Supporting Information for "UV-Visible Lysine-Glutamate Dimer Excitations in Protein Charge Transfer Spectra: TDDFT Descriptions using an Optimally Tuned CAM-B3LYP Functional"

Imon Mandal[†], *Sutanuka Manna*[†], *Ravindra Venkatramani*[†]*

† Department of Chemical Sciences, Tata Institute of Fundamental Research, Homi Bhabha Road, Colaba, Mumbai 400 005, India.

*** Corresponding Author:**

Email: ravi.venkatramani@tifr.res.in

Phone: +91 (022) 2278 2064

CONTENTS

Page Number

A 6-fragment model for the Lys-Glu dimer with possible CT transitions

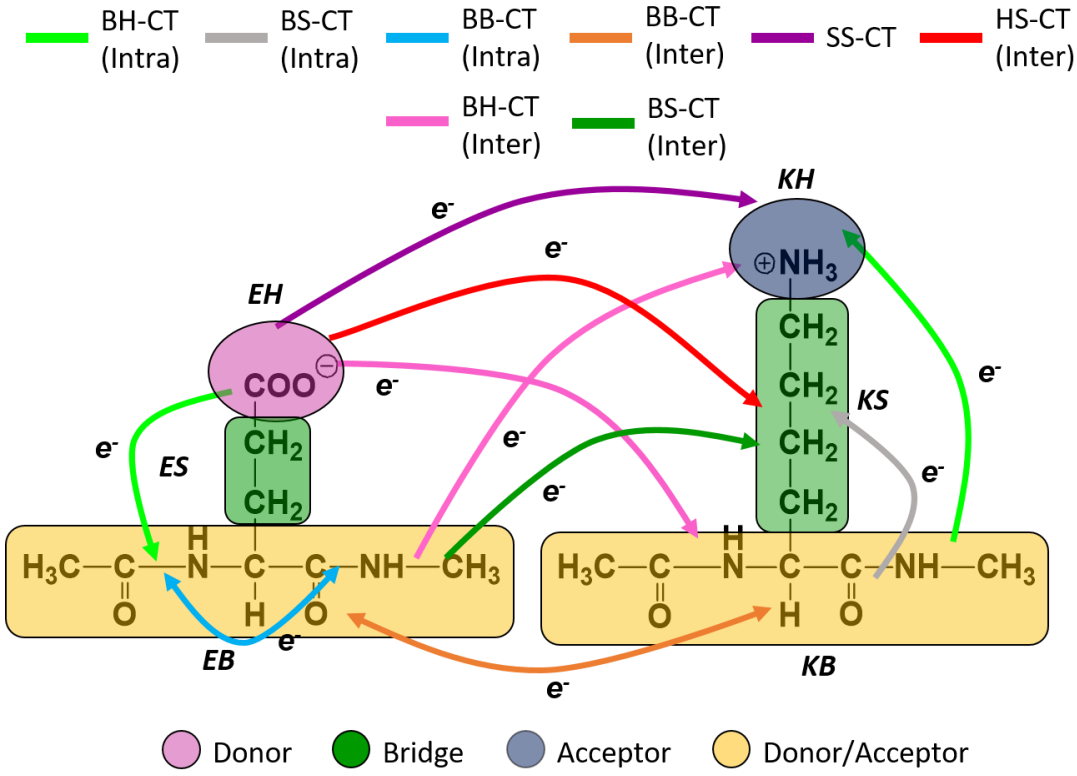

Figure S1: Lys-Glu dimer models considered in the present study depicting symmetric backbones with methyl capping. The shaded regions show the fragments (6-fragment model) used for electron/hole population analysis, wherein each residue (Lys/Glu) is split into headgroup (EH/KH), alkyl sidechain (ES/KS) and the peptide backbone (EB/KB) fragments. Each of the 6 fragments can potentially play the role of donor /acceptor groups for multiple distinct photoinduced electron transfer processes (colored arrows).

