The Journal of Chemical Physics

RESEARCH ARTICLE | FEBRUARY 14 2024

Quantum dynamics of excited state proton transfer in green fluorescent protein

Special Collection: 2024 JCP Emerging Investigators Special Collection

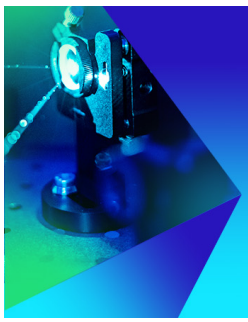
Susannah Bourne-Worster ; Graham A. Worth

Check for updates

J. Chem. Phys. 160, 065102 (2024)
https://doi.org/10.1063/5.0188834

View Online
Export Citation
CrossMark

19 February 2024 16:50:08


The Journal of Chemical Physics
Special Topic: Time-resolved Vibrational Spectroscopy
Submit Today
AIP Publishing

AIP Publishing

The Journal of Chemical Physics ARTICLE pubs.aip.org/aip/jcp

Quantum dynamics of excited state proton transfer in green fluorescent protein

Cite as: J. Chem. Phys. 160, 065102 (2024); doi: 10.1063/5.0188834
Submitted: 24 November 2023 • Accepted: 18 January 2024 •
Published Online: 14 February 2024

View Online Export Citation CrossMark

Susannah Bourne-Worster[a] and Graham A. Worth[b]

AFFILIATIONS

Department of Chemistry, University College London, London WC1H 0AJ, United Kingdom

Note: This paper is part of the 2024 JCP Emerging Investigators Special Collection.
[a]**Author to whom correspondence should be addressed:** susannah.bourne-worster@ucl.ac.uk
[b]**E-mail:** g.a.worth@ucl.ac.uk

ABSTRACT

Photoexcitation of green fluorescent protein (GFP) triggers long-range proton transfer along a "wire" of neighboring protein residues, which, in turn, activates its characteristic green fluorescence. The GFP proton wire is one of the simplest, most well-characterized models of biological proton transfer but remains challenging to simulate due to the sensitivity of its energetics to the surrounding protein conformation and the possibility of non-classical behavior associated with the movement of lightweight protons. Using a direct dynamics variational multiconfigurational Gaussian wavepacket method to provide a fully quantum description of both electrons and nuclei, we explore the mechanism of excited state proton transfer in a high-dimensional model of the GFP chromophore cluster over the first two picoseconds following excitation. During our simulation, we observe the sequential starts of two of the three proton transfers along the wire, confirming the predictions of previous studies that the overall process starts from the end of the wire furthest from the fluorescent chromophore and proceeds in a concerted but asynchronous manner. Furthermore, by comparing the full quantum dynamics to a set of classical trajectories, we provide unambiguous evidence that tunneling plays a critical role in facilitating the leading proton transfer.

© 2024 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/). https://doi.org/10.1063/5.0188834

19 February 2024 16:50:08

I. INTRODUCTION

Long-range transfer of protons along a "wire" via multiple proton transfer reactions is an important process in biology, particularly for pumping protons against a gradient (as occurs in respiration). Green fluorescent protein (GFP), which is the quintessential example of a fluorescent protein,[1] is also highly studied as a small, stable model system for the functioning of proton wires.[2] GFP has an 11-stranded beta-barrel structure with a single alpha helix hanging down the center of the barrel.[3] The GFP chromophore, *p*-hydroxybenzylideneimidazolidinone (HBDI), forms *in situ* halfway along the central helix via a cyclization reaction between exposed serine, tyrosine, and glycine residues.

Wild-type GFP (*wt*-GFP) has two major absorption bands at 398 and 478 nm, corresponding to neutral (protonated) and anionic (deprotonated at the phenolic oxygen) forms of the chromophore.[4]. The neutral, protonated form is the most prevalent in the ground state, although both forms exist and can interconvert via a series of three proton transfers along a wire formed by a structural water molecule and two nearby sidechain residues (Ser205 and Glu222) on the exposed interior of the GFP barrel[3,5] (shown in Fig. 1). Regardless of the excitation wavelength, *wt*-GFP produces a characteristic green fluorescence at around 502 nm. Assignment of this emission to the relaxed anionic form of the chromophore indicates that proton transfer also occurs in the excited state.[4] Fluorescence from the excited anion, which occurs with a quantum yield of 0.8,[3,6] is typically followed by reverse proton transfer to reform the neutral ground state.[4,5]

Structural constraints imposed by the protein pocket,[7] contracted donor–acceptor distances along the proton wire, and the extremely low pKa[2] of the excited neutral chromophore relative to the ground state means that excited state proton transfer (ESPT) occurs much faster than the corresponding ground state process (GSPT).[8] Consequently, ESPT in GFP is less well understood than GSPT, which has been extensively characterized, both experimentally and theoretically.[2] Transient electronic absorption measurements reveal that ESPT occurs over two distinct timescales (2–3 and 10 ps for *wt*-GFP) and has a large kinetic isotope effect (timescales

© Author(s) 2024

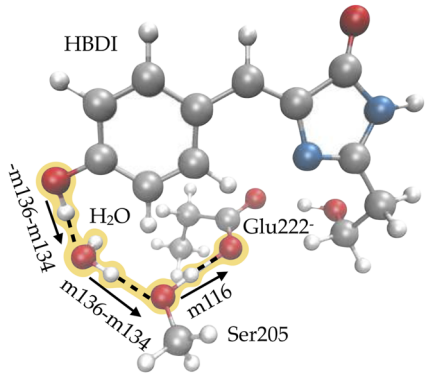


FIG. 1. Cluster model of the GFP chromophore and proton wire. The proton wire is highlighted in yellow. The OH bonds along the proton wire are labeled with the vibrational normal mode coordinates (Sec. III A) that describe their length.

increase to 11–12 and 69 ps for deuterated GFP).[9] The shorter of the two timescales appears to be largely temperature independent, while the longer timescale increases further at lower temperatures.

The interpretation of the biphasic ESPT kinetics is still under debate. One study suggests that the initial fast timescale corresponds to formation of an intermediate state with extended low-barrier hydrogen bonds along the proton wire and that the full proton transfer only occurs over the longer 10 ps timescale.[10] However, transient infrared spectroscopy shows that the fast and slow timescales are associated with different COO^- stretching frequencies in the final proton acceptor, leading to the alternative suggestion that the biphasic kinetics are a result of structural inhomogeneity, with the faster timescale corresponding to stronger hydrogen-bonding around the carboxylate group of Glu222.[11,12]

The debate around whether or not the two proton transfer timescales reflect a multistage process is linked to wider questions around whether the three proton transfers happen in series (stepwise transfer) or as a single concerted process and, if concerted, whether the transfers are triggered in stages (asynchronous transfer) or all start and end at approximately the same time (synchronous transfer). The corresponding GSPT process has been confirmed experimentally[2,8] and theoretically[13,14] to proceed via a concerted, asynchronous mechanism where the initial, rate-limiting step is proton transfer from Ser205 to Glu222. The remaining two proton transfers are barrierless and rapidly follow the first transfer. Since excitation at 398 nm is localized on the GFP chromophore, away from Glu222, it is plausible that ESPT would follow a similar mechanism.

The majority of experimental studies appear to support the idea of a concerted mechanism for ESPT. No intermediate species has yet been unambiguously detected,[2,4,10,15] and transient infrared[15] and electronic[4,16] absorption measurements seem to show matching timescales for the decay of the initially excited chromophore and production of the deprotonated product. Ultrafast transient infrared and fluorescence spectroscopy further suggests that the process is asynchronous with a rate limited by the first proton to start transferring.[17] As the acceptor conformation appears to be correlated with the ESPT rate,[18] this initial, rate-limiting proton transfer is thought to be at the far end of the chain, from Ser205 to Glu222, as in GSPT.

However, attempts to confirm these observations theoretically are inconclusive. High-level [(6,6) CASPT2] calculations of the potential energy along an expected reaction coordinate for the proton wire suggest that the lowest energy path is a stepwise transfer of the protons starting with the transfer of H_{Ser205} from Ser205 to Glu222, although the synchronous, concerted transfer of all three protons is only very slightly higher in energy[19] and may even be the lowest energy route for certain configurations of the chromophore and surrounding residues.[20] These simulations were limited to only 6 degrees of freedom—the relative positions of the three oxygen and three hydrogen atoms of the proton transfer wire. Early dynamical modeling using MCTDH (the multiconfigurational time-dependent Hartree method) with the same 6 degrees of freedom appeared to favor the synchronous, concerted mechanism, although they showed a considerable mismatch between the simulated and measured rate of ESPT.[21] Later studies recognize that the energetic cost of stretching (and breaking) the OH bonds along the proton wire is strongly influenced by the geometry of the rest of the system,[2,22,23] to the extent that certain (low-frequency) vibrational motions of the chromophore may be essential in setting up a favorable geometry for proton transfer.[2,22] QM/MM potential energy surfaces calculated for the six proton wire coordinates described above indicate that the asynchronous, concerted transfer led by the transfer of H_{Ser205} is the most energetically favorable pathway.[23] This suggestion is supported by QM/MM molecular dynamics simulations of the full GFP protein,[22] but attempts to describe the ESPT process using fully quantum MCTDH dynamics with the QM/MM potential in Ref. 23 were unable to successfully observe any proton transfer events, either due to the short simulation timescale or due to freezing all but the six atoms directly involved in the proton wire.

