

Supporting Information for "Monitoring Long-range Electron Transfer Pathways in Proteins by Stimulated Attosecond Broadband X-Ray Raman Spectroscopy"

Yu Zhang,[†] Jason D. Biggs,[†] Niranjan Govind,[‡] and Shaul Mukamel^{*,†}

Dept. of Chemistry, University of California, 450 Rowland Hall, Irvine, California 92697, USA, and William R. Wiley Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, P. O. Box 999, Richland, Washington 99352, USA

E-mail: smukamel@uci.edu

*To whom correspondence should be addressed

[†]Dept. of Chemistry, University of California, 450 Rowland Hall, Irvine, California 92697, USA

[‡]William R. Wiley Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, P. O. Box 999, Richland, Washington 99352, USA

Computational Details

Quantum Chemistry Calculation

Core excitation calculations require all-electron basis sets for heavy atoms such as Re, which makes the computational cost prohibitive for the entire system. To make the calculations feasible, we dissected the system into three molecular fragments which represent the Re complex, the tryptophan group, and the Cu complex, out of the original protein structure (PDB id: 2i7o),¹ as shown in Fig. S5 in Supporting Information. The amino acid residues are terminated by methyl groups. Our X-ray pulses generate localized excited states. Since there is a long distance (19.4 Å) between the Re and Cu center, we only considered the Re-complex and tryptophan fragments in Re and Cl signal calculations; and for the Cu signal calculations, we only included the Cu complex fragment.

Kinetic rate constants are taken Ref. 1. Eqs. 1 are solved with the *Mathematica* software² and the time-dependent populations for all states are obtained. Time-dependent spectroscopy signals are averaged from signals of each states with the corresponding populations. The same initial conditions were taken from Ref. 1: $[A]_0 = 0.8$, $[B]_0 = 0.0$, $[C]_0 = 0.2$ and $[D]_0 = 0.0$. We expect core excitations to be insensitive to vibrational relaxation, therefore the signal contributions from the states A and B (see Fig. 2 in the main text) are lumped together.

We have used the zeroth order regular approximation (ZORA)³⁻⁵ to account for relativistic effects and a relativistic all-electron basis set SARC-ZORA⁶ for Re. The def2-TZVP⁷ and 6-311G** basis sets were used for Cu and Cl, respectively. The 6-31G* basis set was used for all the other atoms. Such a large basis set (~ 800 basis functions for the Re-complex-tryptophan fragment) causes serious SCF convergence difficulties. We used the Δ SCF method to obtain the MLCT and Cu (II) state (see intermediate states A, B and D in Fig. 2 in the main text). There is a recent revival in using Δ SCF for excited state calculations.^{8,9} Better results for core and charge-transfer excitations compared with linear

response TDDFT can be obtained since complete orbital relaxation is included. We have employed occupation number control and the maximum overlap method¹⁰ as implemented in NWChem¹¹ to avoid variational collapse and to converge the excited state wavefunctions.

Core excitations were calculated with restricted excitation window time-dependent density functional theory (REW-TDDFT)¹²⁻¹⁴ and the Tamm-Dancoff approximation.¹⁵ To better describe the charge-transfer states, the long-range corrected density functional LC-PBE0¹⁶⁻¹⁸ was used. The conductor-like implicit solvent model (COSMO)¹⁹ was used to represent the solvent (water) effects. Transition dipole calculation protocols have been documented in Ref. 14,20. All quantum chemistry calculations were performed with a locally modified version of the NWChem code.¹¹

Spectroscopy Signal Calculation

For transient X-ray absorption signals (see Eq. 2 in the main text) we use a Gaussian pulse envelope

$$\mathcal{E}_2(\omega) = \mathcal{E}_2 e^{-\sigma_2^2(\omega-\omega_2)^2/2} \tag{S1}$$

with $\sigma_2 = 200$ as which corresponds to a 10 eV bandwidth (FWHM). The central pulse frequencies for the copper, chlorine, and rhenium L-edge signals are set to 929, 200, and 10,866 eV, respectively. Our pulses are much shorter than the relevant picosecond-to-nanosecond ET timescale. Such short pulses are required not for the time resolution, but for the bandwidth to be broad enough to cover the absorption and emission spectra in order to induce the Raman transitions of interest.

For stimulated X-ray Raman signals we use Gaussian pulses with the same carrier frequencies as in the transient absorption, but a broader bandwidth $\sigma = 100$ as. The pump and the probe pulses are parallel polarized and the signals are averaged over a randomly oriented ensemble.

