The "Carbonyl-Lock" Mechanism Underlying Non-Aromatic Fluorescence in Biological Matter

Gonzalo Díaz Mirón[1], Jonathan A. Semelak[1], Luca Grisanti[2], Alex Rodriguez[3], Irene Conti[4], Martina Stella[3], Nicola Seriani[3], Nada Došlić[2], Ivan Rivalta[4,6], Marco Garavelli[4], Dario A. Estrin[1], Gabriele S. Kaminski Schierle[5], Mariano C. González Lebrero[1], Ali Hassanali[3,*], and Uriel N. Morzan[3,*]

[1]Departamento de Química Inorgánica, Analítica y Química Física, Instituto de Química Física de los Materiales, Medio Ambiente y Energía (INQUIMAE), Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, C1428EHA Buenos Aires, Argentina.
[2]Division of Theoretical Physics, Ruder Bošković Institute, 10000 Zagreb, Croatia
[3]Condensed Matter and Statistical Physics, The Abdus Salam International Centre for Theoretical Physics, 34151 Trieste, Italy
[4]Dipartimento di Chimica industriale "Toso Montanari", Università di Bologna, Viale del Risorgimento 4, 40136 Bologna, Italy
[5]Chemical Engineering and Biotechnology, University of Cambridge, CambridgeCB3 0AS, United Kingdom.
[6]ENSL, CNRS, Laboratoire de Chimie UMR 5182, 46 allée d'Italie, 69364 Lyon France
[*]ahassana@ictp.it umorzan@ictp.it

Abstract

Challenging the basis of our chemical intuition, recent experimental evidence reveals the presence of a new type of intrinsic fluorescence in biomolecules that exists even in the absence of aromatic or electronically conjugated chemical compounds. The origin of this phenomenon has remained elusive so far. In the present study, we identify a mechanism underlying this new type of fluorescence in different biological aggregates. By employing non-adiabatic *ab initio* molecular dynamics simulations combined with an unsupervised learning approach, we characterize the typical ultrafast non-radiative relaxation pathways active in non-fluorescent peptides. We show that the key vibrational mode for the non-radiative decay towards the ground state is the carbonyl elongation. Non-aromatic fluorescence appears to emerge from blocking this mode with strong local interactions such as hydrogen bonds. This *carbonyl-lock* mechanism for trapping the excited state leads to the fluorescence yield increase observed experimentally, and paves the way for design principles to realize novel non-invasive biocompatible probes with applications in bioimaging, sensing, and biophotonics.

1 Introduction

The current paradigm in biophysics and photochemistry dictates that the origin of both UV-visible absorption and fluorescence in proteins is mostly associated with the presence of aromatic amino acids[1] or prosthetic external conjugated moieties.[2, 3, 4, 5] Nevertheless, while the electronic absorption spectrum of proteins is traditionally considered to appear in the ultraviolet region (185–320 nm) [1], emerging experimental and computational research has revealed that proteins void of aromatic amino acids or prosthetic groups can absorb beyond 350 nm and fluoresce in the visible range. Such light emission has been reported for protein aggregates like amyloids, monomeric polypeptides or even single amino acids.[6, 7, 8, 9, 10, 11, 12, 13]. This growing body of evidence calls for a re-evaluation of our photochemical fundamentals on what constitutes a fluorophore and which are the chemical mechanisms that lead to this phenomenon.

So far, the current observations indicate that non-aromatic fluorescence is preceded by a near-visible absorption associated to two alternative electronic transitions: (i) the $n \rightarrow \pi^*$ transitions localized in the carbonyl bonds, which can be shifted towards the visible range when local vibrational fluctuations distort the amide plane and elongate the carbonyl bond (CO) distance,[14] and (ii) charge transfer transitions followed by charge recombination have been also identified as a possible source for the UV-vis absorption.[15, 16] In addition, the role of hydrogen bonds (HBs) between CO and peptide NH groups causing electron delocalization and enabling lower transition energies as well as higher radiative relaxation efficiency, was first suggested[17, 18] and confirmed later on.[19, 7, 12]

The fate of non-aromatic molecules on excited states and how they can possibly get trapped leading to emission of visible light, remains an open challenge. We have recently shown, for example, that glutamine amino-acid (L-glu) crystals can be converted through a chemical reaction, in a supramolecular assembly of pyroglutamine molecules.[20] These pyroglutamine molecules are linked together by very strong hydrogen bonds (SHB)[21] which appear to endow them with a longer excited state lifetime, ultimately leading to fluorescence.

One of the key ingredients of enabling excited state lifetime increase is curbing non-radiative decay from the electronic excited to the ground state. These transitions occur through regions of the potential energy surface (PES), commonly referred to as conical intersections (CoIns), where two or more electronic states become degenerate (isoenergetic). In recent years, several theoretical studies from our group have shown that distortions of the amide groups[14, 20, 22] and hydrogen bonding interactions associated with them, may play a key role in inhibiting non-radiative decay. Since CoIns are intrinsically multidimensional in character, the relevant microscopic fluctuations that can compete with the ability of a molecule to fluoresce, and how the mechanisms change across different systems, have remained elusive to date.

In the present work, we provide a unified mechanism explaining the common origin of the non-aromatic fluorescence in $n \rightarrow \pi^*$ and charge transfer transitions in a series of prototypical biological compounds. Inspired by recent experimental and theoretical studies on amyloid-like aggregates and amino acid supramolecular assemblies,[20, 7, 14] we employ five model systems with a different S_1 excited-state nature (see Figure 6) namely, three systems involving charge transfer excitations and another two involving n$\rightarrow \pi*$ transitions. We demonstrate that the key protagonists in the ensuing optical properties are the carbonyl (CO) bonds, whose elongation lead to S_1-S_0 CoIns, enabling the relaxation towards the ground state. We show that an increased excited state lifetime in biological compounds can be achieved by hindering this CO stretching with strong neighboring chemical interactions such as the presence of SHBs. The ubiquitous nature of carbonyl groups in organic systems and the possibility of using them as optical probes has important implications for the interpretation of optical and spectroscopic fingerprints in biological matter, as well as the design of novel probes for bioimaging and sensing applications.

2 Results

The essential ingredients for fluorescence to arise involve a combination of having a long-lived and bright electronic excited state where non-radiative decay mech-

anisms are hindered. Most non-aromatic compounds in biology exhibit ultrafast non-radiative decay that inhibits light emission. Therefore, the first step towards understanding the origin of fluorescence in non-aromatic biological materials, is to characterize the ultrafast non-radiative decay. As a non-fluorescent model system, we employ a dimer consisting of two L-glu molecules with an initial geometry obtained from the crystallographic structure, and an external potential imitating the effect of the surrounding molecules in the crystal (see Methodology section).[20]

Figure 1 illustrates the nature of the $S_1 \rightarrow S_0$ relaxation in L-glu. We performed 200 independent ab-initio non-adiabatic molecular dynamics (AIMD) simulations employing the decoherence-corrected trajectory surface hopping (DC-TSH) scheme, employing the Time Dependent Density Functional Theory (TDDFT) and the PBE0 exchange-correlation functional (see Methodology section). At time t=0, each trajectory was vertically excited to the S_1 state, emulating the initial photoabsorption, afterwards the time evolution of the system is monitored for 250 fs. Following an initial excitation of 4 eV, the ultrafast non-radiative relaxation is evidenced by 98% of the trajectories decaying to the ground state during the simulation time (see Figure S1 in the SI).

Characterizing the specific nuclear motions associated with the $S_1 \rightarrow S_0$ decay and disentangling them from random thermal fluctuations, by visual inspection of MD trajectories or by a brute-force search of relevant degrees of freedom (DoFs), is a daunting task with no guarantee of success: the collective nature of several modes being possibly activated in the excited state prevents a straightforward identification of the relaxation dynamics. Therefore, in order to elucidate the nuclear rearrangements involved in the $S_1 \rightarrow S_0$ decay, we introduce a linear covariance approximation to the non-radiative relaxation mechanism. This approximation combines the nuclear coordinate fluctuations along the MD trajectories and the diabatic energy difference between the S_0 and S_1 states (see $S_0 \rightarrow S_1$ Relaxation Coordinate section). As a result of this procedure, we identify the nuclear fluctuations in the S_1 state that lead to the S_1-S_0 CoIn where the relaxation takes place. The power of our scheme, combining position and energy-fluctuations, is that it reveals automatically

the complex interplay of all the different relevant modes in the non-radiative decay mechanism. It is important to note that, for molecules with more than three atoms, CoIns are multidimensional seams, and hence the possible relaxation pathways are infinite[23]. Therefore, the $S_1 \rightarrow S_0$ decay pathway determined here corresponds to a statistical average of all the accessible decay motions.