The rates of many biological proton transfers are heavily influenced by tunneling. This is, for example, a very important effect in enzyme catalysis.[24] GFP is an ideal system for studying the role of tunneling in proton transfer. The relatively small range of donor–acceptor distances imposed by its rigid beta-barrel structure makes it much easier to identify tunneling effects on the proton transfer kinetics than in highly flexible enzymes.[2] Furthermore, the GFP proton wire includes a serine residue, whose large ionization energy creates a high potential barrier that, classically, should inhibit proton transfer.[8] Nonetheless, serine residues are frequently involved in biological proton transfer processes.[25,26] It has been suggested that they help control the direction of transfer because their donor and acceptor orbitals point in different directions, giving very different tunneling distances (and, therefore, tunneling rates) for forward and backward transfers.[2,8] The large kinetic isotope effect observed for the ground state proton transfer in GFP (KIE = 40 at room temperature) is well fitted by a model of vibrationally assisted deep tunneling (long-range tunneling through the base of a high potential energy barrier) associated with the transfer of H_{Ser205} from Ser205 to Glu222.[8] However, while ESPT is also associated with a large kinetic isotope effect (about 5, independent of temperature[9,16,27]), the origin of this effect is harder to elucidate. Previous studies have variously suggested that the non-Arrhenius behavior of ESPT could reflect the complex kinetics of multiple proton transfer pathways,[27] the combined effect of isotopic differences

19 February 2024 16:50:08

© Author(s) 2024

in zero-point energy across the three proton transfers,[16] or a role for tunneling.[2,16,27]

In this work, we present a fully quantum simulation of the nuclear dynamics of ESPT in GFP, including many more degrees of freedom than have been considered in previous studies. The model used for the potential surfaces is based on the proton transfer wire in the protein but ignores the protein environment. In this way, the importance of nuclear quantum effects for the transfer mechanism can be observed without mixing in the effects due to the rearrangement of the environment that more sophisticated models have observed to be required for the proton transfer to occur with classical nuclei.[22,28] Our results provide a mechanism that agrees with the earlier work in terms of the order and (a)synchronicity of the three proton transfers and obtains a timescale that can be related to experimental observations. Finally, by comparing to a fully classical description of the dynamics, we demonstrate the potential importance of tunneling in the ESPT mechanism.

II. THEORY AND METHODS

A. Structural model

Since the full GFP protein is too large to model quantum mechanically, a number of cluster models have been developed to capture the chromophore and key elements of surrounding residues that participate in the proton wire.[13,20,21] We use the simple model developed by Zhang *et al.*,[29] shown in Fig. 1, which includes the HBDI chromophore and proton transfer chain from the crystal structure 1EMB.[3] The Ser205 residue is replaced by a representative methanol molecule, while Glu222 is represented by an acetate ion. The cluster geometry was originally optimized at the B3LYP/cc-pvdz level of theory. We reoptimized the geometry using ωB97xd/6-31G*, which gives a better description of the stretched O–H bonds along the proton wire (see Sec. S1 of the supplementary material) and is comparable to the level of theory used in recent *ab initio* MD studies of ESPT.[22]

The coordinates of the reoptimized structure are given in Sec. S2 of the supplementary material along with a comparison of the optimized structure to the original cluster model taken from the crystal structure. The RMSD (root mean square deviation) of the optimized structure from the original cluster model is 0.59 Å, showing that the geometry is similar to that found in the protein. In addition, in Sec. S2, a comparison is made of the optimized structure to an optimized structure at the ωB97xd/6-31++G** level of theory to check the importance of diffuse functions. The RMSD between these structures is 0.35 Å, meaning that the diffuse functions do not radically affect the geometry of the cluster.

B. Quantum dynamics

1. Variational multiconfigurational Gaussian approach

The photoexcited dynamics of the proton wire were simulated using a variational multiconfigurational Gaussian (vMCG) approach, implemented in the Quantics software package.[30,31] This method is detailed in a number of other publications (e.g., Refs. 32–35), so here we simply outline the key principles.

The aim is to solve the time-dependent Schrödinger equation for the full electronic and nuclear wave function,

$$i\hbar\frac{\partial}{\partial t}\Psi(\boldsymbol{R},\boldsymbol{r},t)=\hat{H}\Psi(\boldsymbol{R},\boldsymbol{r},t). \tag{1}$$

Treating the nuclei quantum mechanically is particularly important for understanding proton transfer reactions, where the light weight of the hydrogen atoms increases the likelihood of non-classical behavior such as tunneling.

By expressing the full wave function in a basis of electronic states $\psi_i(\boldsymbol{r};\ \boldsymbol{R})$ (eigenfunctions of the clamped-nucleus electronic Hamiltonian) as

$$\Psi(\boldsymbol{R},\boldsymbol{r},t)=\sum_i \chi_i(\boldsymbol{R},\boldsymbol{t})\psi_i(\boldsymbol{r};\boldsymbol{R},t), \tag{2}$$

Equation (1) can be rewritten as a Schrödinger equation for the nuclear wave functions $\chi_i(\boldsymbol{R},t)$,[36]

$$\left[\frac{-1}{2M}(\nabla+\boldsymbol{F})^2+\boldsymbol{V}\right]\chi_i(\boldsymbol{R},t)=i\hbar\frac{\partial\chi_i(\boldsymbol{R},t)}{\partial t}. \tag{3}$$

The dressed kinetic energy operator $\left(\frac{-1}{2M}(\nabla+\boldsymbol{F})^2\right)$ contains both the standard nuclear kinetic energy operator and the derivative coupling vector $\boldsymbol{F}_{ij}=\langle\psi_i|\nabla\psi_j\rangle$ between electronic states i and j. $\boldsymbol{V}$ is the potential energy operator. Current evidence supports the idea that proton transfer in GFP is electronically adiabatic, so from here on, the subscript on χ will be dropped.[2,8]

In vMCG, the nuclear wave functions are expanded in terms of a basis of N Gaussian wavepackets $\{g_j(\boldsymbol{x},t)\}$ as

$$\chi(\boldsymbol{x},t)=\sum_{j=1}^{N}A_j(t)g_j(\boldsymbol{x},t). \tag{4}$$

Each Gaussian wavepacket is defined by

$$g_j(\boldsymbol{x},t)=\exp\left(\boldsymbol{x}^T\cdot\boldsymbol{\sigma}_j\cdot\boldsymbol{x}+\boldsymbol{\xi}_j\cdot\boldsymbol{x}+\eta_j\right), \tag{5}$$

where the diagonal matrix $\boldsymbol{\sigma}_j$ governs the width of the Gaussian in each dimension, the vector $\boldsymbol{\xi}_j$ dictates the position and momentum of the center of the wavepacket, and the scalar factor η_j contains the phase of the wavepacket.[37] For convenience, vMCG uses "frozen" Gaussians, with a fixed value of $\boldsymbol{\sigma}_j$.[34] We choose the Gaussians to have a width (standard deviation) of $1/\sqrt{2}$ in position along all normal mode coordinates. This corresponds to the width of the $v=0$ eigenfunction for a simple harmonic oscillator on the ground electronic state in the mass-frequency scaled normal mode coordinates used.[33] η_j is constructed to make $g_j(\boldsymbol{x},t)$ phaseless [the phase is factored out into the expansion coefficients $A_j(t)$] and to keep the Gaussians normalized. The variational nature of the wavepacket is preserved in $\boldsymbol{\xi}_j$, which is propagated fully variationally.

At $t=0$, the set of GWPs all have the same center coordinate, here the Franck–Condon point, and different momenta. The first GWP has zero momentum, and subsequent functions are displaced in momentum space sequentially along each mode so that each normal mode has a basis function in which it is excited. The expansion coefficients are then set so that only the GWP with zero momentum is initially populated. The initial wavepacket is thus the ground-state vibrational wave function in the harmonic approximation at 0 K, and the basis set is ready to adapt to configurations with excitations in the normal modes.

19 February 2024 16:50:08

© Author(s) 2024

The Gaussian wavepackets form a nonorthogonal basis, which has to be handled carefully to avoid singularities arising in the case of strongly overlapping functions. The choice to use a Gaussian basis despite this added complication is motivated by their well-defined shape and properties.[38] At any given time, each wavepacket is localized onto a well-defined region of the potential energy surface, within which it is reasonable to approximate the potential as a simple harmonic function (particularly if the width of the Gaussian is sufficiently small). Using this local harmonic approximation to the potential, the matrix elements of the Hamiltonian reduce to simple functions of the moments of the Gaussian functions ($\langle g_i|x^n|g_j\rangle$), which can be evaluated analytically based on the value of the potential at the center of the Gaussian. Since the potential only needs to be calculated at one well-defined point for each Gaussian, there is no need to evaluate the whole potential over a finite grid in advance of any time propagation (as is done in the related multiconfigurational time-dependent Hartree method). This makes it possible to evaluate the potential on the fly, which offers considerable advantages for high-dimensional potentials and bond-breaking reactions of the type being investigated here (see Sec. II B 2).

Equations of motion for the Gaussian wavepackets and expansion coefficients are derived by variationally minimizing the wave function according to the Dirac–Frenkel variational principle,

$$\langle \delta\Psi|H - i\hbar\frac{\partial}{\partial t}|\Psi\rangle = 0. \tag{6}$$

The expansion coefficients $A_j(t)$ are propagated in time according to

$$i\dot{A}_j(t) = \sum_{kl} S^{-1}_{jk}(t)(H_{kl}(t) - i\tau_{kl}(t))A_l(t), \tag{7}$$

where $S_{ij}(t) = \langle g_i(t)|g_j(t)\rangle$ is the overlap between Gaussian basis functions, $\tau_{ij}(t) = \langle g_i(t)|\frac{\partial}{\partial t}g_j(t)\rangle$ is the differential overlap, and $H_{ij} = \langle g_i(t)|\left[\frac{-1}{2M}\nabla^2 + \boldsymbol{V}\right]|g_j(t)\rangle$ is a matrix element of the Hamiltonian constructed with the potential energy operator $\boldsymbol{V}$ appropriate to the electronic state of interest.

