR spectra for each label in DMSO and in 1:1 CH_2Cl_2/CCl_4 is presented in Fig. S1, along with the spectra for the labels covalently bonded to two positions on ubiquitin. Intensity through the diagonal for the labels in polar and nonpolar solvents is shown in Fig. S2. FTIR spectra for the labels in solvents is overlaid in Fig. S3. The 2D IR waiting time intensities are plotted with corresponding single exponential fits for the label in solvents (Fig. S4) and the label attached to ubiquitin (Fig. S5).

Figure S1. 2D IR spectra for each label in DMSO, $CHCl_2/CCl_4$, on K6C, and K63C.

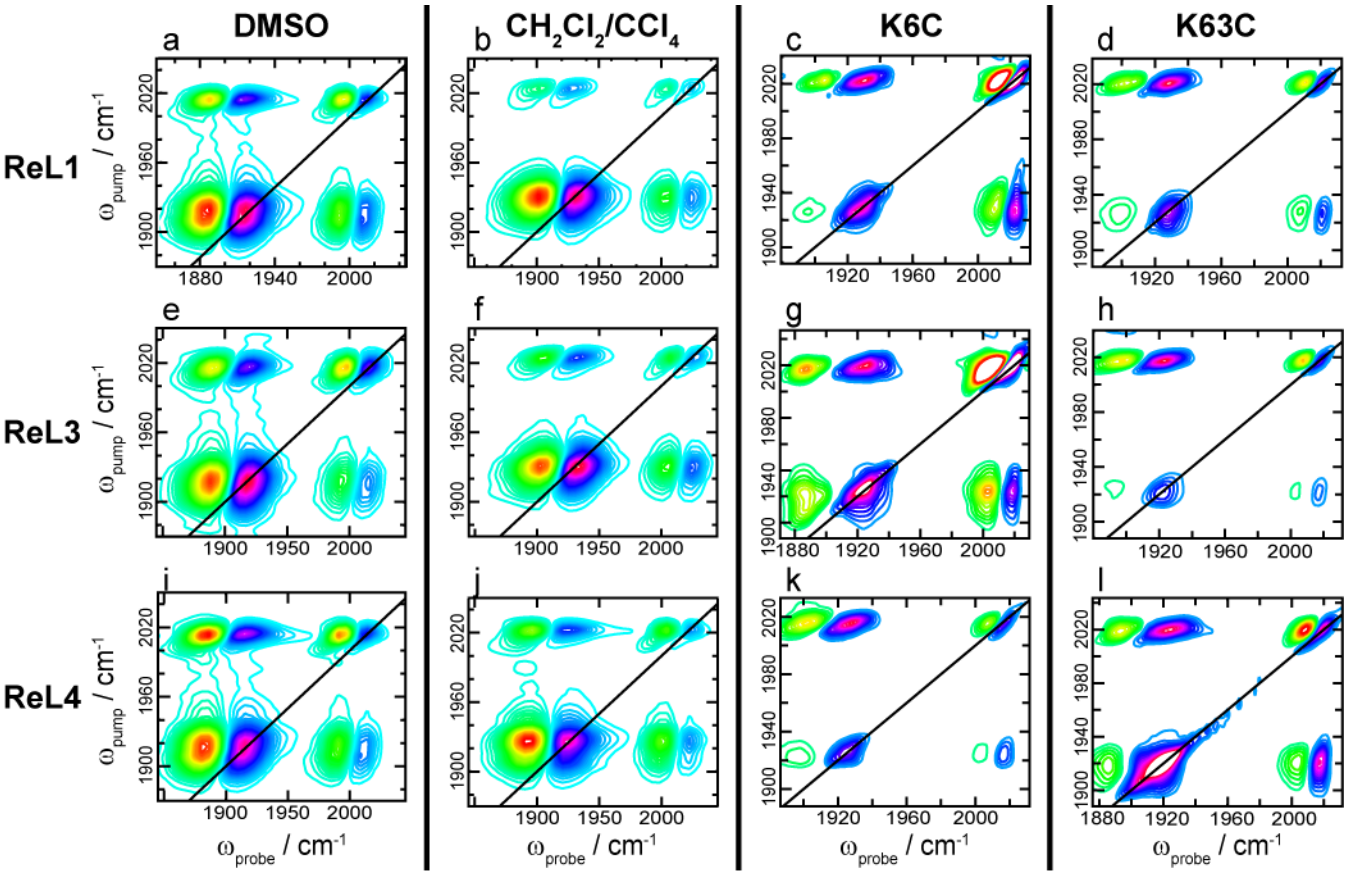


Figure S2. 2D IR diagonal slices for the low (a) and high (b) frequency modes.