Figure 1.**A** shows the main component of the $S_1 \rightarrow S_0$ relaxation pathway projected in the L-glu hydrogen bound dimer model (see also Figure S2 in the SI). The overall collective motion can be decomposed into three main contributions: (i) a concerted event involving an HB weakening along with a contraction of the CO (ii) a planarization of the amide bond (which is deplanarized in the S_1 state), and (iii) a small intermolecular distancing between the hydrogen-bound monomers. The three components of the CoIn pathway are centered around the intermolecular hydrogen bond: the CO contraction reduces the electrostatic interaction between the carbonylic oxygen and the ammonium H, decreasing the HB strength, which causes the intermolecular distancing.

A closer inspection into the de-activating degrees of freedom reveals the essence of the relaxation dynamics along the CoIn: panels **B** and **C** show that accessing the CoIn implies a transient proton transfer (PT) event (the PT coordinate going below 0) significantly increasing in the HB strength. Simultaneously, the amide CO bond stretches, as measured by the CO distance (panel **D**). After this transient HB strengthening and activation of the CO stretching mode associated with the CoIn crossing, the L-glu relaxes to the ground state, where the HB is finally weakened, the CO is contracted and the amide angle is replanarized with respect to the S_1 state conformation (panels **B**, **C** and **D**, see also Figure S3 in the SI). As observed in Figure S4 in the SI, the S_1 state of L-glu has essentially a charge transfer nature, which is stabilized by the subsequent PT. This process, also known as proton-coupled electron transfer (PCET), defines the relaxation process in L-glu, which is also confirmed by our CASPT2 calculations, that show a remarkable agreement in both the nuclear geometry (RMSD $\approx$ 1.7Å) and the electronic structure near the CoIn (see Methodology section and Figure S5 in the SI).

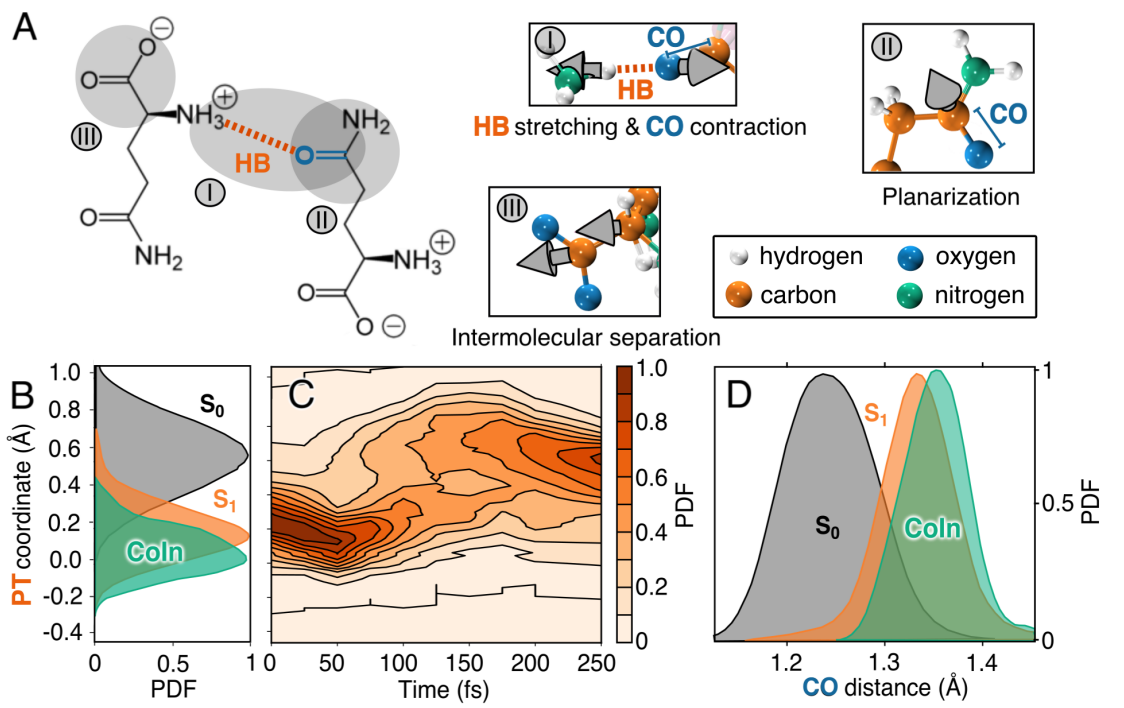

Figure 1: Characterization of L-glu $S_1 \rightarrow S_0$ relaxation pathway. Panel **A**: intermolecular L-glu hydrogen-bonded (HB highlighted in orange) in the zwitterionic state. The $S_1 \rightarrow S_0$ decay coordinate is shown in three panels (I-III): the spheres colored in white, orange, green and blue represent the 3D positions of hydrogen, carbon, nitrogen, and oxygen respectively. The grey arrows illustrate the relaxation coordinate, their length is proportional to the relative contribution of the mode to the $S_1 \rightarrow S_0$ decay. Panel **B**: Proton transfer (PT) coordinate histogram, computed as $d_{\text{O-H}} - d_{\text{N-H}}$, for the S_0 (black) and S_1 states (orange), and the CoIn (green). Panel **C**: Time evolution of the PT coordinate histogram. Panel **D**: CO distance histogram representing the CO distances in the S_0 (black) and the S_1 (orange), as well as in the CoIn (green).

Having characterized the ultrafast $S_1 \rightarrow S_0$ relaxation of a prototypical non-fluorescent compound such as L-glu, the next step towards understanding the mechanism behind non-aromatic fluorescence is to analyze the different possible ways to increase the excited state lifetime. Panels **A**-**D** in Figure 2 show that this can be achieved by constraining independently any of the DoFs associated with the relaxation dynamics. Indeed, by inhibiting the different components of the decay pathway with an external harmonic constraint (see methods section) the access to the CoIn can be blocked, artificially trapping the L-glu in the S_1 state. The four panels show in black the S_1-S_0 energy gap for a selected trajectory that decays to the ground state at $\approx$ 160 fs (when the S_1-S_0 energy gap vanishes). In contrast, when the CO, the HB, or the amide plane DoFs are constrained, the $S_1 \rightarrow S_0$ relaxation is

impeded, as shown by the orange curves in panels **B**, **C** and **D** respectively. It is worth noting that the intermolecular distancing mode depicted in Figure 1.**A**.III was not tested here since it involves a rather large perturbation on all the DoFs of the dimer resulting in a trivial trapping of the excited state. Additionally, we observe that constraining DoFs other than those identified by our covariance approach does not lead to a significant reduction of the relaxation time, even if the atoms involved are adjacent to the amide group (see Figure S6 in the SI).This evidence not only serves as a validation for our covariance decay pathway approximation introduced above, but more importantly it sets the design rules for the development of novel materials with increased excitonic lifetimes.