The equation of motion for the Gaussian center parameter (the only one that is propagated variationally) is[34]

$$i\dot{\boldsymbol{\xi}} = \boldsymbol{C}^{-1}\boldsymbol{Y}, \tag{8}$$

where

$$C_{i\alpha,j\beta} = \rho_{ij}\left(S^{(\alpha\beta)}_{ij} - \left[\boldsymbol{S}^{(\alpha 0)}\boldsymbol{S}^{-1}\boldsymbol{S}^{(0\beta)}\right]_{ij}\right), \tag{9}$$

$$Y_{i\alpha} = \rho_{ij}\left(H^{(\alpha 0)}_{ij} - \left[\boldsymbol{S}^{(\alpha 0)}\boldsymbol{S}^{-1}\boldsymbol{H}\right]_{ij}\right). \tag{10}$$

In Eqs. (9) and (10), $\rho_{ij} = A^*_i A_j$ is referred to as the density matrix. Various differential overlaps are defined with respect to the Gaussian parameters $\{\boldsymbol{\sigma}, \boldsymbol{\xi}, \eta\}$. $\lambda_{i\alpha}$ is the αth parameter for the function g_i. In this notation, $\boldsymbol{S}^{(\alpha\beta)}$ is a matrix with elements

$$S^{\alpha\beta}_{ij} = \left\langle \frac{\delta g_i}{\delta\lambda_{i,\alpha}} \middle| \frac{\delta g_j}{\delta\lambda_{j,\beta}} \right\rangle \tag{11}$$

and

$$H^{\alpha\beta}_{ij} = \left\langle \frac{\delta g_i}{\delta\lambda_{i,\alpha}} \middle| H \middle| \frac{\delta g_j}{\delta\lambda_{j,\beta}} \right\rangle. \tag{12}$$

An index of α or β equal to zero indicates that no derivative is taken.

The equations of motion for A_j and $\boldsymbol{\xi}$ are solved numerically using a fifth-order Runge–Kutta method with an adaptive step size. The error threshold for accepting a step in the propagation is 10^{-6}.

2. Direct dynamics

As has been observed previously,[19,21,23] the challenge of characterizing the full multidimensional potential energy surface for the ESPT is often a limiting factor in dynamical simulations. To help overcome this issue, we use a direct dynamics approach, where points on the potential energy surface are calculated on the fly, as they are explored by the Gaussian wavepackets.

At a given time step, the potential energy surface $\boldsymbol{V}(\boldsymbol{x})$ in the vicinity of each individual Gaussian wavepacket is approximated by a second order Taylor expansion $\boldsymbol{T}_i(\boldsymbol{x})$ around the coordinates $(\boldsymbol{x}_0)$ at the center of the wavepacket,

$$\begin{aligned}\boldsymbol{V}(\boldsymbol{x}) = \boldsymbol{T}_i(\boldsymbol{x}) = \boldsymbol{V}(\boldsymbol{x}_0) + \boldsymbol{g}(\boldsymbol{x}_0)^T \cdot (\boldsymbol{x} - \boldsymbol{x}_0) \\ + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_0)^T \cdot \boldsymbol{H}(\boldsymbol{x}_0)^T \cdot (\boldsymbol{x} - \boldsymbol{x}_0).\end{aligned} \tag{13}$$

This is known as the local harmonic approximation (LHA). Equation (13) requires the energy $\boldsymbol{V}(\boldsymbol{x}_0)$, gradient $\boldsymbol{g}(\boldsymbol{x}_0)$, and Hessian $\boldsymbol{H}(\boldsymbol{x}_0)$ to be evaluated at the center point of each Gaussian wavepacket at each time step of the simulation. Each time a quantum chemical calculation of these quantities is performed, their values are stored in a database and can be reused at future time steps. New calculations are only carried out for structures where the maximum atomic displacement relative to all existing structures in the database exceeds a threshold value `dbmin`. In our simulations, `dbmin` was initially set to 0.05 to populate the database with some starting structures, before being increased to 0.15 to propagate out to longer times.

For structures that lie within the threshold distance of a database point, the local potential energy surface is estimated as a weighted sum of the LHA potential energy surface around all points in the database using a modified Shepard interpolation scheme,

$$\boldsymbol{V}(\boldsymbol{x}) = \sum_{i\in \mathrm{DB}} w_i(\boldsymbol{x})\boldsymbol{T}_i(\boldsymbol{x}). \tag{14}$$

The normalized weights[33,39]

$$w_i(\boldsymbol{x}) = \frac{v_i(\boldsymbol{x})}{\sum_j v_j(\boldsymbol{x})}, \tag{15}$$

$$v_i(\boldsymbol{x}) = \left(\left[\frac{\|\boldsymbol{x} - \boldsymbol{x}_i\|}{\mathrm{rad}_i}\right]^4 + \left[\frac{\|\boldsymbol{x} - \boldsymbol{x}_i\|}{\mathrm{rad}_i}\right]^{24}\right)^{-1}, \tag{16}$$

decrease with the distance $(\|\boldsymbol{x} - \boldsymbol{x}_i\|)$ between the new structure $\boldsymbol{x}$ and the database structure $\boldsymbol{x}_i$. These distances are expressed relative to a confidence radius $\mathrm{rad}_i = \|\boldsymbol{x}_j - \boldsymbol{x}_i\|$, which is the distance between database structure $\boldsymbol{x}_i$ and a reference database structure $\boldsymbol{x}_j$ that is

19 February 2024 16:50:08

© Author(s) 2024

close to $\boldsymbol{x}$. When $\|\boldsymbol{x}-\boldsymbol{x}_i\| < \mathrm{rad}_i$, the second term in Eq. (16) dominates (and tends to zero), meaning that only a handful of database points close to $\boldsymbol{x}$ will contribute to the interpolated potential. This helps keep the potential energy surfaces smooth for larger systems.

Hessians are particularly expensive to calculate, and for this reason, the Hessian is only calculated from scratch for the very first point added to the database. At every other point on the potential energy surface, the Hessian is estimated using the Hessian updating scheme described in Refs. 32 and 40. Under this scheme, the Hessian at a new point $\boldsymbol{x}$ is given by the weighted average

$$\boldsymbol{H}(\boldsymbol{x}) = \frac{\sum_i \|\boldsymbol{x}-\boldsymbol{x}_i\|^{-4}\boldsymbol{H}_{\mathrm{est}}^{(i)}}{\sum_i \|\boldsymbol{x}-\boldsymbol{x}_i\|^{-4}}, \tag{17}$$

where the index i runs over all database points that are closer to the initial (reference) point than $\boldsymbol{x}$. $\boldsymbol{H}_{\mathrm{est}}^{(i)}$ is an estimate of the Hessian extrapolated from point i using the Powell updating scheme,

$$\boldsymbol{H}_{\mathrm{est}}^{(i)} = \boldsymbol{H}(\boldsymbol{x}_i) + \frac{1}{\boldsymbol{\delta}\cdot\boldsymbol{\delta}}(\boldsymbol{\epsilon}\otimes\boldsymbol{\delta}+\boldsymbol{\delta}\otimes\boldsymbol{\epsilon}) - \frac{\boldsymbol{\epsilon}\cdot\boldsymbol{\delta}}{(\boldsymbol{\delta}\cdot\boldsymbol{\delta})^2}\boldsymbol{H}(\boldsymbol{x}_i)\cdot\boldsymbol{\delta}\otimes\boldsymbol{\delta}\cdot\boldsymbol{H}(\boldsymbol{x}_i), \tag{18}$$

where $\boldsymbol{\epsilon} = \boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{g}(\boldsymbol{x}_i)$ and $\boldsymbol{\delta} = \boldsymbol{x} - \boldsymbol{x}_i$. Once $\boldsymbol{H}(\boldsymbol{x})$ has been added to the database, the Hessian at every database point further away from the reference point than $\boldsymbol{x}$ is also updated using the same scheme so that the Hessian at every point is based on all points that lie between it and the reference point.

For the simulations presented in this paper, the energies and gradients of the adiabatic surface are calculated using time-dependent density functional theory (TDDFT) with an ωB97XD functional and 6-31G* basis set. Stretching of the OH bonds in the proton wire and the resulting proton transfer is well-described by range-separated hybrid functionals (see Sec. S1 of the supplementary material). The ωB97XD functional additionally contains an empirical dispersion correction. Inclusion of this correction does not appear to make a significant difference to the description of ESPT energetics, compared to other range-separated functionals, such as CAM-B3LYP,[41] but does improve the description of non-covalent interactions, such as hydrogen bonds, in the ground state optimization used to generate the starting geometry.[42] Polarization functions are included in the basis set to capture the polarization of the OH bonds. However, the basis set is kept small to make it computationally tractable to simulate ESPT over the relatively long experimental timescale.

C. Classical dynamics

To explore the importance of tunneling in promoting proton transfer along the proton wire, the full quantum dynamics were compared to classical dynamics from which all quantum effects had been removed.

A set of frozen Gaussian wavepackets were propagated on the same TDDFT-level potential energy surface, using the structures, energies, gradients, and Hessians stored in the direct dynamics database described above. They were propagated according to a classical equation of motion obtained by factoring out and discarding all second order and higher Gaussian moments from Eq. (10) to leave

$$i\dot{\boldsymbol{\xi}}_j = -2\boldsymbol{\sigma}_j\cdot\dot{\boldsymbol{q}} - i\left(\frac{\partial \boldsymbol{V}}{\partial \boldsymbol{x}}\right)_{\boldsymbol{x}=\boldsymbol{q}}, \tag{19}$$

where $\boldsymbol{q}$ is the position of the center of the Gaussian wavepacket. Equation (19) describes the classical propagation of the wavepacket based only on the momentum and force at the center of the Gaussian, essentially ignoring its shape.