Loop Diagrams for the SXRS Signal

Fig. S1 shows the pulse sequence and loop diagrams for the SXRS experiment. These diagrams are similar to the double-sided ladder diagrams traditionally used to describe non-linear optical spectroscopy, with a few key differences. The loop diagram is read starting with the ground state on the lower left corner and moving upwards, describing forward evolution (propagation of the ket). Interactions with the field are represented by arrows facing into (absorption of a photon) and out of (emission of a photon) the diagram. At the top of the loop we switch to backward time evolution (propagation of the bra), and finally arrive back at the ground state. Thus we are able to work in Hilbert space, rather than the higher dimensional Liouville space required with the density matrix.²¹

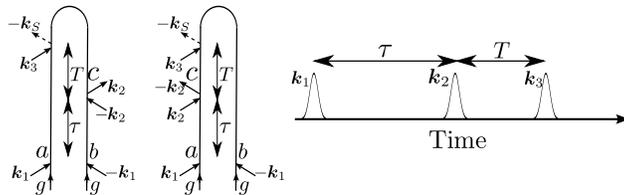


Figure S1: Loop diagrams (left) and pulse sequence (right) for the integrated two-pulse stimulated X-ray Raman Scattering signal. In Ref. 26 this was denoted I2P-SXRS. For diagram rules see.²²

Sampling Time and SXRS Spectral Resolution

In previous work,^{14,23} we assumed an infinite sampling time. In Fig. S2, we illustrate the effects of a finite-duration sampling period on the resulting SXRS spectrum. The signal was collected in the time domain via Eq. 4 in the main text, using a valence linewidth $\Gamma_b = 0.01$ eV (corresponding to an effective valence lifetime of $\hbar/\Gamma_b = 65.82$ fs) for a variable total length of time, using a timestep of 200 as. A Hamming window function is used to remove edge effects.²⁴ Fig. S2 shows how the spectral resolution is increased with the collection time. A collection period of 200 fs was used in the signals presented in Fig. 4 in the main text.

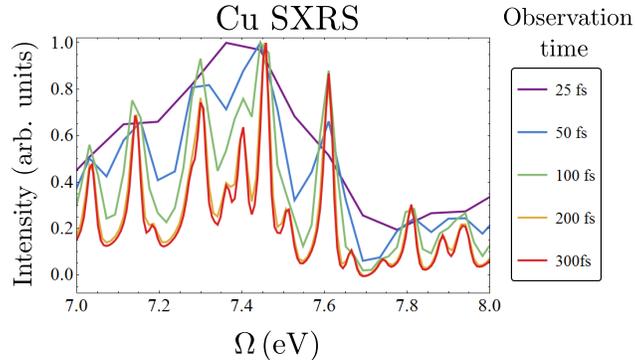


Figure S2: The effects of the collection period on the resolution of the SXRS signal. The signal is collected in the time domain, in steps of 200 as, for the observation periods listed on the right. The frequency resolution is inversely proportional to the sampling period. For the linewidths used in this simulation, $\Gamma_b = 0.01 \text{ eV}$, it is sufficient to use a 200 fs collection period.

X-ray Raman Spectroscopy with a Broadband and a Narrowband Pulse

This different pulse setup compared to Fig. S1 uses three pulses. The first pulse \mathbf{k}_1 is the actinic pump, initiating the electron transfer reaction. After a delay time τ , pulses \mathbf{k}_2 and \mathbf{k}_3 arrive simultaneously at the sample. Pulse \mathbf{k}_2 is narrowband, with frequency ω_2 , and on resonance with a particular core transition. \mathbf{k}_3 is broadband and redshifted from the first as shown in Fig. S3. The signal is given by the dispersed \mathbf{k}_3 pulse spectrum with the Raman pump minus that without. This technique is commonly used in vibrational Raman spectroscopy, where it is known as femtosecond stimulated Raman spectroscopy (FSRS).^{25,26} It is commonly used to report the changing vibrational modes throughout some chemical reaction (i.e. isomerization).