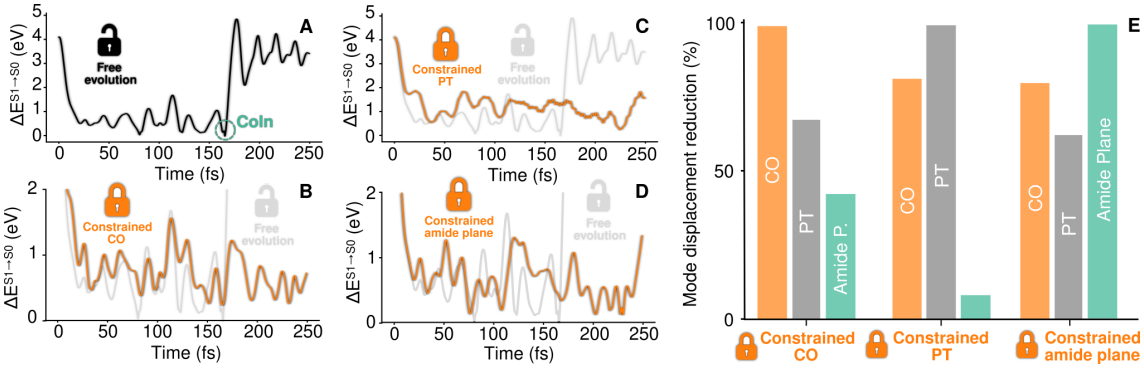

Figure 2: The $S_1 \rightarrow S_0$ relaxation can be delayed by stiffening the decay pathway modes. Panel **A** shows the S_1-S_0 energy difference for a selected L-glu trajectory evolving freely without constraints. Panels **B**, **C** and **D** show the S_1-S_0 energy difference for a selected L-glu trajectory evolving under an applied harmonic constraint (orange curve) on the CO distance (panel **B**), the PT coordinate computed as $d_{\text{O-H}} - d_{\text{N-H}}$ (panel **C**) or the amide bond plane (panel **D**). Panel **E** shows a bar chart indicating the percentual reduction of the displacements in the PT coordinate, CO distance and amide plane degrees of freedom as a result of the constraints applied in panels **B**, **C** and **D** respectively (see methods section). Upon application of the constraints labeled in the horizontal axis, the vertical bars show the percentual reduction in the CO (orange bar), PT coordinate (grey bar) and amide plane (green bar) DoF displacement.

Panels **B-D** confirm that the CO, PT or the amide plane DoFs have a role in the relaxation process. But, are these three DoFs equally important for the $S_1 \rightarrow S_0$ decay? In order to dissect the individual role played by each DoF in the CoIn crossing pathway, panel **E** shows a bar chart quantifying how the constraint in a given DoF affects the displacements in each of the three DoFs (for further details on this estimation see methodology section and Figure S7 in the SI). This

enables establishing a hierarchical ordering between the different DoFs: when the HB distance is constrained, the amide plane DoF remains almost unperturbed, which indicates that the relaxation process can be hindered without altering significantly the natural dynamics of the amide plane. Therefore, the amide planarization by itself is not enough for the relaxation process to take place. Conversely, the CO bond dynamics is the most affected by the three different S_1-trapping constraints, indicating that it is at the core of the relaxation pathway. Similarly, but to a lesser extent, the HB displacement is moderately affected by all the three constraints, showing that it is also a critical fluctuation for the decay process.

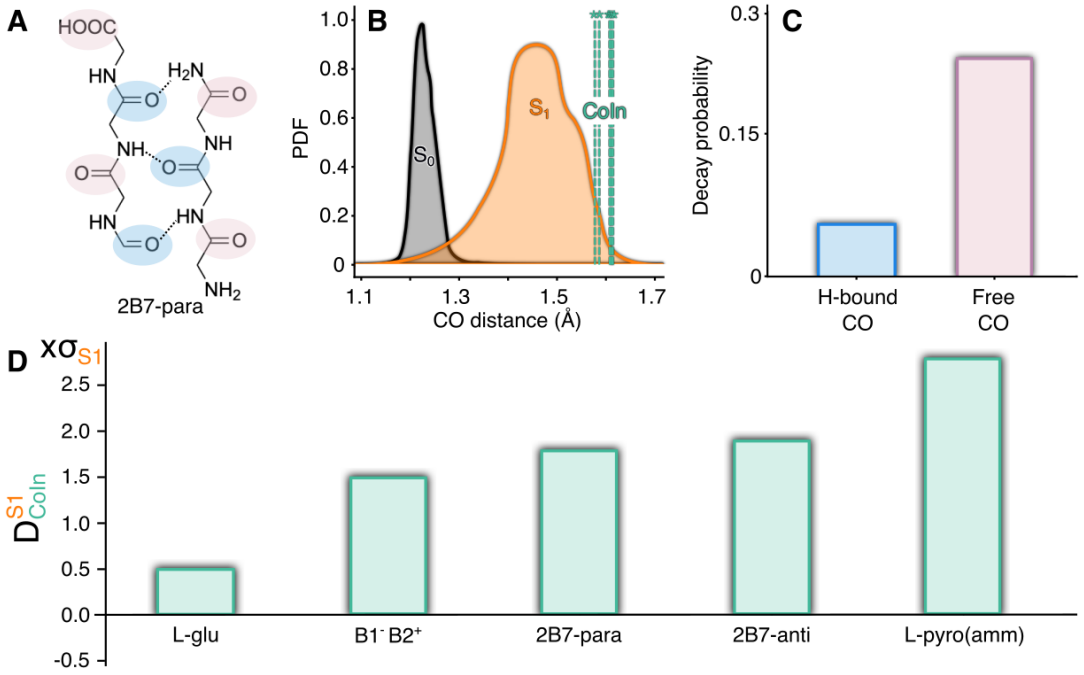

Figure 3: The CO stretching mode role in the $S_1 \rightarrow S_0$ relaxation is ubiquitous among a diversity of prototypical peptide structures. Panel **A** depicts the 2B7 amyloid model composed of 2-strand anti-parallel configuration adopted in the $\beta-$sheet arrangement. The H-bonded and the *free* COs are colored in light blue and pink, respectively. Panel **B** shows the CO distance distribution in the S_1 state (orange), ground state (black) and the distances associated with the CoIn crossings (green) are marked with vertical dashed lines and an asterisk. Panel **C** shows a bar chart indicating the proportion of $S_1 \rightarrow S_0$ relaxation events occurring through H-bonded CO elongations (light blue bar) or non-H-bonded (free) CO elongations (pink bar). Panel **D** shows a bar chart quantifying the difference between the average CO distance, D^{S1}_{CoIn}, between the CO length distributions in the S_1 state and in the CoIn crossings. This distance is computed as the difference between the means of the two distributions normalized by the standard deviation of the excited state CO length.

Analogously, the S_1-S_0 CoIn crossing dynamics in the small model system in-

spired from the amyloid sequence Aβ_{30-35},[7] and dubbed 2B7-para (see Figure 3), involves a strong CO elongation. Amyloids are self-assembled polypeptides characterized by a highly ordered cross β arrangement, where the $\beta-$strands are interconnected by HBs. Because of their involvement in a wide range of human diseases, amyloids have been the focus of attention for numerous experimental and theoretical studies.[12, 14] 2B7-para is characterized by a 2-strand-parallel $\beta-$sheet arrangement,[14] and two types of carbonyl groups can be distinguished in its structure: those that are H-bonded with the NH group of the neighbor strand, and those that are not bonded or *free* (Figure 3.**A**). At variance with the case of L-glu, the nature of the S_1 state in 2B7 is $n\pi^*$ localized on the carbonyl groups.[14] Noteworthy, the CO elongation that leads to the CoIn crossing molecular configuration is approximately 4 times more likely to occur in a free CO than in the H-bonded ones. This prevalence of the free CO deactivations indicates that the HBs can hinder the CO elongation, inhibiting the $S_1 \rightarrow S_0$ relaxation. Furthermore, the decay events localized in the non-bonded COs show that the CO stretching by itself can act as a stand-alone relaxation pathway, without coupling to an HB mode. Thus, the precise mechanisms involved in non-aromatic fluorophores will naturally be fine-tuned by the chemical details of the HB networks involved at the site of photochemical activity.