The full wave function is constructed by averaging over the individual wavepackets according to Eq. (4). However, the quantum nature of the wave function is removed by ignoring the overlap of the Gaussian wavepackets in the propagation of the expansion coefficients. For the system considered in this study, where all the dynamics take place on a single electronic state, this makes the expansion coefficients time-independent. This approximation defines the independent multiconfigurational Gaussian (iMCG) method,[37] which describes the evolution of the full wave function by averaging over a swarm of Gaussian wavepackets evolving classically and completely independently on an Ehrenfest potential energy surface.

In iMCG, the expansion coefficients act as weights in the average over the individual wavepacket trajectories and are set at the beginning of a simulation by fitting the full initial wave function. However, properly sampling and fitting the full high-dimensional wave function for the GFP chromophore cluster is challenging, and for simplicity, in this work, we assigned equal weights to each basis Gaussian wavepacket. The wavepackets were all initially centered on the Franck–Condon point. Each wavepacket was given an initial momentum along one normal mode, with an equal number of wavepackets assigned to each mode and symmetrically spaced about zero. This does not give us a true description of the full wave function but provides properties as an average over a swarm of classical trajectories. However, this is sufficient to determine whether proton transfer can still occur at the same rate when the option of tunneling is removed.

D. Calculating time-dependent properties

Properties such as average geometries, or bond lengths, must be obtained from expectation values of the full wave function. The expectation value of a general operator $\hat{O}$ using the vMCG wave function, Eq. (4), is

$$\langle\hat{O}\rangle = \sum_{ij} A_i^* \langle g_i(\mathbf{x}_i)|\hat{O}|g_j(\mathbf{x}_j)\rangle A_j, \tag{20}$$

where $\mathbf{x}_i$ and $\mathbf{x}_j$ are the center coordinates of the GWPs. Properties such as bond lengths, however, cannot be easily written as an operator in coordinate space, particularly as the simulations are performed in normal mode coordinates. However, if the operator is taken at the center of the GWP on the RHS, then

$$\langle\hat{O}\rangle = \sum_j O(\mathbf{x_j}) A_i^* \langle g_i(\mathbf{x}_i)|g_j(\mathbf{x}_j)\rangle A_j + \text{c.c.} \tag{21}$$

$$= \sum_j \mathrm{GGP}_j O(\mathbf{x_j}) + \text{c.c.}, \tag{22}$$

where c.c. is the complex conjugate and we have defined the *gross Gaussian populations* as

$$\mathrm{GGP}_j = \sum_i A_i^* S_{ij} A_j, \tag{23}$$

which are analogous to the Mulliken population analysis used in quantum chemistry calculations.[43] In this way, properties of any

19 February 2024 16:50:08

© Author(s) 2024

function of the molecular geometry can be obtained by knowing the trajectory followed by the centers of the GWPs and the Gaussian overlap matrices.

III. RESULTS

A. Vibrational analysis

It is not computationally feasible to simulate the full quantum dynamics of all 144 degrees of freedom of the cluster model shown in Fig. 1 (x, y, z coordinates of 48 atoms). Rather than focusing on the motion of the proton wire atoms, as has previously been done in Refs. 21 and 23), we treat all 48 atoms in the cluster but only consider motion along vibrational normal mode vectors that we expect to be significantly involved in the ESPT process.

Vibrational frequencies and normal modes were calculated for the cluster model shown in Fig. 1 at the ωB97xd/6-31G* level of theory using Gaussian09.[44] Displacements along the O–H bond coordinates of the proton wire (used as the reaction coordinates by Refs. 21 and 23) are described by a Ser205 O–H bond stretch at 2873 cm^{-1} (mode 116) and symmetric and antisymmetric combinations of the $(O-H)_{HBDI}$ and $(O-H)_{H_2O}$ bond stretches at 3340 cm^{-1} (mode 134) and 3437 cm^{-1} (mode 136), respectively.

The primary GFP absorption at 398 nm is an $n \rightarrow \pi^*$ transition centered on the conjugated ring system of HBDI and does not directly excite the OH stretching modes of the proton wire. The vibrational modes most strongly excited by the electronic absorption are alternating C–C (C=C) bond stretches around the HBDI ring system at 1665 cm^{-1} (mode 109), 1685 cm^{-1} (mode 110), 1742 cm^{-1} (mode 112), and 1750 cm^{-1} (mode 113), as can be seen by the absolute magnitude of the potential energy gradient with respect to each normal mode at the Franck–Condon point (shown in Fig. 2).

The size of the off-diagonal elements of the Hessian at the Franck–Condon point reveals the strength of coupling between different vibrational modes. Figure 2 shows the absolute values of the ten largest coupling values for each of the proton-wire OH stretches and four initially excited C–C stretches identified above. These two sets of vibrations couple to many of the same lower frequency modes, creating a route for absorbed energy to be transferred to the proton wire, triggering ESPT.

Based on this analysis, the 23 vibrational modes in Table I were selected as active degrees of freedom in simulations of the ESPT dynamics. Further analysis of the forces experienced by each normal mode during the course of the simulation (Sec. S4 of the supplementary material) confirms that this set of normal modes captures the most important motions of the chromophore cluster during ESPT.

B. Proton transfer dynamics

The dynamics of the ESPT process were simulated for 2 ps (comparable to the observed experimental timescale), with a basis of 46 Gaussian wavepackets to describe the wave function. This corresponds to two basis functions per vibrational mode and was found to be sufficient to allow each mode to vibrate freely (see Sec. S5 of the supplementary material). The simulation was started by vertically exciting the neutral ground state eigenfunction (a single Gaussian wavepacket centered on the optimized ground state geometry) onto the S1 state with an average momentum of zero. The remaining 45

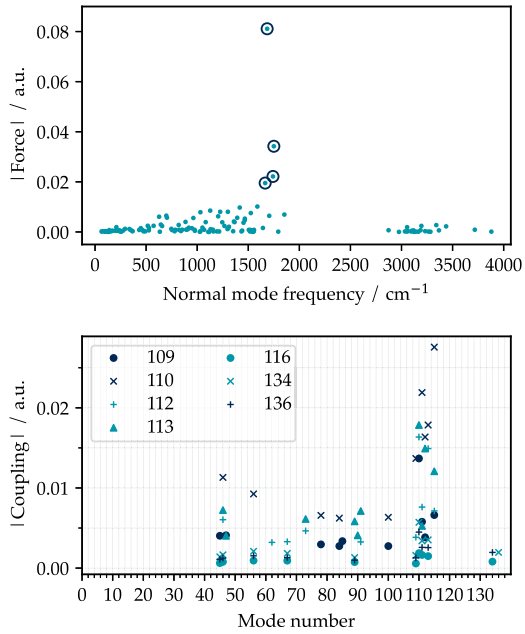


FIG. 2. Top panel: Absolute values of the initial force (negative gradient of the energy with respect to displacement) on each normal mode. Modes 109, 110, 112, and 113 (circled) experience a significantly larger initial force than other modes and are, therefore, selected for inclusion in the simulation. Bottom panel: Absolute values of the coupling (off-diagonal elements of the Hessian) for the ten normal modes most strongly coupled to modes 109, 110, 112, 113, 116, 134, and 136 (the initially excited and reaction coordinate modes). All quantities are reported in atomic units. Displacement along a normal mode is measured in mass-frequency weighted coordinates.

basis wavepackets were given an initial coefficient of zero in the total wave function. The dynamics were assumed to proceed exclusively on the bright S1 state (398 nm absorption), in keeping with previous studies demonstrating that transfer along the proton wire is electronically adiabatic.[8]

As expected, photoexcitation into the S1 state activates a set of strongly coupled vibrational modes, largely centered on the central conjugated π system of the HBDI chromophore. This does

TABLE I. Index and frequencies of the 23 vibrational normal modes included in simulations of the proton wire ESPT dynamics. Visualizations of the selected modes are presented in Sec. S3 of the supplementary material.

Mode	Frequency (cm^{-1})	Mode	Frequency (cm^{-1})	Mode	Frequency (cm^{-1})
45	656	84	1300	111	1705
46	697	85	1303	112	1742
47	704	89	1366	113	1750
56	885	90	1392	115	1853
62	962	91	1416	116	2873
67	1031	100	1525	134	3340
73	1127	109	1665	136	3437
78	1205	110	1685		