In the X-ray extension of the FSRS technique, herein dubbed AXRS (A for attosecond), it is the response of the valence electrons to transient excitation of the core that is probed. The Raman pump \mathbf{k}_2 promotes a core electron to a virtual orbital. The Raman probe then stimulates a valence electron to fill the core hole. Since the Raman pump and probe do not overlap in frequency, the electron that fills the core hole will come from an occupied valence

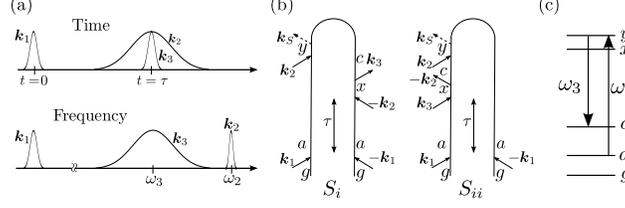


Figure S3: (a) Pulse sequence for the narrow pump/broad probe stimulated Raman experiment, shown in the time and frequency domains. (b) The two contributing loop diagrams. (c) Energy level diagram: Shown are the ground state g , valence-excited states a, b , and c , and core-excited states x and y .

orbital. The broad bandwidth of the Raman probe stimulates many such transitions, and the signal will contain peaks where $\omega_2 - \omega = \omega_{ca}$, where ω is the dispersed frequency and ω_{ca} is a valence excitation frequency.

The general expression for the signal is given by Eq. (11) of Ref.²⁶ Here we do not treat the interaction with the actinic pump explicitly. Rather, we take the results from the kinetic model described above and assume that the electronic Hamiltonian does not change considerably during interaction with the Raman pump and probe. This assumption is well justified, as the X-ray Raman process is limited by the core lifetime (~ 5 fs), and the nuclei do not move significantly during that time.

The signal is given by

$$S(\tau, \omega) = S_i(\tau, \omega) + S_{ii}(\tau, \omega) \quad (\text{S2})$$

where

$$S_i(\tau, \omega) = \frac{2}{\hbar^4} \Im \sum_{ac} \rho_{aa}(\tau) \frac{|\alpha_{ca}(\omega_2) \mathcal{E}_2 \mathcal{E}_3(\omega)|^2}{\omega - (\omega_2 - \omega_{ca}) + i\gamma_{ca}} \quad (\text{S3})$$

$$S_{ii}(\tau, \omega) = \frac{2}{\hbar^4} \Im \sum_{ac} \rho_{aa}(\tau) \frac{|\alpha_{ca}(\omega) \mathcal{E}_2 \mathcal{E}_3(\omega)|^2}{\omega - (\omega_2 + \omega_{ca}) + i\gamma_{ca}}$$

and

$$\alpha_{ij}(\omega) = \sum_z \frac{V_{iz} V_{zj}^\dagger}{\omega - \omega_{zj} + i\gamma_{zj}} \quad (\text{S4})$$

By selecting $\mathbf{k}_3(\omega)$ is to the red of $\mathbf{k}_2(\omega)$, and they do not overlap. As we can see from Eqs. S3, under the condition that $\mathbf{k}_3(\omega)$ is only nonzero for frequencies below ω_2 , S_i corresponds to a Stokes process ($\omega_{ca} > 0$), and S_{ii} corresponds to an anti-Stokes process ($\omega_{ca} < 0$). S_i therefore represents time domain RIXS, where we use shaped pulses (broad and narrow) to discriminate against elastic contributions.

Fig. S4 shows the AXRS signal taken at the copper L-edge for the two different oxidation states. We can see that the AXRS and SXRS signals are very similar. The differences arise from the different weighting applied by the broadband pulse envelopes. Note that the SXRS signal in the lower right panel of Fig. S4, the low-energy peaks between 1 and 2 eV are dominant, while in the AXRS signal they are weaker than the peaks between 3 and 4 eV. This is due to the fact that, in AXRS the broadband pulse is chosen to not overlap with the narrowband frequency, resulting in a discrimination against low-frequency terms. There is no such spectral selection in SXRS.

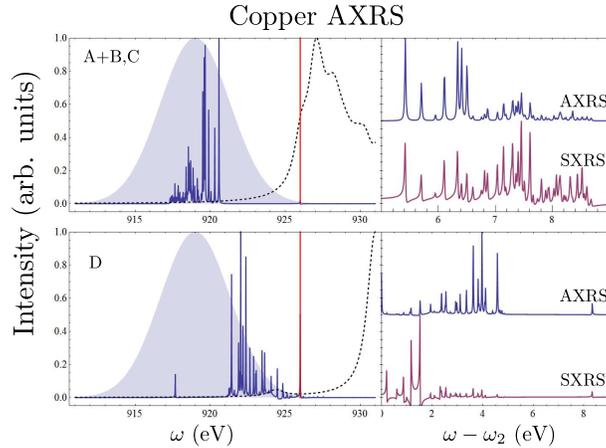


Figure S4: AXRS signals taken at the copper L-edge for the different states in the ET process. The top panel shows the signal from A+B, and C states, while the bottom corresponds to the D state. On the left we show the XANES spectrum (dashed black line), narrowband \mathbf{k}_2 frequency (red line), broadband \mathbf{k}_3 pulse power spectrum (shaded blue region), and AXRS signal (blue line). The right panel shows the AXRS signal plotted as a function of $\omega - \omega_2$ on top and the corresponding SXRS signal below.