Control calculations on Glycine: Characterization of BB-CT and LE transitions

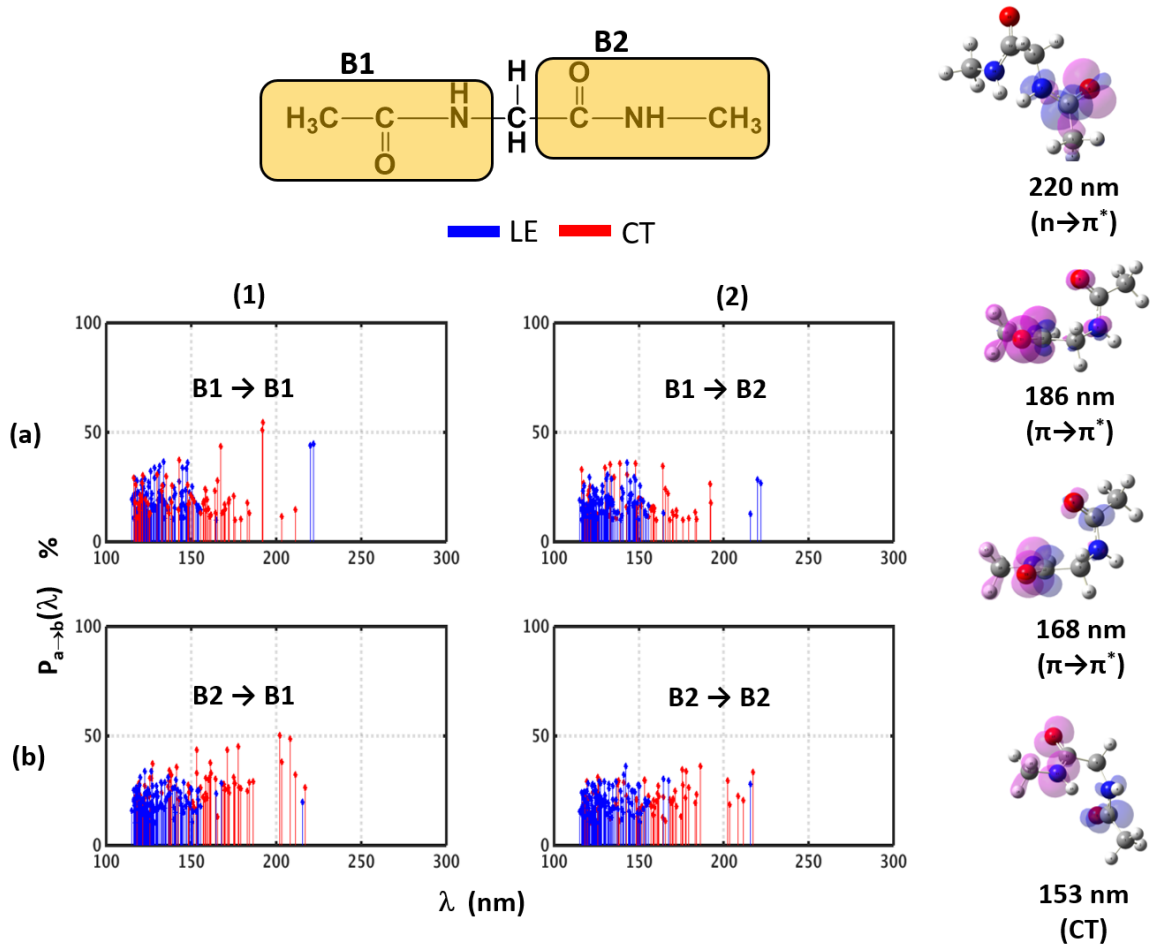

Figure S2: Percentage fragment pair contributions $P_{a\to b}$ (λ) (Eqn. 9) to ProCharTS transitions from a 2-fragment model for 2 representative conformations of Gly ((φ, ψ) =(96,11) and (68,23)) sampled from the α_3C MD trajectory (structure and fragments drawn on top). Each row (a-b) /column (1-2) panel shows $P_{a\to b}$ (λ) for a specific fragment pair (a=B1/B2)/(b=B1/B2) to hole/electron density of transitions classified as CT/LE (charge separation measures in section 2.2.1) and colored according to their character. Each panel only shows transitions with $P_{a\to b}$ (λ) > 10 %. For 4 selected transitions at 220 nm, 186 nm, 168 nm, and 153 nm, we show a Symm-BB (Gly) fragment with a difference density plot (Pink/blue lobes represent the hole/electron density).

Characterization of Lys-Glu dimer (weak interaction) spectra using the 6-fragment model