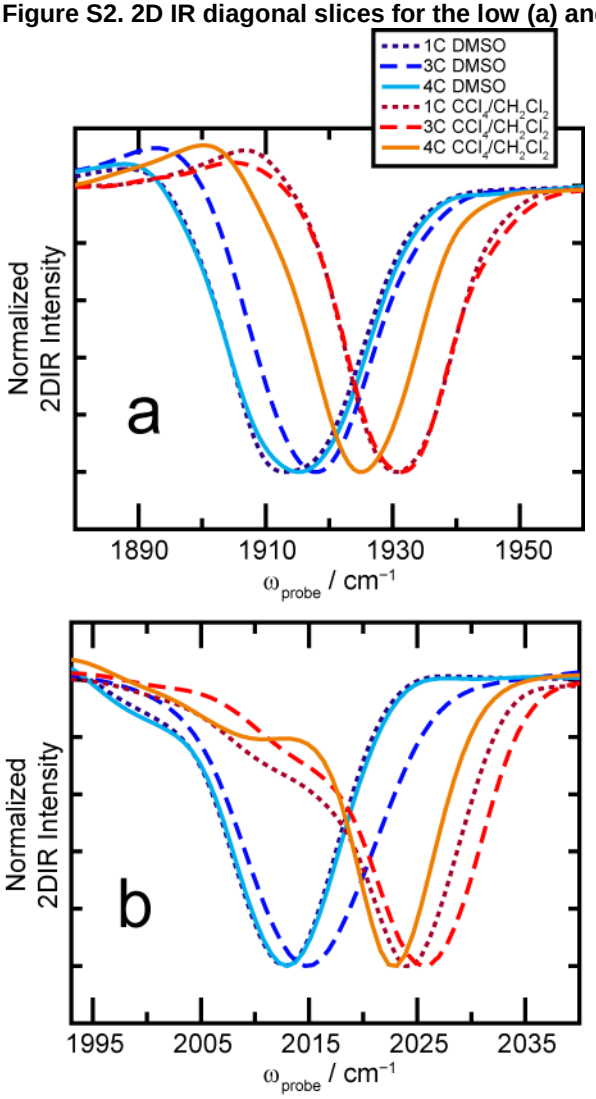


Figure S3. FTIR spectra of the labels in DMSO or 1:1 CH2Cl2/CCl4.

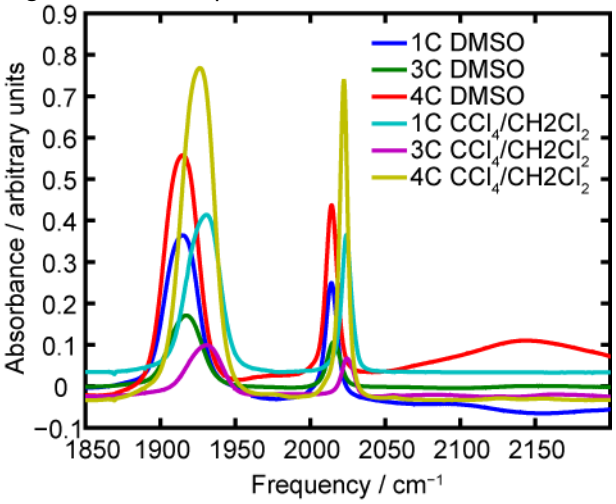


Figure S4. Symmetric stretch waiting time data for each label in DMSO or 1:1 CH_2Cl_2/CCl_4.