Video S1

We include a movie file to show the correlation between the electronic density and the SXRS signals. In this movie, the top row shows the time-dependent electron density difference, which has the ground-state density subtracted out, for the three molecular fragments: the rhenium complex on the left, the chlorine-substituted tryptophan in the middle, and the copper complex on the right. Red denotes a negative sign (hole) and blue denotes a positive sign (electron). The middle row shows the interpulse delay and the time-dependent populations of the various species involved in the ET process. The bottom row shows the SXRS signals at the rhenium, chlorine, and copper L-edges.

Molecular Fragment Model of the Re-modified Azurin System

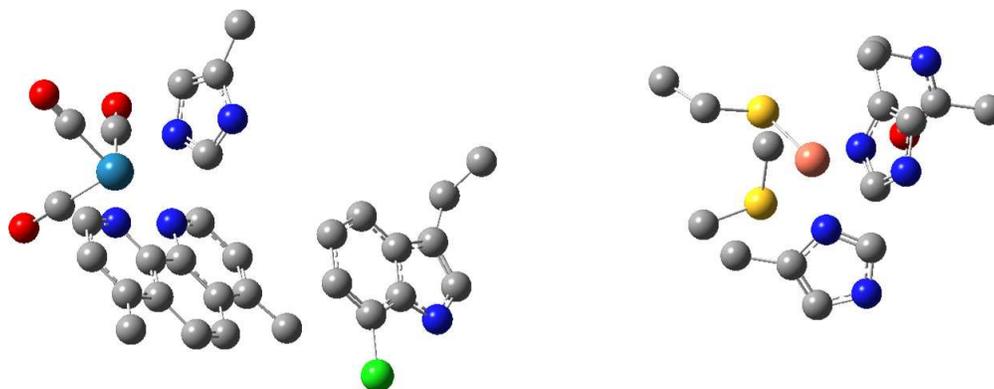


Figure S5: Molecular fragment model of the Re-modified azurin system. On the left is the Re-complex, in the middle is the tryptophan bridge, and on the right is the Cu-complex. O is in red, Re is in light blue, N is in deep blue, Cl is in green, S is in yellow, Cu is in pink and C is in gray. All H atoms are omitted for simplicity.

References

- (1) Shih, C.; Museth, A. K.; Abrahamsson, M.; Blanco-Rodriguez, A. M.; Bilio, A. J. D.; Sudhamsu, J.; Crane, B. R.; Ronayne, K. L.; Towrie, M.; Vlček Jr., A. et al. Tryptophan-Accelerated Electron Flow Through Proteins. *Science* **2008**, *320*, 1760–1762.
- (2) 2013; Wolfram Research, Inc., Mathematica, Version 9.0, Champaign, IL.
- (3) Chang, C.; Pelissier, M.; Durand, M. Regular Two-Component Pauli-Like Effective Hamiltonians in Dirac Theory. *Phys. Scr.* **1986**, *34*, 394–404.
- (4) van Lenthe, E.; Baerends, E. J.; Snijders, J. Relativistic Total Energy Using Regular Approximations. *J. Chem. Phys.* **1994**, *101*, 9783–9792.
- (5) Faas, S.; Snijders, J.; van Lenthe, J.; van Lenthe, E.; Baerends, E. The ZORA formalism Applied to the Dirac-Fock Equation. *Chem. Phys. Lett.* **1995**, *246*, 632–640.
- (6) Pantazis, D. A.; Chen, X.-Y.; Landis, C. R.; Neese, F. All-electron Scalar Relativistic Basis Sets for Third-row Transition Metal Atoms. *J. Chem. Theory Comput.* **2008**, *4*, 908–919.
- (7) Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
- (8) Besley, N. A.; Gilbert, A. T. B.; Gill, P. M. W. Self-consistent-field Calculations of Core Excited States. *J. Chem. Phys.* **2009**, *130*, 124308.
- (9) Kowalczyk, T.; Yost, S. R.; Van Voorhis, T. Assessment of the Δ SCF Density Functional Theory Approach for Electronic Excitations in Organic Dyes. *J. Chem. Phys.* **2011**, *134*, 054128.