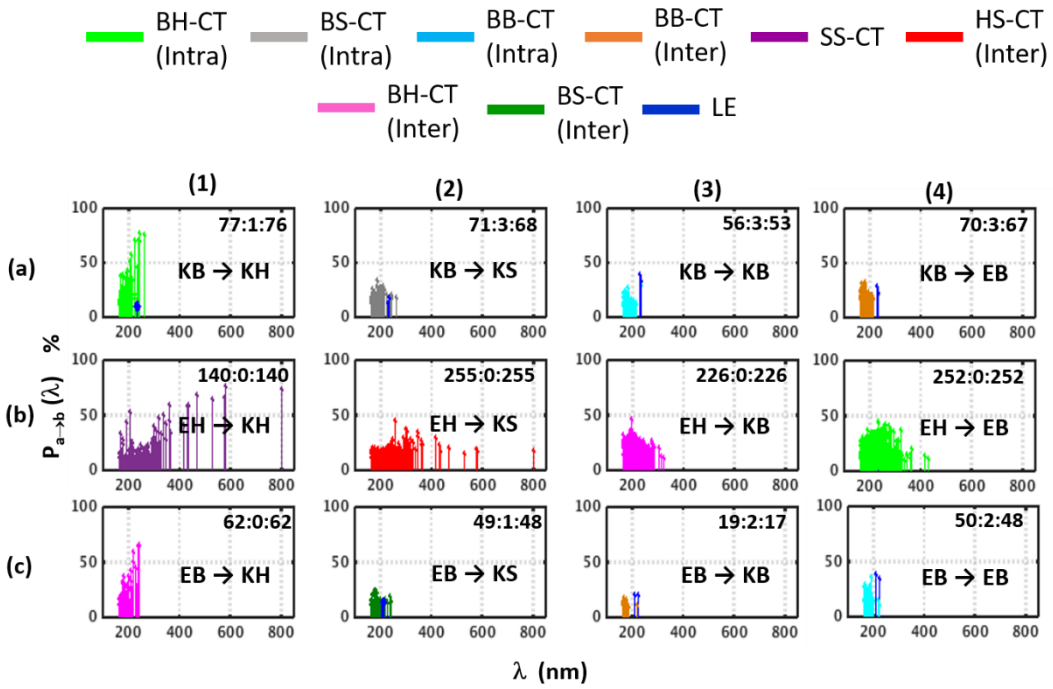

Figure S3: Percentage fragment pair contributions $P_{a\to b}$ (λ) (Eqn. 9) to ProCharTS transitions from 2 representative conformations of Lys-Glu dimers with sidechain separations of 5-6 Å using the 6-fragment model (Fig S1). Each row (a-c) /column (1-4) panel shows $P_{a\to b}$ (λ) for a specific fragment pair (*a*=KB/EH/EB)/(*b*=KH/KS/KB/EB) to hole/electron density of transitions classified as CT/LE (based on charge separation measures in section 2.2.1) and colored according to their character. The number of CT: LE:Total transitions (only those with $P_{a\to b}$ (λ) > 10 %) in each panel is indicated.

Characterization of Lys-Glu dimer (intermediate interaction) spectra using the 4-fragment model

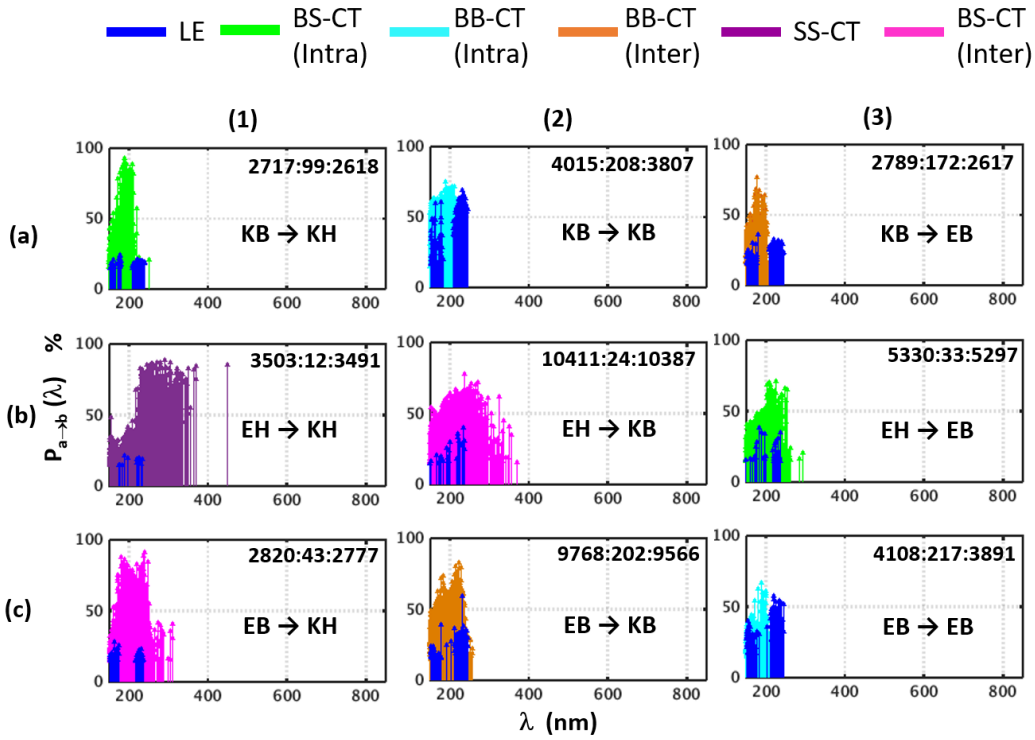

Figure S4: Percentage fragment pair contributions $P_{a\to b}$ (λ) (Eqn. 9) to ProCharTS transitions from 100 conformations of Lys-Glu dimers with sidechain separations of 3-4 Å (intermediate interactions) using the 4-fragment model. Each row (a-c) /column (1-3) panel shows $P_{a\to b}$ (λ) for a specific fragment pair (a=KB/EH/EB)/(b=KH/KB/EB) to hole/electron density of transitions classified as CT/LE (based on spatial charge separation measures in section 2.2.1) and colored according to their character. The number of CT:LE:Total transitions (only those with $P_{a\to b}$ (λ) > 10 %) in each panel is indicated.