19 February 2024 16:50:08

© Author(s) 2024

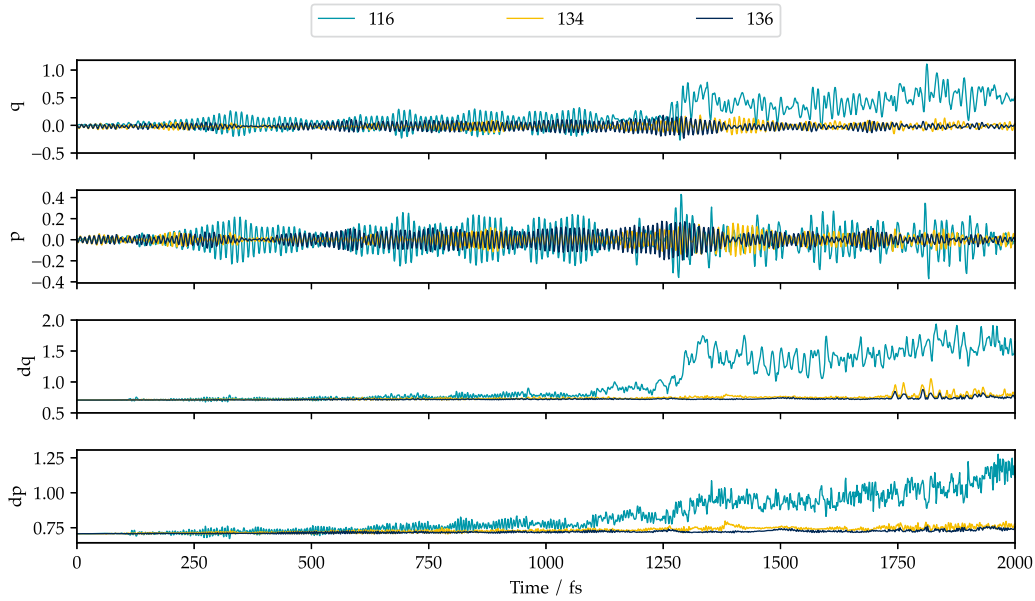


FIG. 3. Top two panels: Expectation value of the displacement (q) and momentum (p) for the three proton transfer modes as a function of time. Mass-frequency weighted coordinates are used to measure the displacement of each normal mode from its ground state-optimized geometry. The expectation value is calculated as a weighted average over the centers of each Gaussian wavepacket basis function, with weights equal to the population associated with the wavepacket (see Sec. II D for details). Lower two panels: Standard deviations of the displacement (dq) and momentum (dp) of the three proton transfer modes.

not include any of the three proton transfer modes, which initially remain completely stationary (although, in reality, they would be expected to have some thermal motion). The oscillation amplitude of the activated modes varies as energy spreads from the initially excited modes to the lower frequency coupled modes identified in Fig. 2 and eventually starts to accumulate in the three proton transfer modes (116, 134, and 136).

The slow filtering of energy into the three proton transfer modes is shown in Fig. 3, in the increasing amplitude of oscillations in their average position and momentum. Figure 4, which shows the time-dependent position of the center of each Gaussian basis wavepacket (in normal mode coordinates), similarly shows increasing motion of the proton transfer modes over time. A closer examination of Fig. 4 indicates that energy is transferred from the initially excited vibrations of the HBDI ring system (modes 109, 110, 112, and 113) to OH stretching modes along the proton wire within ~100 fs. Modes 134 and 136 start vibrating fractionally before mode 116, reflecting their stronger coupling to the lower frequency vibrations of the HBDI side groups (see Fig. 2). Inspection of the normal mode vectors for the vibrations listed in Table I suggests

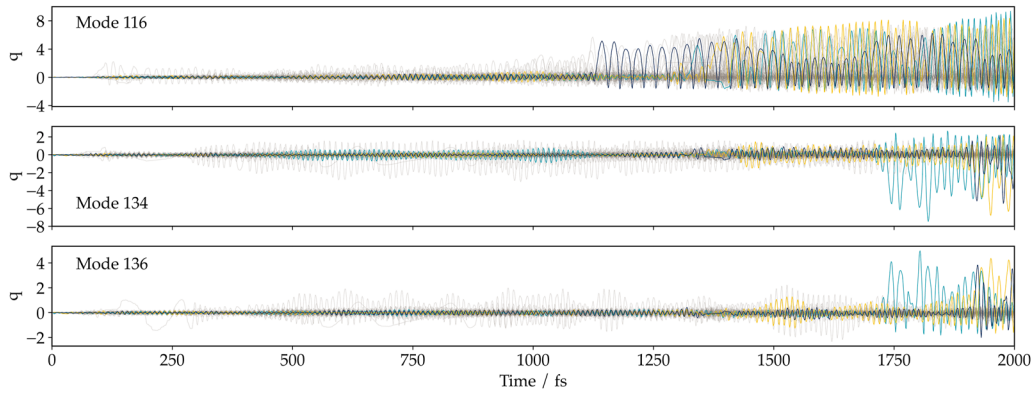


FIG. 4. Trajectories describing the position of the center of each Gaussian wavepacket basis function along the normal mode axes 116, 134, and 136 (the proton transfer modes). To guide the eye, a small number of trajectories with interesting features have been picked out in color. The normal mode coordinates are mass-frequency weighted.

19 February 2024 16:50:08

© Author(s) 2024

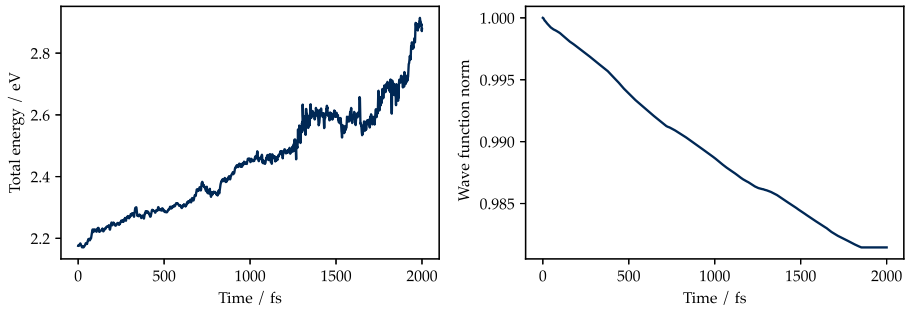


FIG. 5. Change in total system energy and wave function norm over the course of the simulation. Energies are given relative to the energy of the S1 state at the Franck–Condon point. The gradual drift in both values is an indication of numerical errors in the description of the dynamics.

the following mechanism for photoactivation of the OH stretches along the GFP proton wire: photoexcitation from an n to π^* orbital weakens the double bonds in the conjugated ring system of HBDI, causing fluctuations in the C–C bond lengths (i.e., activating alternating C–C/C=C bond stretching modes). This, in turn, activates low frequency ring-breathing modes and wagging motion of the HBDI side groups (including the OH group at the near end of the proton wire). These movements create fluctuations in the interactions between HBDI and other nearby residues (e.g., in the strength of hydrogen bonding between the OH side group and the structural H_2O molecule in the proton wire), ultimately triggering OH stretches along the proton wire.

By construction, the vMCG equations of motion conserve the total energy of the system and the wave function norm, and ideally, these quantities will remain constant during a simulation even as energy is redistributed between different vibrational modes. However, the error introduced by using the local harmonic approximation to obtain the Hamiltonian matrix elements, combined with the limited accuracy of the numerical integrator, introduces small errors, leading to a gradual upward drift in the total energy and a corresponding decrease in the norm of the wave function throughout the simulation. Figure 5 shows the trend in total energy and wave function over the 2 ps of the simulation. The clear drift in both quantities indicates that the description of the dynamics becomes less reliable at later times. This energy and norm drift are, however, spread over the 46 GWPs of the simulation, which means that this is a relatively small overall change in total energy (approximately less than 0.01 eV/GWP) and norm (less than 2% overall) from start to finish, offering reassurance that the qualitative behavior of the system will be correctly described.

Although the $(O{-}H)_{Ser205}$ bond is (marginally) the slowest of the OH bonds in the proton wire to be affected by photoexcitation, it quickly starts vibrating with a noticeably greater amplitude than the rest of the proton wire. Moreover, a clear jump in the average displacement of mode 116 is observed around 1250 fs (Fig. 3), indicating transfer of the H_{Ser205} proton toward Glu222. This transfer can also clearly be seen as an increase in the average $(O{-}H)_{Ser205}$ bond length by 0.1 Å (Fig. 6). Since the distance between O_{Ser205} and

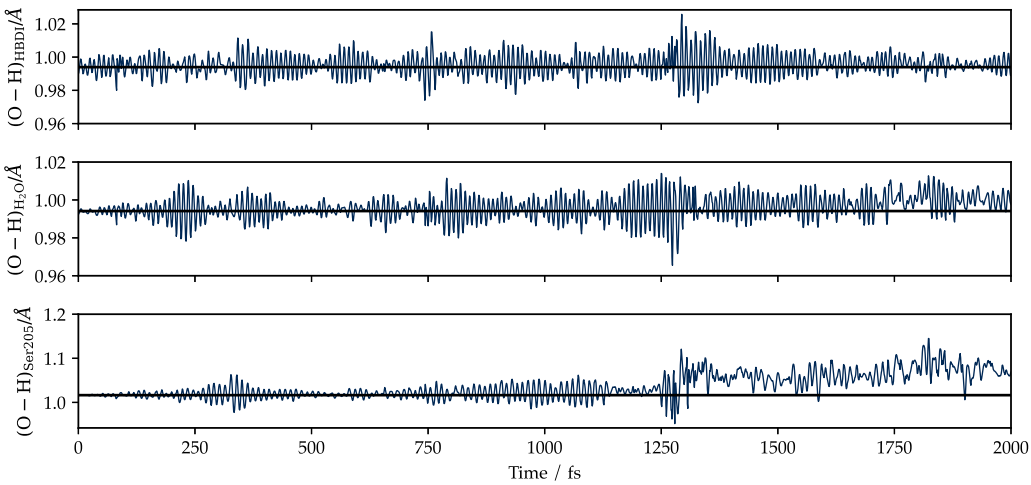


FIG. 6. Change in the average length of the three OH bonds along the proton wire over time. The average bond length is calculated as an average over the bond length explored by each of the 46 basis Gaussian wavepackets, weighted by the population of the wavepacket. Populations are calculated from the density in an analogous way to a Mulliken population analysis (see Sec. II D for details).