In order to assess the ubiquitous role of the CO-elongation in the non-radiative relaxation dynamics of non-aromatic fluorescent systems, Figure 3, panel **D** compares the characteristic CO elongation associated with the S_1-S_0 CoIn crossing of 5 representative model systems (see also Figure 6 for detailed molecular structures). Among them, L-glu is the only non-fluorescent model. The remaining model systems are some of the most relevant non-aromatic fluorescent cases that have been reported so far.[12] The fluorescent counterpart of L-glu, L-pyro(amm), shares the same charge transfer S_1 character and a very similar crystal intermolecular arrangement. In addition, three models associated with the non-aromatic fluorescent 2Y3J (amyloid segment) peptide are employed: (i) 2B7-para, (ii) 2B7-anti, which is the antiparallel analog of 2B7-para, and (iii) $B1^-$ $B2^+$, which represents two zwitterionic

H-bonded termini residues of 2Y3J. The S_1 character of (i) and (ii) is $n\pi^*$, while that of (iii) has a charge transfer nature. The vertical axis in the bar chart quantifies the distance between the CO length distributions in the S_1 state and in the CoIn crossing (D^{S1}_{CoIn}). This distance is an estimation of how rare the CoIn crossing event is for the S_1 state dynamics. Loosely speaking, it indirectly provides a measure of the magnitude of fluctuation needed along the CO in the S_1 state in order to elongate it up to a S_1−S_0 crossing point. The requirement of a CO elongation in the excited state as a precondition for the S_1 →S_0 relaxation is verified in the five different model systems, suggesting that it is a general fingerprint for the S_1 →S_0 non-radiative decay in polypeptides.

At the two extremes of Figure 3.D L-glu and L-pyro(amm) show the lowest and the highest D^{S1}_{CoIn} value respectively. Figure S8 in the SI shows that the fluorescence in L-pyro(amm), as compared to its precursor L-glu, does not arise from an increase in the S_1 →S_0 transition dipole moment (i.e. the instantaneous emission probability), which indicates that destabilization of the non-radiative decay pathways is the main origin of its fluorescence.

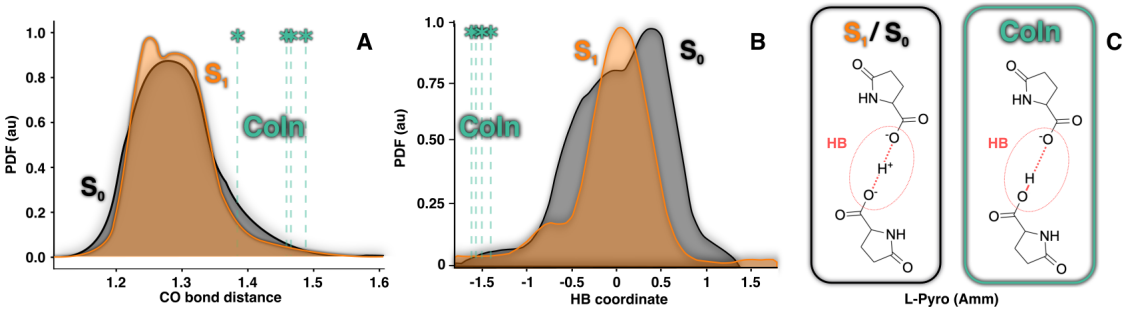

Figure 4: The excited state lifetime of L-pyro(amm) increases by "locking" the CO stretch with a strong HB. Panels **A** and **B** show the distribution of CO distance and PT coordinate values, defined as $d_{O\text{-}H} - d_{O'\text{-}H}$ (where O and O' identify the two carboxyl oxygens involved in the HB), in the S_1 state (orange), ground state (black). The CoIn configurations are represented in green vertical bars (only 2.5 % of the AIMD trajectories decay to the ground state). Panel **C** depicts the molecular arrangement of the L-pyro(amm) dimer model system in the S_0 and S_1 states (black framed panel), and the transient proton transfer arrangement associated to the CoIn crossing configuration (green framed panel).

A closer inspection into the L-pyro(amm) dynamics (Figure 4), indicates that the HB interaction plays a crucial role in hindering the access to the S_1−S_0 CoIn, and hence in the ensuing fluorescence. The main structural difference between L-glu and

L-pyro(amm) is the presence of a very strong HB between the carboxyl groups (with a length of ≈2.45 Å) in the latter, while L-glu presents a more conventional HB (with a length of ≈2.85 Å). As in the case of L-glu, the S_1 excited state of the crystalline L-pyro(amm) is characterized by a charge transfer transition between the H-bonded residues (see Figure S4 in the SI). Only 2.5% of L-pyro(amm) NAMD trajectories decay to the ground state within 250 fs (see Figure S1 in the SI),showing that its excited state lifetime is considerably increased with respect to L-glu. Panels **A** and **B** provide a clear explanation for this: at variance to the case of L-glu, both the HB coordinate and the CO distance in L-pyro(amm) are not considerably altered upon $S_0 \rightarrow S_1$ excitation. The S_1 structural arrangement remains very similar to that in the S_0 state. This hampers the access to the CoIn crossing conformations which, as in L-glu, have a PCET nature that requires further CO elongation coupled with a proton donation (panel **C**). Therefore, the S_1 lifetime in L-pyro(amm) is enhanced with respect to that of L-glu by destabilizing the $S_1 \rightarrow S_0$ relaxation pathway.

3 Discussion

Two alternative types of electronic transitions have been identified behind the non-aromatic fluorescence phenomenon: n$\rightarrow \pi^*$ and charge transfer.[12] In both cases the absorption and fluorescence occur in the near-UV to visible range. In the present study we show that the CO stretching is the key molecular distortion occurring during the $S_1 \rightarrow S_0$ relaxation in both types of excitations, by means of non-adiabatic excited state MD simulations of a series of prototypical peptide systems. Importantly, this is equally valid when the S_1 state character is either n$\rightarrow \pi^*$ (2B7-para and -anti case) or charge transfer (L-pyro(amm), L-glu, and $B1^-B2^+$ case), suggesting that a common decay pathway could be ubiquitous among proteins and peptide aggregates.

Furthermore, we have shown that strong local interactions, such as SHBs, can prevent the CO from stretching, hindering the relaxation towards the ground state, and hence increasing the S_1 excited state lifetime. We propose this "CO-locking" mechanism as the origin of the non-aromatic fluorescence. Importantly, the ubiq-

uity of SHBs in biological and inorganic matter[24] added to recent experimental evidence revealing fluorescence in carbonyl containing compounds,[25, 11] suggest that this "CO-locking" fluorescence mechanism might be widespread among biology and beyond.

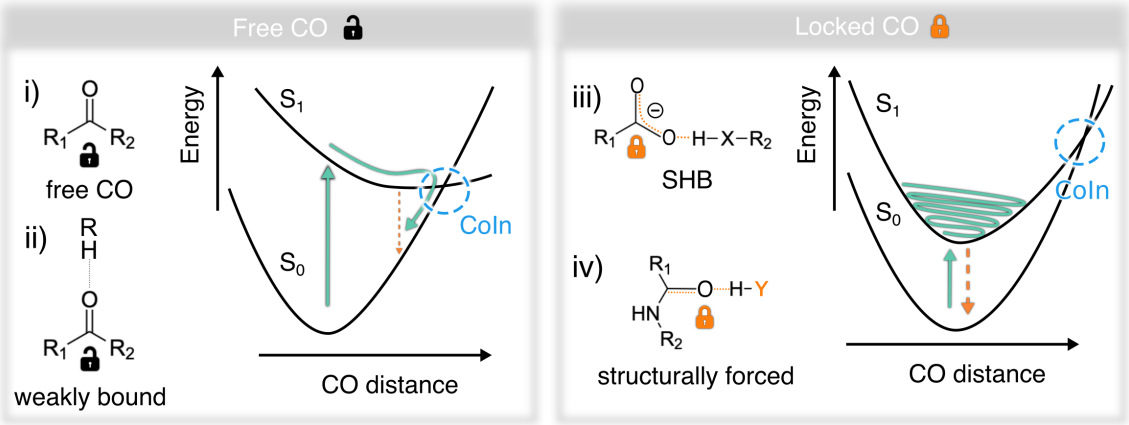

Figure 5: The CO "lock" mechanism of non-aromatic fluorescence. The left panel shows the typical relaxation pathway triggered by a CO elongation in two possible scenarios where the CO is non-bound (case i) or weekly bound (case ii) or non-bound (case ii). The right panel shows the CO-lock mechanism in which a strong local interaction blocks the large amplitude CO elongations preventing the relaxation towards the ground state. Again, two alternative CO-locking scenarios are presented: in case iii an electron withdrawal HB donor strongly interacts with the carboxyl group, limiting the resonance of the double bond and preventing its elongation. In case iv an HB is established with the CO group imposing a direct restriction to large elongations.