Characterization of Lys-Glu dimer (strong interaction) spectra using the 4-fragment model

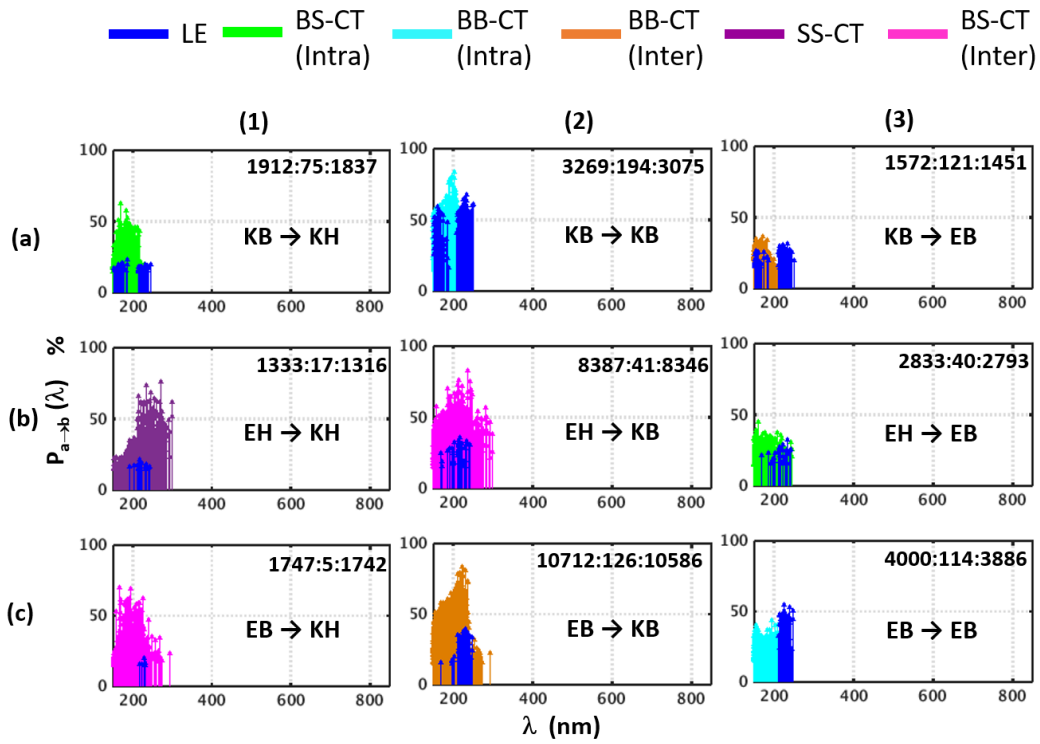

Figure S5: Percentage fragment pair contributions $P_{a\to b}$ (λ) (Eqn. 9) to ProCharTS transitions from 80 conformations of Lys-Glu dimers with sidechain separations of 2-3 Å (strong interactions) using the 4-fragment model. Each row (a-c) /column (1-3) panel shows $P_{a\to b}$ (λ) for a specific fragment pair (a=KB/EH/EB)/(b=KH/KB/EB) to hole/electron density of transitions classified as CT/LE (based on spatial charge separation measures in section 2.2.1) and colored according to their character. The number of CT:LE:Total transitions (only those with $P_{a\to b}$ (λ) > 10 %) in each panel is indicated.

CAM-B3LYP J-score landscape for Lys-Glu dimers: sensitivity to interaction strength