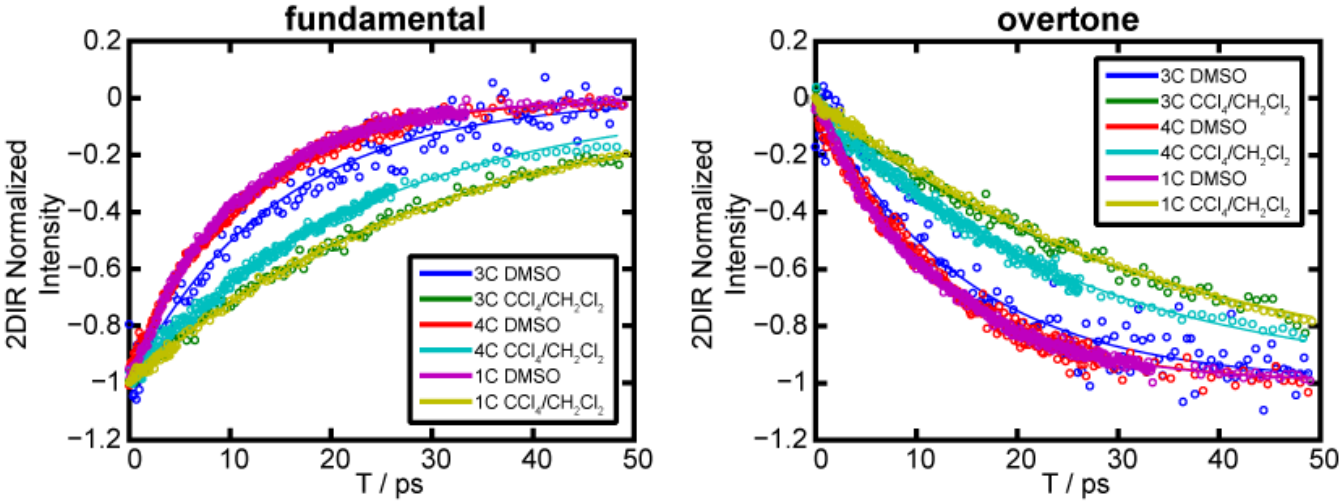
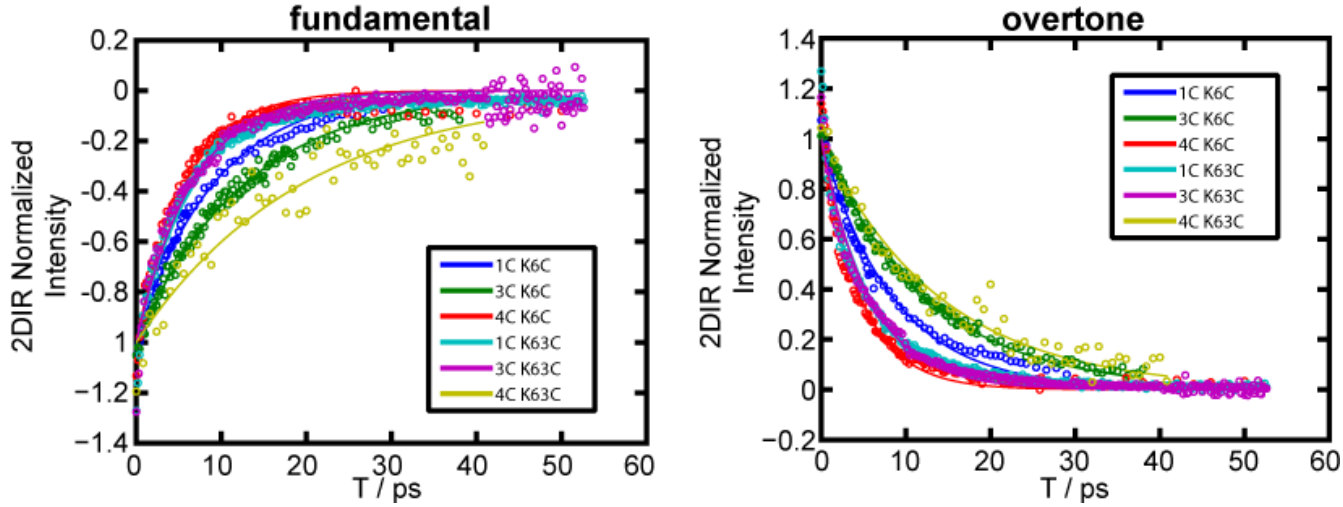


Figure S5. Symmetric stretch waiting time data for each label on K6C or K63C.



References

(1) Agbossou, F.; Ramsden, J. A.; Huang, Y. H.; Arif, A. M.; Gladysz, J. A. *Organometallics* **1992**, *11*, 693-701.
(2) Herberhold, M.; Biersack, M. *J. Organomet. Chem.* **1995**, *503*, 277-287.
(3) Mavunkal, I. J.; Moss, J. R.; Bacsa, J. *J. Organomet. Chem.* **2000**, *593-594*, 361-368.
(4) Yang, G. K.; Bergman, R. G. *Organometallics* **1985**, *4*, 129-138.
(5) Pangborn, A. B.; Giardello, M. A.; Grubbs, R. H.; Rosen, R. K.; Timmers, F. J. *Organometallics* **1996**, *15*, 1518-1520.
(6) Still, W. C.; Kahn, M.; Mitra, A. *J. Org. Chem.* **1978**, *43*, 2923-2925.
(7) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Comput. Chem* **2004**, *25*, 1605-1612.
(8) Fitzpatrick, P. J.; Lepage, Y.; Butler, I. S. *Acta Crystallogr B* **1981**, *37*, 1052-1058.
(9) Kyte, J.; Doolittle, R. F. *J. Mol. Biol.* **1982**, *157*, 105-132.