Figure 5 shows a schematic representation of the CO-locking mechanism. The left panel represents the non-radiative decay mechanism based on the elongation of a free (case i) or weakly bound (case ii) CO bond. The right panel shows two alternative scenarios for hindering the CO-relaxation. In both cases the HB stiffens the carbonyl vibrations, impeding its stretching. Case iii depicts the trapped charge transfer S_1 state exemplified in the previous section with the case of L-pyro(amm). Here the electron withdrawal effect of the SHB limits the internal double bond resonance in the carboxyl group, inhibiting the double-bond elongation. The relaxation coordinate in this case consists essentially in a PCET where the electron-hole recombination is preceded by a concerted proton transfer and CO elongation.

The second scenario depicted in the right panel corresponds to the 2B7 amyloid model, where the HB is directly established with the CO group and the nature of

the S_1 excited state is n$\rightarrow \pi^*$. Here, the CO and NH groups belong to opposite β-strands in the amyloid-like structure, hence the CO elongation forces an energetically unfavorable inter-strand separation, limiting the access to the non-radiative relaxation.

The generality of this mechanism is reinforced by employing a combination of *ab initio* non-adiabatic molecular dynamics simulations at multiple levels of theory: going from TDDFT, to ADC(2), to CASPT2. We have combined these techniques with a simple data-driven approach pinpointing the key structural and dynamical aspects of the $S_1 \rightarrow S_0$ decay, and enabling the recognition of the CO elongation as the essential nuclear fluctuation associated with the CoIn configurations that lead to the non-radiative relaxation.

Similar aggregation-induced emission (AIE) has been reported in the context of aromatic synthetic molecular materials, where it was hypothesized that steric restrictions in the aggregate phase induce the emissive response.[26, 27] The "CO-locking" mechanism identified here can be considered as a particular form of AIE, with the singularity that it enables the fluorescence of non-aromatic naturally occurring biomolecular aggregates. Whether this mechanism could be operative in a broader variety of synthetic and aromatic molecular or polymeric materials reminds an open question that will be the focus of our future research.

Traditionally, autofluorescence has been primarily considered a source of noise for fluorescence imaging methods, arising from different aromatic biomolecules.[28, 29] However, very often these signals show absorption and emission in the spectral range of non-aromatic fluorophores.[12, 30] In this context, our findings offer a new interpretation for the fluorescence fingerprints of complex biological systems, paving the way for the development of non-invasive measurements monitoring structure and conformational dynamics of proteins inside living cells. Furthermore, the interactions originating non-aromatic fluorescence can be linked to specific secondary structural arrangements, for instance, it has been shown that the characteristic fluorescence and absorption features in amyloid proteins emerges at the beginning of the $\beta-$sheet formation process.[31] This fold-sensitive biooptical effect could be har-

nessed to design label-free medical diagnosis technology and precise phototherapy treatment protocols for amyloidosis at its early stages.[8]

Overall, the findings reported in this work lay down the ground principles behind non-aromatic fluorescence in biological materials. This simple molecular-level picture might enable the rational design of a new generation of bioinspired peptide integrated optical devices with unique photonic and electronic properties and an intrinsic biocompatibility. Specifically, the development of biological aggregates with a supramolecular arrangement imposing a constraint on the carbonyl stretching modes could lead to bright materials with adjustable optical properties.

4 Acknowledgement

GDM and JAS gratefully acknowledges CONICET for the fellowship. UNM acknowledges CINECA supercomputing centre (IsC85 project), and Elettra-TeraFERMI project 20224056. AAH would like to acknowledge the European Commission for funding on the ERC Grant HyBOP 101043272. DAE, JAS and UNM would like to acknowledge founding from PICT 2020 01828, Agencia I-+d+i. IR gratefully acknowledges the use of HPC resources of the "Pôle Scientifique de Modélisation Numérique" (PSMN) of the ENS-Lyon, France. We also acknowledge Prof. Sir John Walker, Prof. David Palmer, Dr. Johannes Schmidt, Dr. Zeinab Ebrahimpour, Dr. Pablo Videla, Prof. Victor S. Batista and Dr. Marcello Coreno for useful discussions.

5 Methodology

Five different prototypical model systems were employed throughout this work (Figure 6). Among them, two different types of $S_0 \rightarrow S_1$ transitions characterize these compounds: n$\rightarrow \pi^*$ or charge transfer (left and right panel respectively).

Three of the simulated systems correspond amiloid-based structures: 2B7(-para and -anti) were obtained starting from available crystal structure of hexapeptide Aβ_{30-35}[32] (PDB code: 2Y3J) replacing all side-chains with H, removing 1-C and

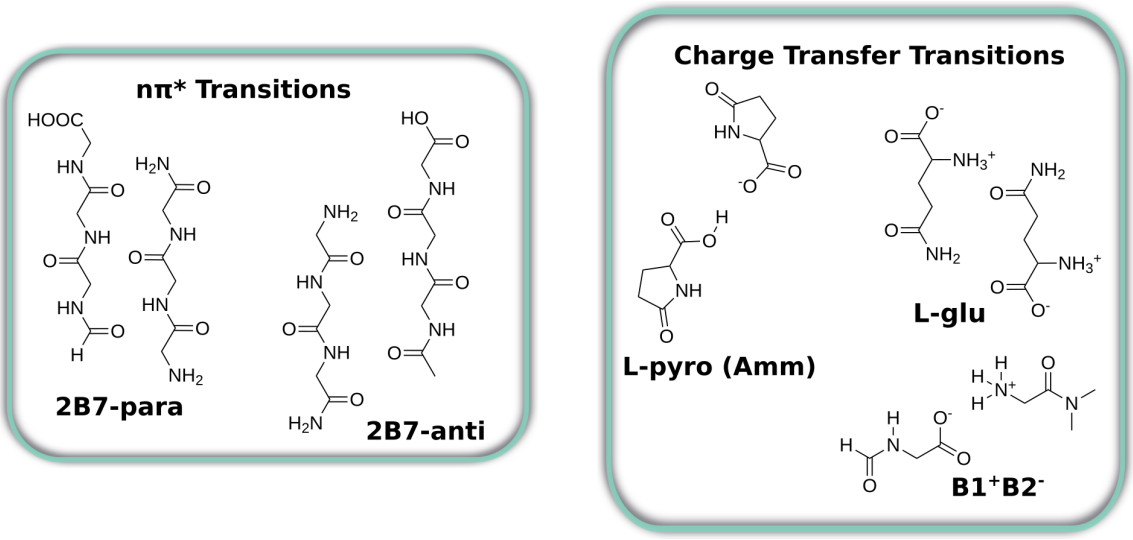

Figure 6: Molecular structures for the five model systems studied in this work separated by the electronic transition character.

1-N termini from each chain, and capping respectively with -CHO and -$CONH_2$. The geometry of two chains were optimized by either keeping parallel configuration 2B7-para (as in the crystal) or by preparing it in antiparallel fashion 2B7-anti (more details about these systems can be found in reference[14]). $B1^+B2^-$ model was prepared starting from the same crystal structure, but selecting only the head-to-tail termini-fragments (C-termini capped with CHO and N-termini capped with $N(CH_3)_2$). In the optimization and the ground state trajectories for $B1^+B2^-$ the proton was constrained to remain on the N-side.

For L-glutamine (L-glu) and L-pyro-glutamine-ammonium (L-pyro(amm)) we extracted a dimer conformation from the crystal structures (see Figure S9 in the SI), taking care that they accurately reproduce the optical properties of the solid.[20] In order to preserve the molecular arrangement of the crystal structure we applied soft harmonic constraining potentials (with a constant of 100 $Kcal/mol$Å) to the atoms indicated in Figure S9 in the SI. This procedure does not suppress any vibrational motion, but rather approximates the steric effect of neighboring molecules in the crystal structure. In the case of L-pyro(amm), our model does not explicitly include the ammonium ion, but due to the presence of the soft constraining potentials described above the system retains the same molecular arrangement as in the presence of ammonium.