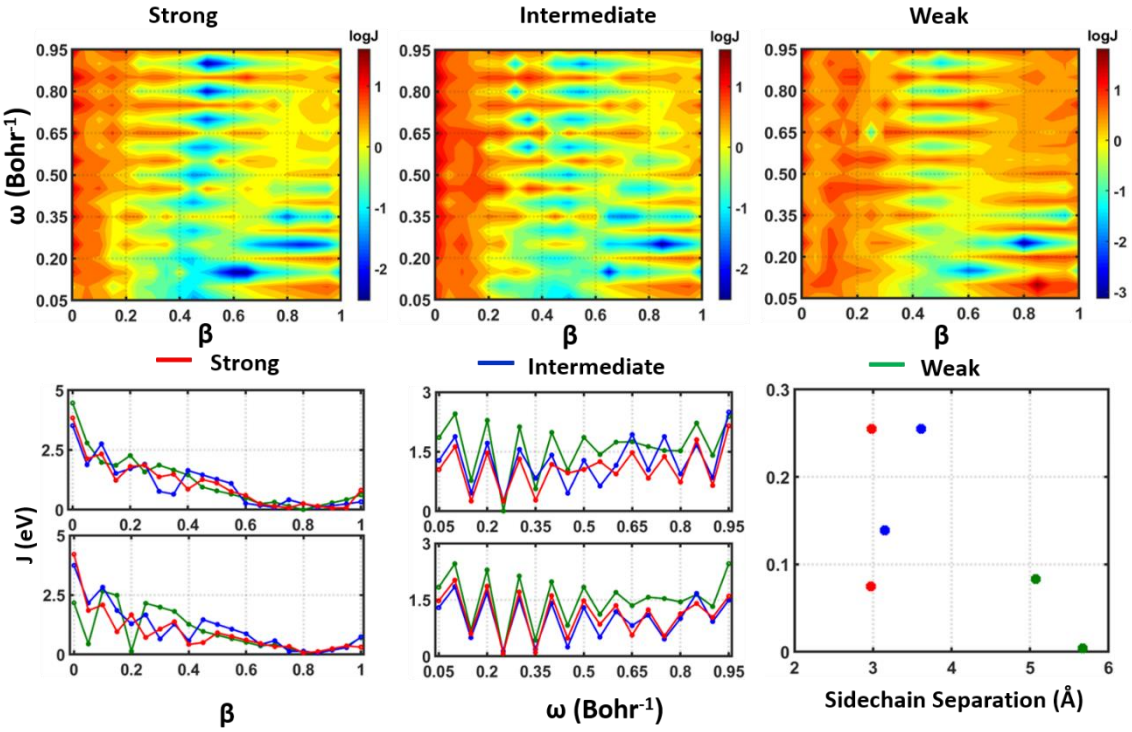

Figure S6: Top row: 2D plots of average $J(\omega,\beta)$ for 2 Lys-Glu dimer conformations sampled from α_3C MD structures at three sidechain separations (strong=2-3 Å, intermediate=3-4 Å, weak=5-6 Å). In all cases *J-score* minima are found at or very close to the optimal parameter values β=*0.80*, ω=*0.25*. **Bottom row:** 1D slices of $J(\beta, \omega=0.25)$ and $J(\beta=0.80, \omega)$ for the data in the top row (left and centre panels). *J*-score data for the 2 Lys-Glu conformations (top and bottom) with different sidechain separations (red, blue, green colors). Variation in optimal $J(\beta=0.80,\omega=0.25$ Bohr$^{-1})$ for each of the 6 Lys-Glu dimer structures as a function of Lys-Glu sidechain separations (right panel).

CAM-B3LYP J-score landscape for Lys-Glu dimers: sensitivity to sidechain separation

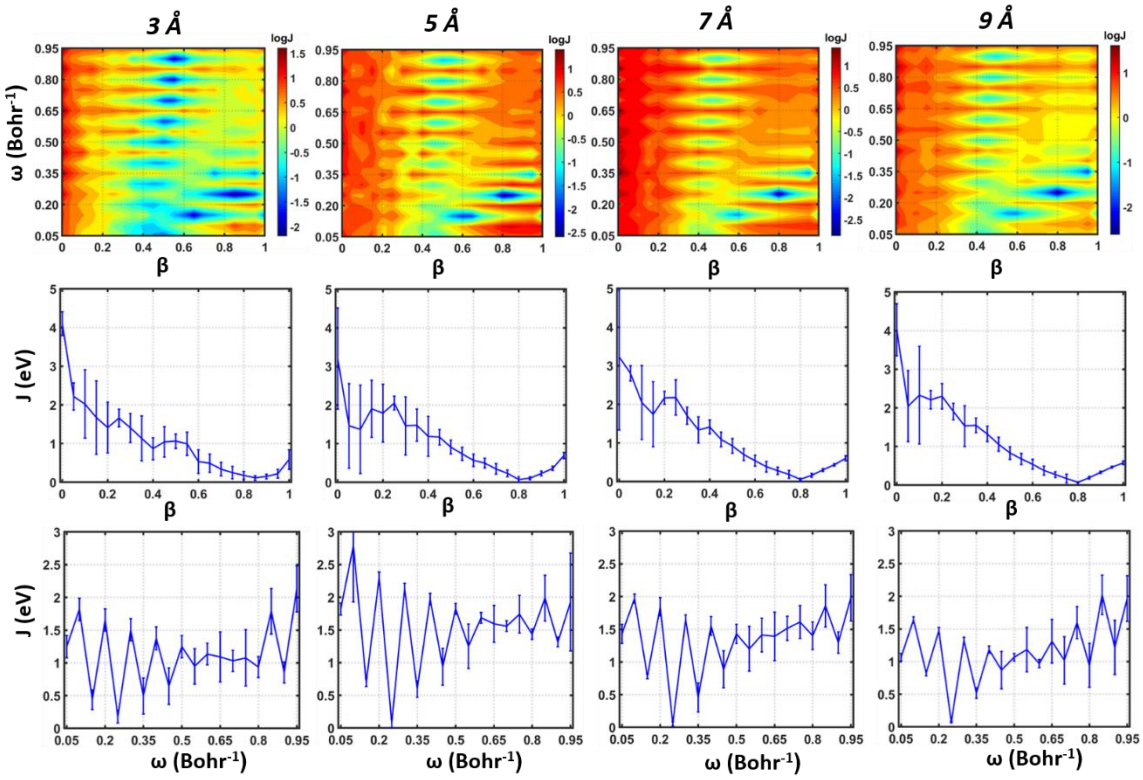

Figure S7: Top row: 2D plots of average $J(\omega,\beta)$ for 6 Lys-Glu dimer conformations sampled from α_3C MD structures at 4 sidechain separations ranging from 3-9 Å. In all cases *J*-score minima are found at or very close to the optimal parameter values β=*0.80*, ω=*0.25*. **Middle row:** 1D $J(\beta, \omega$=0.25) slices (averages and SD over 6 conformations) at different sidechain separations derived from the data in the top row. **Bottom row:** 1D $J(\beta$=*0.80*, $\omega)$ slices (averages and SD over 6 conformations) at different sidechain separations derived from the data in the top row.