19 February 2024 16:50:08

© Author(s) 2024

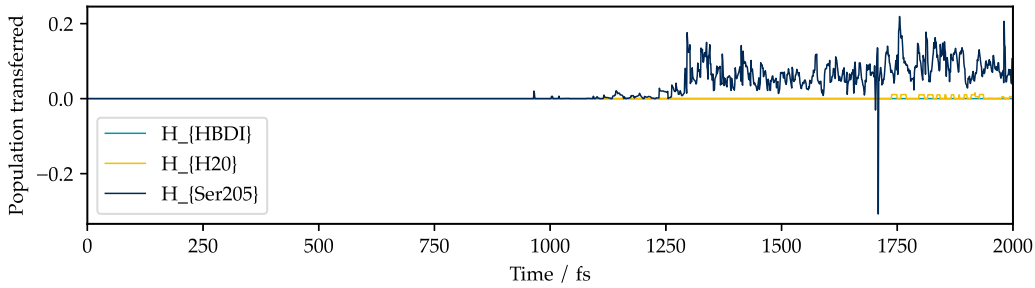


FIG. 7. Total fractional populations of Gaussian wavepacket trajectories in which a proton transfer has taken place, as a function of time. The proton is considered to have transferred if it sits more than halfway between the oxygen atom to which it is initially bonded and the oxygen atom to which it is transferring. Populations are calculated from the density in a manner analogous to Mulliken populations. We note that negative populations are a well-known artifact that can arise from Mulliken population analysis and, in this case, have led to an apparently negative percentage of population transfer at around 1700 fs. No physical significance should be attached to this.

O_{Glu222} is 2.6 Å, the proton transfer is clearly incomplete (a full transfer would be expected to increase the $(O-H)_{Ser205}$ bond length by ~0.6 Å). However, note that the average bond length reflects an average over parts of the wave function that explore the full transfer of H_{Ser205} and parts where H_{Ser205} remains very closely associated with O_{Ser205}, as can be seen from the trajectories of the individual Gaussian basis functions in Fig. 4. Therefore, the small increase in average bond length should not be taken as an indication that H_{Ser205} never explores larger displacements from its original position.

No equivalent jump in average displacement is seen for modes 134 and 136 within the 2 ps timescale, confirming the suggestion of previous studies that the ESPT process is asynchronous and starts from the "far end" of the proton wire with the transfer of H_{Ser205} from Ser_{205} to Glu_{222}. In other words, the GFP proton wire acts as a photobase, rather than a photoacid, with the proton transfers being driven by abstraction of a proton by the Glu_{222} anion rather than proton donation by the photoexcited HBDI.

The start of a second proton transfer is observed in the last 250 fs of the simulation. Although no significant change is seen in the average displacement of modes 134 and 136, a sharp increase in the variation of the displacement (dq) and momentum (dp) is observed for these modes around 1750 fs (Fig. 3). Figure 4 shows that this corresponds to significant displacement of modes 134 and 136 in a small number of individual (highlighted) Gaussian wavepacket

19 February 2024 16:50:08

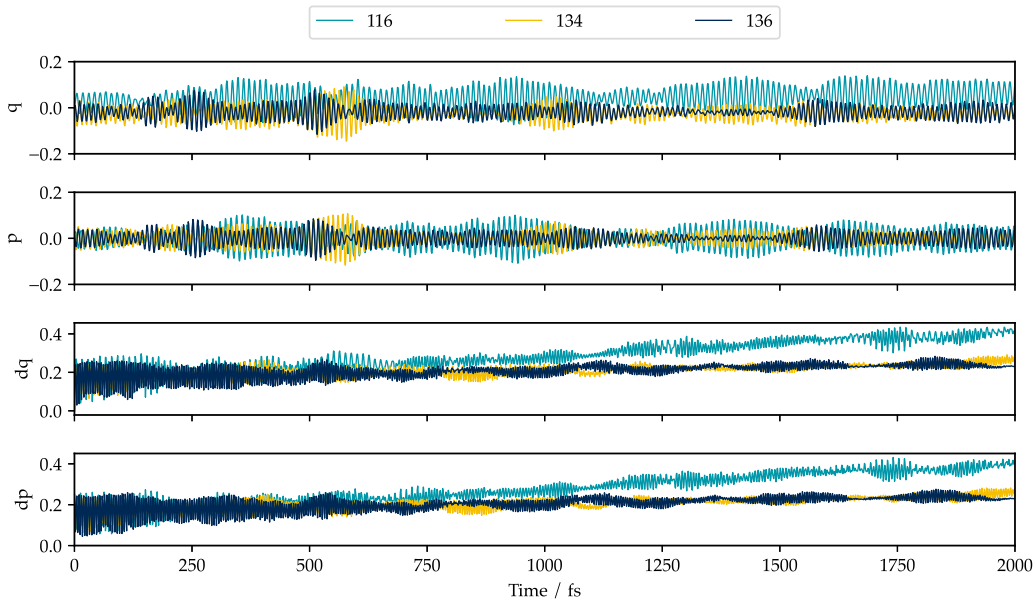


FIG. 8. Top two panels: Expectation value of the displacement (q) and momentum (p) for the three proton transfer modes over the course of a set of 2 ps classical trajectories. Mass-frequency weighted coordinates are used to measure the displacement of each normal mode from its ground state-optimized geometry. The expectation value is calculated as an (unweighted) average over the centers of each Gaussian wavepacket basis function. Lower two panels: Standard deviation of the average displacement (dq) and momentum (dp) of the three proton transfer modes.

© Author(s) 2024

trajectories. Since modes 134 and 136 are, respectively, symmetric and asymmetric combinations of the $(O-H)_{HBDI}$ and $(O-H)_{H_2O}$ stretches, the combination of modes 136 and 134 corresponds (primarily) to the $(O-H)_{H_2O}$ stretch (transfer of H_{H_2O}) and, indeed, a small increase is seen in the average length of the $(O-H)_{H_2O}$ bond at around 1750 fs (Fig. 6). We anticipate that a fuller transfer of H_{H_2O} would occur in the next few hundred femtoseconds, following the pattern of the earlier H_{Ser205} transfer.

The proton transfer observed at 1250 fs is not complete. A number of the individual Gaussian wavepacket trajectories plotted in Fig. 4 only explore small displacements of mode 116, meaning that some proportion of the total wave function describes a structure where H_{Ser205} remains closely associated with O_{Ser205}. That the trajectories with no H_{Ser205} transfer are significantly populated is evidenced by the fact that, at any one time, less than 20% of the population is exploring a conformation with the proton transferred (see Fig. 7). Consequently, the average displacement of mode 116, shown in Fig. 3, only reaches a fraction of the maximum displacements observed for individual trajectories in Fig. 4. (Note that the displacements reported in these figures are in mass-frequency weighted coordinates.) Furthermore, those trajectories that do explore the H_{Ser205} transfer show the proton hopping continuously back and forth between Ser_{205} and Glu_{222}, rather than remaining closely associated with Glu_{222} (the minimum displacement of mode 116 in the oscillating trajectories in Fig. 4 does not noticeably increase). The fact that the first proton transfer is still in progress when the second proton transfer begins supports the hypothesis that the overall ESPT process has a concerted (rather than stepwise) asynchronous mechanism.

No proton transfer is observed within 2 ps for swarms of classically evolving Gaussian wavepackets. Figure 8 shows the evolution of the average position and momentum of a swarm of 92 trajectories (to test convergence, a smaller swarm of 46 trajectories with the same initial momentum and position as the vMCG trajectories was also considered but showed no appreciable difference). Both quantities oscillate around their average with no sign of any increase, sudden or gradual, that characterized proton transfer in the quantum trajectories shown in Fig. 3. Not only is there no increase in the average displacement of any of the three proton transfer modes but larger displacements are not explored in any of the individual Gaussian wavepacket trajectories (see Sec. S6 of the supplementary material). Toward the end of the 2 ps simulation, a small increase is observed in the variation of the average position and momentum of mode 116, which is matched by a very small increase in the average $(O-H)_{Ser205}$ bond length (Fig. 9). However, this is a significantly smaller increase in bond length than seen in the quantum trajectory (Fig. 6), indicating that quantum tunneling is essential to allow proton transfer to take place within the experimentally observed timescale.

C. Potential energy surfaces

The potential energy landscape for OH bond stretches along the proton wire fluctuates continuously with the changing structure of the rest of the system. However, certain cuts through the full 23-dimensional potential energy surface offer some insight into why the excited state proton transfer proceeds in the manner described above.

Figures 10(a) and 10(b) show 2D slices through the potential energy surface along modes 116 and 134 or 136, respectively. The displacements of all the other vibrational modes are fixed at their time-averaged values (values given in Sec. S3 of the supplementary material). The surface is calculated using the Shepard interpolation between database points, as described in Sec. II B 2. While stretching (and certainly breaking) any of the three OH bonds in the proton wire has an associated energy cost, it is clear from the elongated minima of the potential energy surfaces that stretching the $(O-H)_{Ser205}$ bond (positive displacement along mode 116) is significantly easier than stretching the other two OH bonds. This is likely because of the electrostatic attraction between H_{Ser205} and the Glu222 anion. Thus, the proton transfer at the "far end" of the proton wire is the first to commence.