5.1 Non-Adiabatic Molecular Dynamics: Trajectory Surface Hopping

In order to examine the excited state dynamics and relaxation of the systems described in the previous sections, we employed the trajectory surface-hopping (TSH) approach[33, 34, 35] using two different electronic theory levels: TDDFT for L-glu and L-pyro and ADC(2) for amiloid-like systems (see next section).

TSH employs a swarm of independent classical trajectories, each one evolving on a single potential energy surface (PES). In every MD step, the hop probability to other PES is computed according to the following expression:

$$P_{i\to j} = -2\int_{t}^{t+dt} \frac{c_i^* c_j \dot{R}\,\overrightarrow{d}_{ij}}{|c_i|^2}. \tag{1}$$

The coefficients of each electronic state c_i evolve according to the Time Dependent Schrödinger Equation (TDSE):

$$i\frac{dc_j(t)}{dt} = c_j(t)E_j - i\sum_i c_i(t)\dot{R}\,\overrightarrow{d}_{ij}, \tag{2}$$

where $\dot{R}$ denotes the time derivative of nuclear coordinates R. In contrast with the electrons, the nuclei are propagated using classical mechanics following Newton's law:

$$M\frac{\partial^2}{\partial t^2}R = -\nabla_R E_j \tag{3}$$

The term $\overrightarrow{d}_{ij}$ in equations 1 and 2 is the key variable in NAMD called Non-Adiabatic Coupling Vector (NACV).

$$\overrightarrow{d}_{ij} = \langle\Psi_i|\frac{\partial}{\partial\overrightarrow{R}}|\Psi_j\rangle \tag{4}$$

Ψ_i is the Born-Oppenheimer (BO) wavefunction of PES i. Whether the system change its PES or not is controlled by stochastic decision algorithm[36, 37].
We included a decoherence correction (DC) developed by Granucci et al[38], where the electronic coefficients c_j are damped with the following equations:

$$c_j(t) = c_j(t)e^{-\Delta t/\tau_{ji}}, \tag{5}$$

where the decoherence time τ is defined by:

$$\tau_{ji} = \frac{\hbar}{|E_j - E_i|}\left(1 + \frac{C}{E_{kin}}\right). \tag{6}$$

The state i denotes the actual PES, C is an adjustable parameter (0.1 in this work) and E_{kin} is the nuclear kinetic energy.

5.2 NAMD Simulation Protocol

The initial conditions for our L-glu and L-pyro(amm) NAMD simulations were generated by extracting a dimer conformation from a ground state optimized structure, followed by 200 ns of classical MD in the NVT ensemble at 300 K with a 1 fs time-step, employing the *AMBER* package[39]. From this trajectory, 200 nuclear conformations were employed as initial configurations for a 1 ps *ab initio* ground state MD simulation in the NVT ensemble at 300 K with a time step of 0.5 fs. The calculations were performed using TDDFT at the PBE0/6-31G[40] level as implemented in the **LI**$\hat{\mathcal{O}}$ package by our group.[41, 42] The nuclei were evolved classically by employing *AMBER*.[39] After the initial sampling on the ground state, the system was vertically excited onto the S_1 electronic state, and then evolved for 250 fs, employing the TSH scheme with a timestep of 0.5 fs in NVE ensemble. Excitation energies and oscillator strengths were calculated using LR-TDDFT[43] and the Tamm-Dancoff approximation[44]. The NACVs were calculated at the same theory level using the method developed by Furche et al[45, 46]. The functional and basis set used for the NAMD trajectories were the same as those employed in ground state dynamics. A total of 200 NAMD trajectories were performed L-glu system and 100 NAMD trajectories for L-pyro.

The amyloid model systems (Figure 6 left panel) were prepared by first performing ground state *ab initio* molecular dynamics (AIMD) simulations employung the CP2K package[47]. A convergence criterion of $5\times\ 10^{-7}$ a.u. was used for the op-

timization of the wave function. Using the Gaussian and plane wave methods, the wave function was expanded in the Gaussian double-ζ valence polarized (DZVP) basis set, and the Becke-Lee-Yang-Parr (BLYP)[48, 49] functional with the D3(0) Grimme dispersion corrections for van der Waals interactions[50]. TSH dynamics for 2B7(-para and -anti) and $B1^{+}B2^{-}$ was performed according to the scheme presented before, employing an in-house version of the Zagreb surface hopping code[51] based on the fewest switches surface hopping algorithm[33] at the the ADC(2) level of theory. The initial conditions (positions and velocities) were prepared by randomly selecting frames from GS AIMD and by vertically exciting to the S_1-S_5 manifold. In the case of 2B7, a total of 21 trajectories were obtained, 15 with parallel and 6 with anti-parallel β-strand configuration. For $B1^{+}B2^{-}$ 22 trajectories were performed. These simulations were obtained for 250 fs or until the energy gap between the S_1 and S_0 states dropped below 0.1 eV. More information can be found in the Supplementary material of reference [14].

5.3 Validation of TDDFT results at the CASPT2 level

Describing the electronic structure and nuclear configurations around CoIns represents a challenging task for most electronic structure methods. One of the most robust and accurate approaches for the characterisation of CoIns is the CASPT2 method.[52, 53] In this study, in order to validate our TDDFT-based simulations near the S_1-S_0 crossing regions we performed SS- and MS-CASPT2/6-31G* calculations averaging over three states (i.e. SA-3). The active space included the orbitals mainly involved in the photoreactive molecular region (see Figure S4 in the SI), corresponding to 10 electrons in 8 orbitals in the case of L-glu and 14 electrons in 10 orbitals in the case of L-pyro(amm). CoIn optimizations were performed with numerical gradients at the MS-CASPT2 level, utilizing the gradient projection algorithm of Bearpark et al.[54, 55] as implemented in COBRAMM.[56, 57] The active space of the CoIn optimizations included 8 electrons in 7 orbitals for both the systems. Thereby, the ionization-potential-electron-affinity (IPEA) shift[58] was set to 0.0, and an imaginary shift[59] of 0.2 a.u. was used throughout. All CASPT2 calcula-

tions were performed using the Gaussian16 code[60] and OpenMolcas package[61, 62] through its interface with COBRAMM. The validation of our TDDFT results in the proximities of the CoIn was performed by comparing both the structures of the MS-CASPT2 optimized CoIns and those obtained from NAMD at the TDDFT level (see the previous method sections), as well as the symmetry and orbitals involved in the $S_0 \leftrightarrow S_1$ electronic transition.

5.4 $S_1 \rightarrow S_0$ Relaxation Coordinate

In order to elucidate the nuclear rearrangements involved in the $S_1 \rightarrow S_0$ relaxation, we employed a modified principal component analysis (PCA) combining the nuclear coordinate fluctuations $(x - \overline{x})$ along the AIMD trajectories and the energy difference between the diabatic electronic states (ΔE^D). If N is the number of atoms in the system, we define the $3N-$dimensional $S_1 \rightarrow S_0$ relaxation pathway vector $\mathbf{c}$ as:

$$\mathbf{c_i} = \frac{\langle [x_i(t) - \overline{x_i}] \mathrm{Sign}[-\Delta E^D(t)] \exp^{-\frac{|\Delta E^D(t)|}{\alpha k T}} \rangle}{\sqrt{\langle [x_i(t) - \overline{x_i}]^2 \rangle \langle [\exp^{-\frac{|\Delta E^D(t)|}{\alpha k T}}]^2 \rangle}}, \tag{7}$$

where the index i spans over the $3N$ Cartesian coordinates of the system, $x_i(t)$ represents the i-th component of the cartesian position vector at time t, ΔE^D is the diabatic energy difference between S_0 and S_1, and the angular brackets as well as the over-bar represent a time average. The diabatic energies were approximated by the adiabatic ones before the CoIn crossing and swapping the S_1 and S_0 identities after the CoIn passage. The second term in the numerator, $\mathrm{Sign}[-\Delta E^D(t)]$, sets the direction of the relaxation pathway vector from the S_1 to the S_0 configurations. The third term, $\exp^{-\frac{|\Delta E^D(t)|}{\alpha k T}}$, is an Arrhenius-like factor, where T is the room temperature, k is the Boltzmann constant, and α is an adjustable parameter that controls the width of the exponential term with respect to $\Delta E^D(t)$ (throughout this work α was fixed to 100, see Figure S10 in the SI). This factor enables disentangling the thermal fluctuations from the relaxation process. As the nuclear configurations get closer to the CoIn, the $|\Delta E^D(t)|$ tends to zero and the Arrhenius-like term peaks for these configurations, increasing their weight in the ensemble average. In this way, the vector $\mathbf{c}$ is a linear estimator of the nuclear fluctuations in the S_1 state that lead

to the S_1-S_0 crossing. It is important to note that usually the excited state landscape is characterized by many accessible CoIns, and several different decay pathways can be accessible. In these cases, the relaxation pathway vector **c** represents a statistical average of all the decay motions.