CAM-B3LYP J-score landscape for Lys-Glu dimers: sensitivity to sidechain separation

Distance (monomer Separation), Å	Min J In eV					
	Conf 1	Conf 2	Conf 3	Conf 4	Conf 5	Conf 6
3	0.0220*	0.0540§	0.0183§	0.0443*	0.0255*	0.005*
5	0.0507	0.0216	0.0388 ζ	0.0353	0.0271	0.0973*
7	0.0443	0.0738	0.0752 ζ	0.0517	0.0429	0.0149
9	0.0756	0.0916	0.0771	0.0761	0.0703	0.0493

Min J scores in table were obtained using optimal tuned parameters: α=0.20 β=0.80 & ω=0.25 Bohr^{-1}

*Min J scores were obtained using optimal tuned parameters: α=0.35 β=0.65 & ω=0.15 Bohr^{-1}

§Min J scores were obtained using optimal tuned parameters: α=0.40 β=0.60 & ω=0.15 Bohr^{-1}

#Min J scores were obtained using optimal tuned parameters: α=0.30 β=0.70 & ω=0.15 Bohr^{-1}

ζMin J scores were obtained using optimal tuned parameters: α=0.15 β=0.85 & ω=0.25 Bohr^{-1}

Table S8: Optimal $J(\beta, \omega)$ values obtained from a complete parameter scan α/β and ω values subject to the constraints ($0 \leq \alpha \leq 1$; $\alpha + \beta = 1$; $0 < \omega < 1$) for the CAM-B3LYP functional. $J(\beta, \omega)$ were computed for 6 conformations x 4 sidechain separations = 24 Lys-Glu dimer structures to produce the 2D landscapes shown in Fig S7. The values shown in the table represent the minima in the 2D plots.

Effect of TDDFT Functional and Backbone Model on the Lowest Energy Transitions in the ProcharTS of weakly interacting Lys-Glu dimers