Figures 10(c) and 10(d) show a comparison of the 2D slices through the potential energy surface along modes 134 and 136 when the displacement of mode 116 is fixed at either 0.19 units (its time-averaged value) or 8 units (the approximate maximum value explored by individual trajectories). The displacements of all other modes are again set to their time-averaged values. It is apparent from these slices that the transfer of H_{Ser205} significantly reduces the energy barrier for the remaining two proton transfers (indeed,

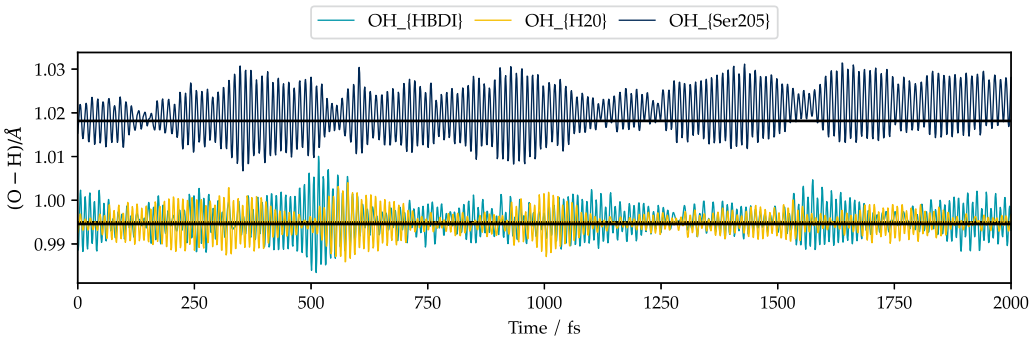


FIG. 9. Change in the average length of the three OH bonds along the proton wire over the course of a set of 2 ps classical trajectories. The average bond length is calculated as an (unweighted) average over the bond length explored by each of the 92 basis Gaussian wavepackets.

19 February 2024 16:50:08

© Author(s) 2024

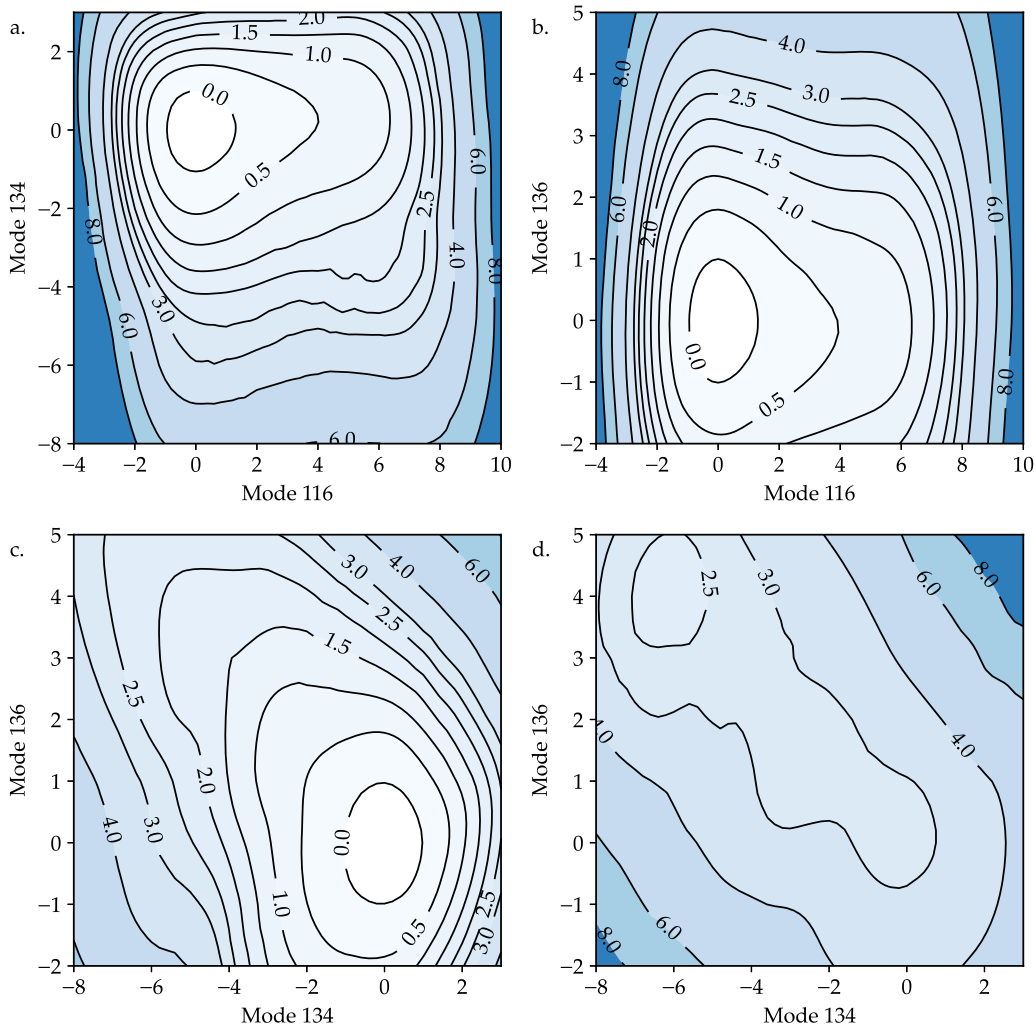


FIG. 10. 2D slices of the 23-dimensional potential energy surface. (a) Slice along modes 116 and 134 with the displacement of all other modes fixed at their time-averaged values (see Sec. S3 of the supplementary material). (b) Slice along modes 116 and 136 with the displacement of all other modes fixed at their time-averaged values. (c) Slice along modes 134 and 136 with the displacement of all other modes fixed at their time-averaged values. (d) Slice along modes 134 and 136 with mode 116 displaced by 8 units and the displacement of all other modes fixed at their time-averaged values. Displacements along each normal mode are given in mass-frequency weighted coordinates. Each contour line is labeled with a value for the potential energy in eV, relative to the energy of the photoexcited S1 state at the Franck–Condon point.

in the case of a complete H_{Ser205} transfer, the second proton transfer would become energetically downhill). Furthermore, the lowest energy pathway is the transfer of H_{H_2O} alone (simultaneous displacement of mode 134 in the negative direction and mode 136 in the positive direction). There is a slight widening of the minimum around the top left corner of the potential energy surface shown in Fig. 10(d). This suggests that once H_{H_2O} has (partially) transferred, it may become easier to displace mode 136 [the symmetric combination of $(O–H)_{HBDI}$ and $(O–H)_{H_2O}$ stretches], making it possible for the final proton transfer (transfer of H_{HBDI}) to occur.

The shape of the potential energy surface helps explain why tunneling might be necessary to achieve proton transfer within a reasonable timeframe. The energetic cost of stretching any of the OH bonds along the proton wire creates a large activation barrier to the proton transfer process that is difficult to overcome classically. However, a quantum wave function is able to explore regions of the potential energy surface that are classically inaccessible, allowing it to sample the downhill gradient created by transferring H_{Ser205} (and the presumed minimum associated with transfer of all three protons) and tunnel in this direction.

The potential energy surfaces shown in Fig. 10 are constructed by interpolating between the energies of the structures stored in the database built up during the simulation (see Sec. II B 2). The transfer of H_{HBDI} was not observed during the 2 ps simulation, so the database does not contain information about this part of the potential energy surface. Therefore, this simulation does not

19 February 2024 16:50:08

© Author(s) 2024

provide a picture of the potential minimum associated with the product species where all three proton transfers have occurred. However, we note that those parts of the potential energy surface explored by our simulation also do not show any local minima associated with intermediate species where the first or second proton transfers have occurred. This is in keeping with experimental studies, suggesting that there is no observable intermediate for the ESPT process.

IV. DISCUSSION

By selecting a subset of vibrational normal modes strongly coupled to the electronic photoexcitation and proton transfer reaction coordinate of *wt*-GFP, we have been able to perform a fully quantum simulation of the nuclear dynamics of excited state proton transfer on a higher dimensional potential than has previously been used. The importance of considering more than the six degrees of freedom describing the relative position of the atoms in the proton wire has been evident from the failure of previous theoretical studies to reproduce the experimentally measured ESPT timescale. By contrast, we observe partial transfer of two of the three protons along the wire within 2 ps, consistent with the 2–3 ps ESPT timescale measured using ultrafast electronic transient absorption spectroscopy.[9] This good agreement with the experimental timescale indicates that our model is sufficiently large to provide a qualitatively accurate description of the mechanism of excited state proton transfer, clarifying a number of key mechanistic details.

First, we observe that ESPT proceeds via a concerted, asynchronous mechanism analogous to that for ground state proton transfer:[13,14] the three proton transfers all begin at different times, but the first transfer is not complete before subsequent transfers begin. This is in contrast to the stepwise, asynchronous[16] and concerted, synchronous[20,21] mechanisms that appear favorable when only energetic changes along the O–H coordinates of the proton wire are considered, emphasizing the importance of the wider system in shaping the reactive pathway through the potential energy surface. We note that Nadal-Ferret *et al.* previously demonstrated that proton transfer could not occur without structural relaxation of the chromophore and residues surrounding the proton wire.[23] Based on embedded QM/MM energies of the stretched O–H bonds along the proton wire (with all other atoms frozen in the optimized ground-state geometry), they predicted that allowing for a wider structural relaxation would reveal the ESPT mechanism to be concerted and asynchronous. Our result confirms their prediction.

Deprotonation of HBDI via the proton wire is initiated by proton transfer from Ser205 to Glu222 (at the far end of the wire). Based on 2D slices through the full-dimensional potential energy surface, we observe, in agreement with a number of previous studies,[11,18,19,21,22] that $(O-H)_{Ser205}$ is by far the easiest part of the wire to extend (stabilized by its proximity to the Glu222 anion) and that the transfer of H_{Ser205} substantially reduces the energetic cost of subsequent proton transfers. The transfer of H_{Ser205} as the rate-limiting step is consistent with transient infrared measurements, showing a strong correlation between ESPT rate and the conformation of the Glu222 acceptor.[2,18]

The 2D PES slices analyzed in this study are based on a time-averaged structure for the bulk of the chromophore system. They do not illustrate the time-dependent availability of different pathways through the full-dimensional PES (which dictates the dynamical behavior of the system) and, therefore, give a limited view of the energetic cost of each individual proton transfer.