5.5 Hierarchical Ordering of the $S_1 \rightarrow S_0$ Relaxation Degrees of Freedom

The bar chart in Figure 2 estimates the extent in which the dynamics of each degree of freedom (DoF) is influenced by a constrain in a chosen DoF. This magnitude is quantified as the normalized variance $\tilde{\text{Var}}_{ij}(Y) = 1 - [\text{Var}(Y_i^j) - \text{Var}(Y_i^0)]/\text{Var}(Y_i^0)$, where $\text{Var}(Y_i^j)$ is the variance of the displacements in the $i-$th DoF obtained from a NAMD simulation where the $j-$th DoF is being constrained, both i and j indexes label the CO, HB and amide plane modes, and the 0 index refers to the unconstrained simulation (see Figure S7 in the SI). The values of $\tilde{\text{Var}}_{ij}(Y)$ indicate how much the dynamics in the i-th DoF is affected by the constraint in the $j-$th DoF.

References

[1] Lakowicz, J. R. Principles of Fluorescence Spectroscopy. *Springer* **2006**

[2] Teale, F. W. J. The ultraviolet fluorescence of proteins in neutral solution. *Biochem J.* **1960**, *76*, 381–388

[3] MR., E. The use of fluorescence methods to monitor unfolding transitions in proteins. *Biochem J.* **1994**, *66*, 482–501

[4] Lapidus, L. J.; Eaton, W. A.; Hofrichter, J. Measuring the rate of intramolecular contact formation in polypeptides. *Proceedings of the National Academy of Sciences* **2000**, *97*, 7220–7225

[5] Sizemore SM, R. A. G. G. V. S., Cope SM Slow Internal Dynamics and Charge Expansion in the Disordered Protein CGRP: A Comparison with Amylin. *Biophys J.* **2015**, *109*, 1038–1048

[6] Pinotsi, D.; Buell, A. K.; Dobson, C. M.; Kaminski Schierle, G. S.; Kaminski, C. F. A Label-Free, Quantitative Assay of Amyloid Fibril Growth Based on Intrinsic Fluorescence. *ChemBioChem* **2013**, *14*, 846–850

[7] Pinotsi, D.; Grisanti, L.; Mahou, P.; Gebauer, R.; Kaminski, C. F.; Hassanali, A.; Kaminski Schierle, G. S. Proton Transfer and Structure-Specific Fluorescence in Hydrogen Bond-Rich Protein Structures. *Journal of the American Chemical Society* **2016**, *138*, 3046–3057, PMID: 26824778

[8] Pansieri, J. et al. Ultraviolet–visible–near-infrared optical properties of amyloid fibrils shed light on amyloidogenesis. *Nature Photonics* **2019**, *13*, 473–479

[9] Johansson, P. K.; Koelsch, P. Label-free imaging of amyloids using their intrinsic linear and nonlinear optical properties. *Biomed. Opt. Express* **2017**, *8*, 743–756

[10] TN, T.; NR, R.; AY, Z.; NN, S.; AA, R.; AS, O.; EN, N.; VV, F.; VN, U.; EA., S. Dissection of the deep-blue autofluorescence changes accompanying amyloid fibrillation. *Arch Biochem Biophys.* **2018**, *651*, 13–20

[11] Niyangoda, C.; Miti, T.; Breydo, L.; Uversky, V.; Muschol, M. Carbonyl-based blue autofluorescence of proteins and amino acids. *PLOS ONE* **2017**, *12*, 1–15

[12] Morzan, U. N.; Díaz Mirón, G.; Grisanti, L.; González Lebrero, M. C.; Kaminski Schierle, G. S.; Hassanali, A. Non-Aromatic Fluorescence in Biological Matter: The Exception or the Rule? *The Journal of Physical Chemistry B* **2022**, *126*, 7203–7211, PMID: 36128666

[13] Arnon, Z. A.; Kreiser, T.; Yakimov, B.; Brown, N.; Aizen, R.; Shaham-Niv, S.; Makam, P.; Qaisrani, M. N.; Poli, E.; Ruggiero, A.; Slutsky, I.; Hassanali, A.; Shirshin, E.; Levy, D.; Gazit, E. On-off transition and ultrafast decay of amino acid luminescence driven by modulation of supramolecular packing. *iScience* **2021**, *24*, 102695

[14] Grisanti, L.; Sapunar, M.; Hassanali, A.; Došlić, N. Toward Understanding Optical Properties of Amyloids: A Reaction Path and Nonadiabatic Dynamics Study. *Journal of the American Chemical Society* **2020**, *142*, 18042–18049, PMID: 32966740

[15] Kumar, A.; Ahari, D.; Priyadarshi, A.; Ziauddin Ansari, M.; Swaminathan, R. Weak Intrinsic Luminescence in Monomeric Proteins Arising from Charge Recombination. *The Journal of Physical Chemistry B* **2020**, *124*, 2731–2746, PMID: 32155070

[16] Prasad, S.; Mandal, I.; Singh, S.; Paul, A.; Mandal, B.; Venkatramani, R.; Swaminathan, R. Near UV-Visible electronic absorption originating from charged amino acids in a monomeric protein. *Chem. Sci.* **2017**, *8*, 5416–5433

[17] Shukla, A.; Mukherjee, S.; Sharma, S.; Agrawal, V.; Radha Kishan, K.; Guptasarma, P. A novel UV laser-induced visible blue radiation from protein crystals and aggregates: scattering artifacts or fluorescence transitions of peptide electrons delocalized through hydrogen bonding? *Archives of Biochemistry and Biophysics* **2004**, *428*, 144–153

[18] del Mercato, L. L.; Pompa, P. P.; Maruccio, G.; Torre, A. D.; Sabella, S.; Tamburro, A. M.; Cingolani, R.; Rinaldi, R. Charge transport and intrinsic fluorescence in amyloid-like fibrils. *Proceedings of the National Academy of Sciences* **2007**, *104*, 18019–18024

[19] Ye, R.; Liu, Y.; Zhang, H.; Su, H.; Zhang, Y.; Xu, L.; Hu, R.; Kwok, R. T. K.; Wong, K. S.; Lam, J. W. Y.; Goddard, W. A.; Tang, B. Z. Non-conventional fluorescent biogenic and synthetic polymers without aromatic rings. *Polym. Chem.* **2017**, *8*, 1722–1727

[20] Stephens, A. D. et al. Short hydrogen bonds enhance nonaromatic protein-related fluorescence. *Proceedings of the National Academy of Sciences* **2021**, *118*, e2020389118

[21] Dereka, B.; Yu, Q.; Lewis, N. H. C.; Carpenter, W. B.; Bowman, J. M.; Tokmakoff, A. Crossover from hydrogen to chemical bonding. *Science* **2021**, *371*, 160–164

[22] Mališ, M.; Loquais, Y.; Gloaguen, E.; Jouvet, C.; Brenner, V.; Mons, M.; Ljubić, I.; Došlić, N. Non-radiative relaxation of UV photoexcited phenylalanine residues: probing the role of conical intersections by chemical substitution. *Phys. Chem. Chem. Phys.* **2014**, *16*, 2285–2288

[23] Matsika, S.; Krause, P. Nonadiabatic events and conical intersections. *Annual review of physical chemistry* **2011**, *62*, 621–643