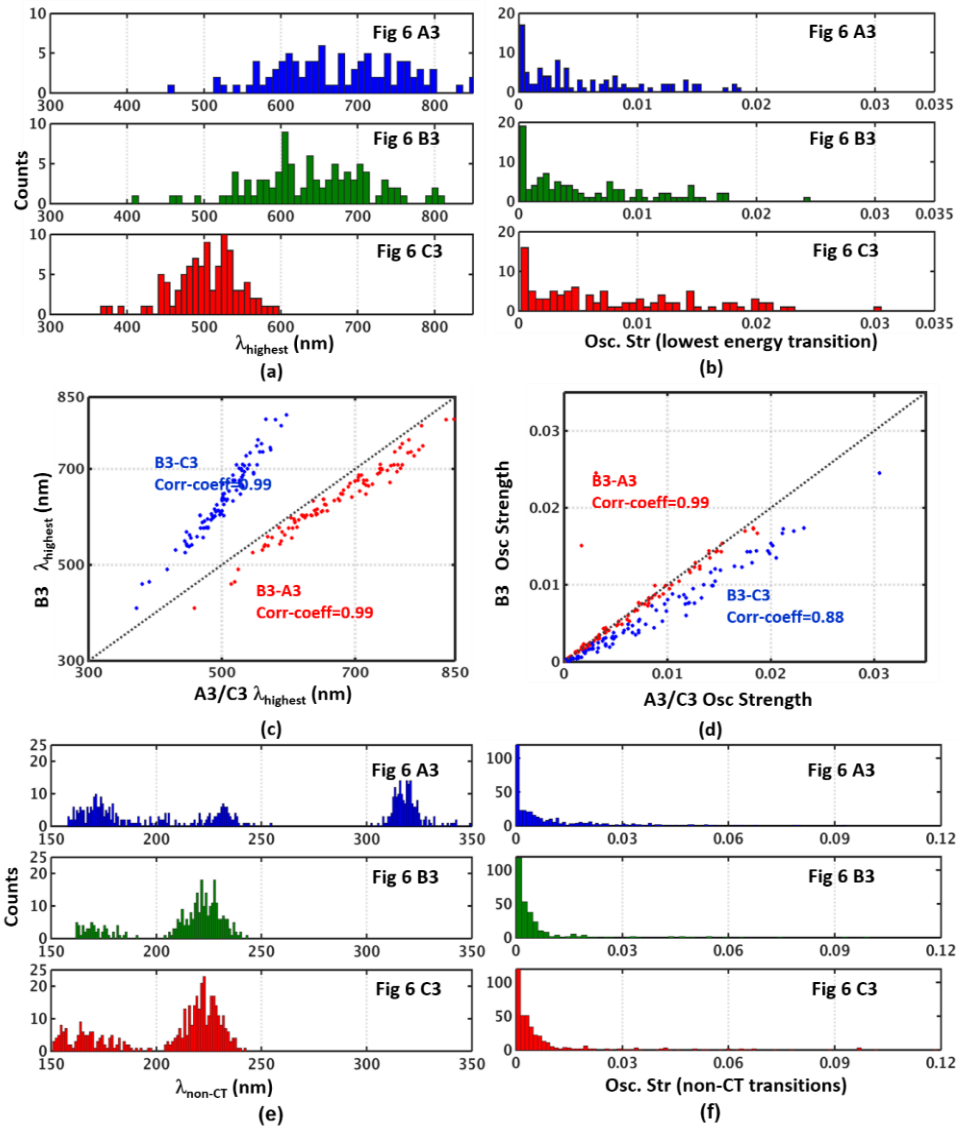

Figure S9: (a) Histograms of absorption wavelengths for the lowest energy transitions obtained in Fig 6 a3-c3. (b) Histograms of oscillator strengths (intensities) for the lowest energy transitions obtained in Fig 6 a3-c3. (c) Scatter plots correlating the absorption wavelengths obtained for the lowest energy transitions in Fig 6 b3 with that in Fig 6 a3 (red) and Fig 6 c3 (blue). (d) Scatter plots correlating the absorption wavelengths obtained for the lowest energy transitions in Fig 6 b3 with that in Fig 6 a3 (red) and Fig 6 c3 (blue). (e) Histograms of absorption wavelengths for LE transitions obtained in Fig 6 a3-c3 (f) histograms of oscillator strengths (intensities) for LE transitions obtained in Fig 6 a3-c3.

TDDFT CAM-B3LYP spectra and statistics for pure CT transitions ($P_{a\to b}(\lambda)$ > 75 %) in Lys-Glu dimers.