- (10) Gilbert, A. T. B.; Besley, N. A.; Gill, P. M. W. Self-Consistent Field Calculations of Excited States Using the Maximum Overlap Method (MOM). *J. Phys. Chem. A* **2008**, *112*, 13164–13171.
- (11) Valiev, M.; Bylaska, E.; Govind, N.; Kowalski, K.; Straatsma, T.; van Dam, H.; Wang, D.; Nieplocha, J.; Apra, E.; Windus, T. et al. NWChem: A Comprehensive and Scalable Open-source Solution for Large Scale Molecular Simulations. *Comput. Phys. Commun.* **2010**, *181*, 1477–1489.
- (12) Stener, M.; Fronzoni, G.; de Simone, M. Time Dependent Density Functional Theory of Core Electrons Excitations. *Chem. Phys. Lett.* **2003**, *373*, 115–123.
- (13) Lopata, K.; Van Kuiken, B. E.; Khalil, M.; Govind, N. Linear-response and Real-time Time-dependent Density Functional Theory Studies of Core-level Near-edge X-ray Absorption. *J. Chem. Theory. Comput.* **2012**, *8*, 3284–3292.
- (14) Zhang, Y.; Biggs, J. D.; Healion, D.; Govind, N.; Mukamel, S. Core and Valence Excitations in Resonant X-ray Spectroscopy Using Restricted Excitation Window Time-dependent Density Functional Theory. *J. Chem. Phys.* **2012**, *137*, 194306.
- (15) Hirata, S.; Head-Gordon, M. Time-dependent Density Functional Theory within the Tamm-Dancoff Approximation. *Chem. Phys. Lett.* **1999**, *314*, 291–299.
- (16) Tawada, Y.; Tsuneda, T.; Yanahisawa, S.; Yanai, T.; Hirao, K. A Long-range-corrected Time-dependent Density Functional Theory. *J. Chem. Phys.* **2004**, *120*, 8425–8433.
- (17) Yanai, T.; Tew, D. P.; Handy, N. A New Hybrid Exchange-correlation Functional Using the Coulomb-Attenuating Method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393*, 51–57.
- (18) Adamo, C.; Barone, V. Toward Reliable Density Functional Methods without Adjustable Parameters: The PBE0 Model. *J. Chem. Phys.* **1998**, *110*, 6158–6170.

- (19) Klamt, A.; Schüürmann, G. COSMO: A New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and Its Gradient. *J. Chem. Soc. Perkin Trans. 2* **1993**, 799–805.
- (20) Zhang, Y.; Healion, D.; Biggs, J.; Mukamel, S. Double-core Excitations in Formamide Can Be Probed by X-ray Double-quantum-coherence Spectroscopy. *J. Chem. Phys.* **2013**, *138*, 144301.
- (21) Mukamel, S.; Rahav, S. Ultrafast nonlinear optical signals viewed from the molecule’s perspective: Kramers–Heisenberg transition-amplitudes versus susceptibilities. *Advances in Atomic, Molecular, and Optical Physics* **2010**, *59*, 223–263.
- (22) Biggs, J.; Voll, J.; Mukamel, S. Coherent Nonlinear Optical Studies of Elementary Processes in Biological Complexes; Diagrammatic Techniques Based on the Wavefunction vs. the Density Matrix. *Phil. Trans. R. Soc. A* **2012**, *370*, 3709–3727.
- (23) Mukamel, S.; Healion, D.; Zhang, Y.; Biggs, J. D. Multidimensional Attosecond Resonant X-Ray Spectroscopy of Molecules: Lessons from the Optical Regime. *Annu. Rev. Phys. Chem.* **2013**, *64*, 101–127.
- (24) Oppenheim, A. V.; Schaffer, R. W. *Discrete Time Signal Processing*, 3rd ed.; Prentice Hall Signal Processing; Prentice Hall, New Jersey, USA, 2009.
- (25) Kukura, P.; McCamant, D. W.; Mathies, R. A. Femtosecond Stimulated Raman Spectroscopy. *Annu. Rev. Phys. Chem.* **2007**, *58*, 461–488.
- (26) Dorfman, K. E.; Fingerhut, B. P.; Mukamel, S. Time-resolved Broadband Raman Spectroscopies: A Unified Six-wave-mixing Representation. *J. Chem. Phys.* **2013**, *139*, 124113.