[24] Zhou, S.; Wang, L. Unraveling the structural and chemical features of biological short hydrogen bonds. *Chem. Sci.* **2019**, *10*, 7734–7745

[25] Tang, S.; Yang, T.; Zhao, Z.; Zhu, T.; Zhang, Q.; Hou, W.; Yuan, W. Z. Nonconventional luminophores: characteristics, advancements and perspectives. *Chem. Soc. Rev.* **2021**, *50*, 12616–12655

[26] Crespo-Otero, R.; Li, Q.; Blancafort, L. Exploring Potential Energy Surfaces for Aggregation-Induced Emission—From Solution to Crystal. *Chemistry – An Asian Journal* **2019**, *14*, 700–714

[27] Peng, X.-L.; Ruiz-Barragan, S.; Li, Z.-S.; Li, Q.-S.; Blancafort, L. Restricted access to a conical intersection to explain aggregation induced emission in dimethyl tetraphenylsilole. *J. Mater. Chem. C* **2016**, *4*, 2802–2810

[28] Surre, J.; Saint-Ruf, C.; Collin, V.; Orenga, S.; Ramjeet, M.; Matic, I. Strong increase in the autofluorescence of cells signals struggle for survival. *Scientific Reports* **2018**, *8*, 12088

[29] Jun, Y. W.; Kim, H. R.; Reo, Y. J.; Dai, M.; Ahn, K. H. Addressing the autofluorescence issue in deep tissue imaging by two-photon microscopy: the significance of far-red emitting dyes. *Chem. Sci.* **2017**, *8*, 7696–7704

[30] Rosenman, G.; Apter, B. Bioinspired materials: Physical properties governed by biological refolding. *Applied Physics Reviews* **2022**, *9*, 021303

[31] Apter, B.; Lapshina, N.; Lapsker, I.; Handelman, A.; Accardo, A.; Diaferia, C.; Morelli, G.; Rosenman, G. Fold-Sensitive Visible Fluorescence in β-Sheet Peptide Structures. *Advanced Optical Materials* **2021**, *9*, 2002247

[32] Colletier, J.-P.; Laganowsky, A.; Landau, M.; Zhao, M.; Soriaga, A. B.; Goldschmidt, L.; Flot, D.; Cascio, D.; Sawaya, M. R.; Eisenberg, D. Molecular basis for amyloid-β polymorphism. *Proceedings of the National Academy of Sciences* **2011**, *108*, 16938–16943

[33] Tully, J. C. Molecular dynamics with electronic transitions. *Journal of Chemical Physics* **1990**, *93*, 1061–1071

[34] Crespo-Otero, R.; Barbatti, M. Recent Advances and Perspectives on Nonadiabatic Mixed Quantum-Classical Dynamics. *Chemical Reviews* **2018**, *118*, 7026–7068

[35] Tavernelli, I.; Curchod, B. F.; Rothlisberger, U. Nonadiabatic molecular dynamics with solvent effects: A LR-TDDFT QM/MM study of ruthenium (II) tris (bipyridine) in water. *Chemical Physics* **2011**, *391*, 101–109

[36] Tapavicza, E.; Tavernelli, I.; Rothlisberger, U. Trajectory surface hopping within linear response time-dependent density-functional theory. *Physical Review Letters* **2007**, *98*, 1–4

[37] Barbatti, M. Nonadiabatic dynamics with trajectory surface hopping method. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2011**, *1*, 620–633

[38] Granucci, G.; Persico, M. Critical appraisal of the fewest switches algorithm for surface hopping. *The Journal of chemical physics* **2007**, *126*, 134114

[39] Case, D. A.; Cheatham III, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz Jr, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular simulation programs. *Journal of computational chemistry* **2005**, *26*, 1668–1688

[40] Hehre, W. J.; Ditchfield, R.; Pople, J. A. Self—consistent molecular orbital methods. XII. Further extensions of Gaussian—type basis sets for use in molecular orbital studies of organic molecules. *The Journal of Chemical Physics* **1972**, *56*, 2257–2261

[41] Perdew, J. P.; Ernzerhof, M.; Burke, K. Rationale for mixing exact exchange with density functional approximations. *The Journal of chemical physics* **1996**, *105*, 9982–9985

[42] Nitsche, M. A.; Ferreria, M.; Mocskos, E. E.; Gonzalez Lebrero, M. C. GPU accelerated implementation of density functional theory for hybrid QM/MM simulations. https://github.com/MALBECC/lio. *Journal of Chemical Theory and Computation* **2014**, *10*, 959–967

[43] Casida, M. E. *Recent Advances In Density Functional Methods: (Part I)*; World Scientific, 1995; pp 155–192

[44] Hirata, S.; Head-Gordon, M. Time-dependent density functional theory within the Tamm–Dancoff approximation. *Chem. Phys. Lett.* **1999**, *314*, 291–299

[45] Send, R.; Furche, F. First-order nonadiabatic couplings from time-dependent hybrid density functional response theory: Consistent formalism, implementation, and performance. *Journal of Chemical Physics* **2010**, *132*

[46] Tapavicza, E.; Bellchambers, G. D.; Vincent, J. C.; Furche, F. Ab initio nonadiabatic molecular dynamics. *Physical Chemistry Chemical Physics* **2013**, *15*, 18336–18348

[47] Hutter, J.; Iannuzzi, M.; Schiffmann, F.; VandeVondele, J. cp2k: atomistic simulations of condensed matter systems. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2014**, *4*, 15–25

[48] Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Physical review A* **1988**, *38*, 3098

[49] Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical review B* **1988**, *37*, 785

[50] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *The Journal of chemical physics* **2010**, *132*, 154104

[51] Coonjobeeharry, J.; Spinlove, K. E.; Sanz Sanz, C.; Sapunar, M.; Došlić, N.; Worth, G. A. Mixed-quantum-classical or fully-quantized dynamics? A unified code to compare methods. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **2022**, *380*, 20200386

[52] Curchod, B. F.; Martínez, T. J. Ab Initio Nonadiabatic Quantum Molecular Dynamics. *Chemical Reviews* **2018**, *118*, 3305–3336

[53] Pulay, P. A perspective on the CASPT2 method. *International Journal of Quantum Chemistry* **2011**, *111*, 3273–3279

[54] Bearpark, M. J.; Robb, M. A.; Bernhard Schlegel, H. A direct method for the location of the lowest energy point on a potential surface crossing. *Chemical Physics Letters* **1994**, *223*, 269–274

[55] Bearpark, M. J.; Larkin, S. M.; Vreven, T. Searching for Conical Intersections of Potential Energy Surfaces with the ONIOM Method: Application to Previtamin D. *The Journal of Physical Chemistry A* **2008**, *112*, 7286–7295, PMID: 18636693

[56] Avagliano, D.; Bonfanti, M.; Nenov, A.; Garavelli, M. Automatized protocol and interface to simulate QM/MM time-resolved transient absorption at TD-DFT level with COBRAMM. *Journal of Computational Chemistry* **2022**, *43*, 1641–1655

[57] Conti, I.; Bonfanti, M.; Nenov, A.; Rivalta, I.; Garavelli, M. In *QM/MM Studies of Light-responsive Biological Systems*; Andruniów, T., Olivucci, M., Eds.; Springer International Publishing: Cham, 2021; pp 77–142

[58] Zobel, J. P.; Nogueira, J. J.; González, L. The IPEA dilemma in CASPT2. *Chem. Sci.* **2017**, *8*, 1482–1499

[59] Forsberg, N.; Åke Malmqvist, P. Multiconfiguration perturbation theory with imaginary level shift. *Chemical Physics Letters* **1997**, *274*, 196–204

[60] Frisch, M. J. et al. Gaussian˜16 Revision C.01. 2016; Gaussian Inc. Wallingford CT

[61] Fdez. Galván, I. et al. OpenMolcas: From Source Code to Insight. *Journal of Chemical Theory and Computation* **2019**, *15*, 5925–5964, PMID: 31509407

[62] Aquilante, F. et al. Modern quantum chemistry with [Open]Molcas. *The Journal of Chemical Physics* **2020**, *152*, 214117