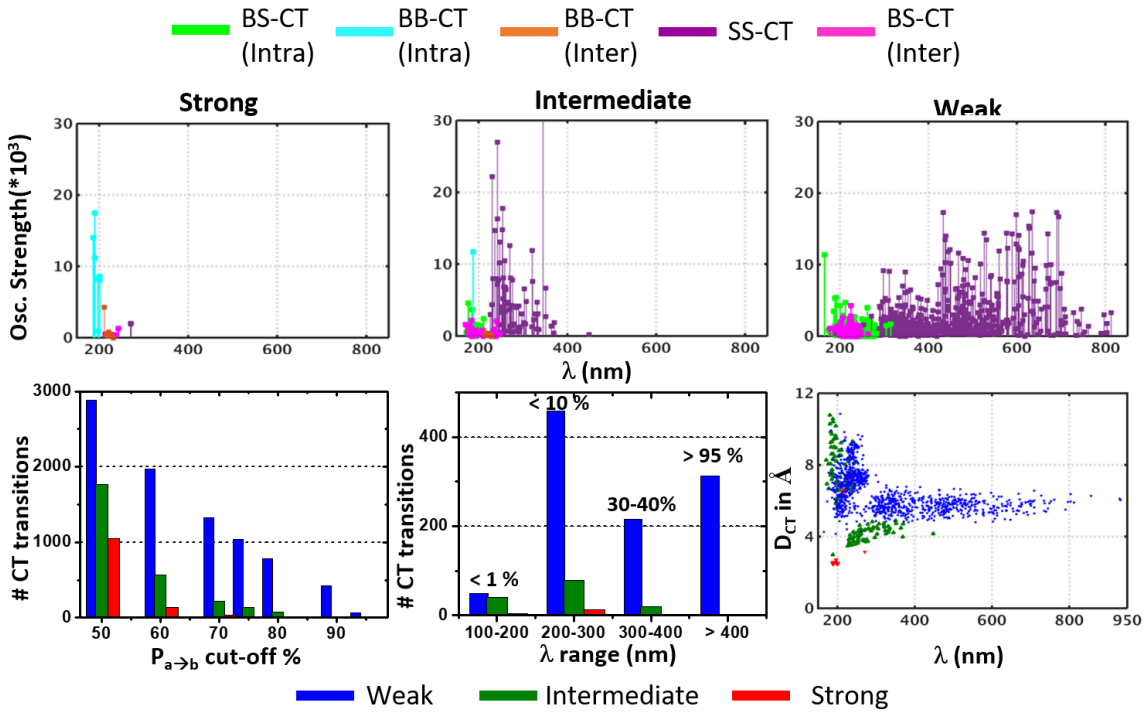

Figure S10: Filtered ProCharTS spectra of pure CT transitions (colored by CT type) computed using TDDFT (CAM-B3LYP/6-31G++(d)) for Lys-Glu dimers with the Symm-backbone model (top row) with strong, intermediate, and weak residue sidechain interactions. The transitions (20,000 in each panel) for the spectra in Fig. 6 b1-b3 were first characterized into different classes of CT transitions using the 4-fragment model as shown in Fig 4 and then selected CT transitions with $P_{a\to b}$ (λ) values > 75 % are displayed in the top row of panels. The bottom row shows statistics associated with pure CT transitions in the Lys-Glu dimer ProCharTS (colored by residue sidechain interaction strength). The total number of CT transitions over the 150-950 nm wavelength range as a function of $P_{a\to b}$ (λ) cut-off is shown in the left panel. Only transitions with $P_{a\to b}$ (λ) values above the cut-off are counted. The middle panel shows number of pure CT transitions observed in specific wavelength ranges corresponding to peptide backbone absorption (100-200 nm), aromatic amino acids (200-300 nm), dye labels (300-400), and protein metal-ligand co-factors/prosthetic groups (> 400 nm). For each wavelength range we also show the fraction of CT transitions as a percentage of total transitions in that window. The right panel shows a correlation plot between the charge separation index D_{CT} (Eqns 2-3) and absorption wavelength for pure CT transitions in the 150-950 nm range.

Statistics for pure CT transitions ($P_{a\rightarrow b}(\lambda)$ > 75 %) in Lys-Glu dimers: CAM-B3LYP versus OT-CAM-B3LYP

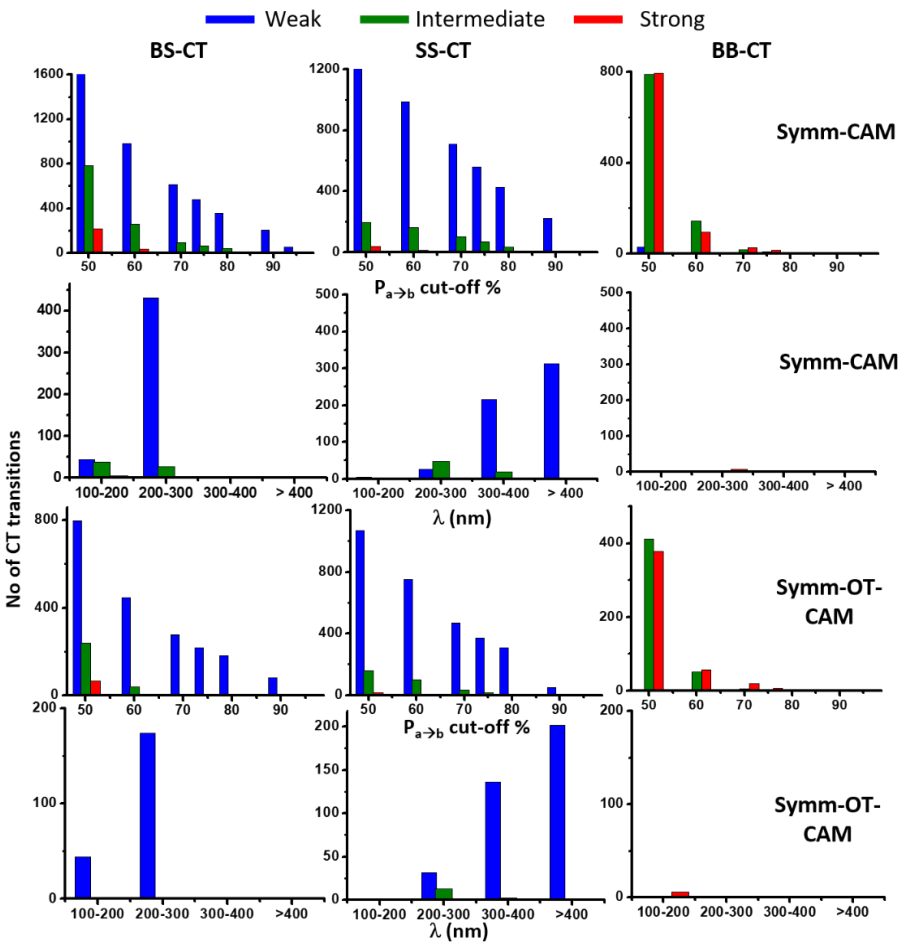

Figure S11: Statistics associated with pure CT transitions in the Lys-Glu dimer ProCharTS (colored by residue sidechain interaction strength). Each column corresponds to data for a specific type of CT transitions (BS-CT, SS-CT, and BB-CT). The total number of CT transitions over the 150-850 nm wavelength range as a function of $P_{a\rightarrow b}$ (λ) cut-off is shown in rows 1 and 3. Only transitions with $P_{a\rightarrow b}$ (λ) values above the cut-off are counted. Rows 2 and 4 show number of pure CT transitions observed in specific wavelength ranges corresponding to peptide backbone absorption (100-200 nm), aromatic amino acids (200-300 nm), labelling dyes (300-400), and protein metal-ligand co-factors/prosthetic groups (> 400 nm). Rows 1 and 2 correspond to data obtained using the Symm-Backbone model and the CAM-B3LYP functional. Rows 3 and 4 correspond to data obtained using the Symm-Backbone model and the OT-CAM-B3LYP functional as indicated.