In particular, this selective view of the PES does not straightforwardly reveal the size of the energetic barrier for the reaction. However, we note that stretching and breaking the $(O–H)_{Ser205}$ bond (displacing mode 116) at the start of the proton transfer takes on the order of 1 eV of energy (see Fig. 10). This is an important observation in light of the upward drift in total energy caused by the buildup of small numerical errors across the simulation. Since the energy drift over the whole simulation only corresponds to ~0.01 eV of additional energy per GWP (see Fig. 5), it is unlikely to have artificially triggered the proton transfer reaction. Furthermore, the continuous upward slope of the potential surface along mode 116 means that, classically, additional energy should result in a smooth increase in the $(O–H)_{Ser205}$ bond length. Instead, we observe a sudden increase in $(O–H)_{Ser205}$ bond length around 1250 ps that is not preceded by any sudden increase in the total energy, offering some reassurance that the observed proton transfer dynamics are not an artifact of the energy drift.

However, the quantitative accuracy of our potential energy surfaces could be improved by further extensions to the model. Most significantly, the potential energy surface can be sensitive to the configuration of the surrounding protein environment. In this study, we allow motion of only 23 normal modes out of the 136 vibrational degrees of freedom of the chromophore cluster model. The surrounding protein environment is not considered beyond the selected side-chain residues involved in the proton wire. Inspection of the forces on the normal modes at each geometry explored in the direct dynamics database shows that the frozen modes in our simulation are almost never subject to large forces that could trigger significant motion (see Sec. S4 of the supplementary material). This confirms our assertion that the degrees of freedom included in our simulation are those that are most important in describing the ESPT dynamics. Nevertheless, additionally considering the motion of the frozen modes, even at the level of a classical force field, would likely influence the ESPT timescale, both through small changes to the potential energy surface and by providing additional routes for energy dissipation. Slightly larger forces on a small handful of frozen normal modes in a small number of database structures also indicate that considering additional modes (or a different selection of modes) may be necessary for exploring the dynamics of the proton wire at longer times. Previous QM/MM studies of the GFP proton wire have, moreover, demonstrated that the potential energy surface can be sensitive to the configuration of the surrounding protein environment.[22,23] In particular, the proximity of a nearby histidine residue (His148) to O_{HBDI} may play an important role in triggering a favorable conformation for proton transfer.[22] However, the added cost of increasing the system size or including an atomistic environment makes it unfeasible at the present time to explore the full quantum dynamics of the proton wire plus its surroundings over the timescale of proton transfer. While our results give a good overall picture of the proton transfer mechanism, alternative QM/MM approaches describing the full system and environment may be suitable for exploring nuanced structural details at critical stages of the proton transfer.

Within a cell, the GFP proton wire operates not only within a complex environment but also at body temperature. By contrast,

19 February 2024 16:50:08

© Author(s) 2024

our simulation is effectively carried out at low temperatures in the sense of propagating a single wave function, rather than a density matrix of thermally populated states. However, we do not expect this to have a significant effect. The vibrational states of the lowest frequency vibrational normal mode included in our simulation will have a spacing of 656 cm^{-1}, already several times the Boltzmann energy at 310 K (37 °C). Therefore, the population of vibrational states higher than $v = 0$ will be very small and should not noticeably affect the dynamics.

An important outstanding question around the GFP ESPT mechanism is the origin of its two fluorescence lifetimes. Suggested explanations have included a two-stage process with the formation of an intermediate species containing a long $O_{Ser205} - H_{Ser205} - O_{Glu222}$ hydrogen bond, or different ESPT timescales associated with two different structures around the Glu222 end of the proton wire. Intermediate formation is generally considered to be a less likely explanation since a variety of experiments have failed to detect any distinct intermediate species.[2,4,10,15] Our results similarly do not show any evidence of an intermediate species forming. The transfer of H_{Ser205} observed in the second picosecond of our simulation is incomplete, giving an average picture of a slightly lengthened $(O - H)_{Ser205}$ bond, similar to that postulated by Di Donato *et al.*[10] However, this intermediate bond length is a weighted average over individual wavepacket trajectories that show either no proton transfer or full transfer of H_{Ser205} back and forth between O_{Ser205} and O_{Glu222}. This suggests that the wave functions explore a wide region of the geometric space, with some probability of finding H_{Ser205} in most locations between O_{Ser205} and O_{Glu222} (but strongly weighted toward finding it in close association with one of the two oxygens). This is different to forming an identifiable intermediate, where the wave function would be expected to describe a narrower distribution around an intermediate H_{Ser205} position. Furthermore, the fact that signs of an imminent H_{H_2O} transfer emerge at close to 2 ps appears inconsistent with the suggestion that the longer 10 ps timescale is required to complete the final two proton transfers. However, simulating the dynamics of the proton wire for a full 10 ps is not feasible at the level of theory used in this study, in terms of both computational cost and the degree of numerical error that would accumulate over a longer simulation. For this reason, we also do not attempt to explore the dynamics of alternative proton wire structures thought to be associated with slower ESPT.

The fact that not even partial proton transfer is observed within our 2 ps classical trajectories is strong evidence that the proton wire relies on tunneling to operate. However, the static 2D slices that we obtain through the PES do not allow us to calculate accurate barrier heights. For this reason, we do not attempt to calculate the tunneling rate or quantitatively rationalize the experimentally observed kinetic isotope effect. Characterizing the full, dynamic PES associated with the H_{Ser205} transfer, for both wild-type and mutant species of GFP, would be extremely interesting for understanding the potential role of Ser205 in controlling or gating the ESPT process.

Furthermore, as we do not observe transfer of H_{H_2O} or H_{HBDI} within the timeframe of our simulation, we cannot determine whether tunneling is important for all three proton transfers or, as previously suggested,[8] only for the transfer of H_{Ser205}. However, while H_{Ser205} transfer is an energetically uphill process, our calculated slices through the PES suggest that subsequent transfers may be entirely downhill. Therefore, we anticipate that tunneling is likely to play a much smaller, even negligible, role in promoting H_{H_2O} and H_{HBDI} transfers.

V. CONCLUSIONS

Using direct quantum dynamics, we have confirmed that excited state proton transfer along the proton wire in GFP proceeds in a concerted, asynchronous fashion, led by abstraction of a proton from Ser205 by the Glu222 anion. There is a significant activation barrier to deprotonating Ser205, and tunneling through this barrier plays a critical role in allowing the proton transfer to proceed at a reasonable rate. Deprotonating Ser205 makes subsequent proton transfers along the wire energetically more favorable—possibly even barrierless.

Direct dynamics means that the nuclei move over the anharmonic potential surfaces of the excited state as given by the level of quantum chemistry selected with no restrictions. However, the 2D slices through the PES obtained in this study do not fully capture the complexity of the time-evolving potential as the model used just included the proton wire and ignored the protein and solvent environment. While this minimal model shows that the proton wire alone can undergo proton transfer if the quantum nature of the nuclei is taken into account, a more detailed analysis of the pathway taken through the full-dimensional PES is required to quantitatively characterize the kinetics of ESPT and understand how these are controlled by the structure of the proton wire and vibrational motion of the wider system. In addition to using a bigger basis set to improve the quality of the dynamics simulation, the role of specific residues, such as histidine 128, and re-organization of the protein environment should be examined, as they are known to enhance the proton transfer.[22] For simulations of this size, further improvements are required of the DD-vMCG method to increase the efficiency in order to treat the larger systems. This includes more efficient and stable integration schemes and a more efficient handling of the quantum chemistry data providing the surfaces. This work is presently underway.

SUPPLEMENTARY MATERIAL

See the supplementary material for benchmarking data for the electronic structure methods, optimized structure of the GFP cluster model, characterization of the selected normal modes included in the dynamics, initial forces on normal modes, benchmarking of the number of Gaussian wavepackets required for dynamics, GWP trajectories from classical simulations, and input and output files for the geometry optimization and dynamics calculations, including the full database of geometries explored by the dynamics.

ACKNOWLEDGMENTS

The authors would like to thank Thomas Ashley and Iakov Polyak for initial exploratory calculations characterizing the electronic structure of the cluster model. This research was funded by the EPSRC as part of the program "Ultrafast Photochemical Dynamics in Complex Environment" (Grant No. EP/V026690/1). The APC was funded by University College London.

19 February 2024 16:50:08

© Author(s) 2024

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Susannah Bourne-Worster: Conceptualization (equal); Data curation (lead); Formal analysis (lead); Investigation (lead); Methodology (equal); Project administration (lead); Validation (equal); Visualization (lead); Writing – original draft (lead); Writing – review & editing (lead). **Graham A. Worth**: Conceptualization (equal); Funding acquisition (lead); Methodology (equal); Resources (lead); Supervision (lead); Writing – review & editing (supporting).

DATA AVAILABILITY

The files (input files, output files, and DD-vMCG database) for the calculations in this study are openly available in UCL Research Data Repository at https://rdr.ucl.ac.uk at http://doi.org/10.5522/04/24190683. Additional details, validation, and analysis of the data that support the findings of this study are available in the supplementary material.

REFERENCES

[1] R. Y. Tsien, "The green fluorescent protein," Annu. Rev. Biochem. **67**(1), 509–544 (1998).

[2] J. J. van Thor and P. M. Champion, "Photoacid dynamics in the green fluorescent protein," Annu. Rev. Phys. Chem. **74**(1), 123–144 (2023).

[3] K. Brejc, T. K. Sixma, P. A. Kitts, S. R. Kain, R. Y. Tsien, M. Ormö, and S. James Remington, "Structural basis for dual excitation and photoisomerization of the *Aequorea victoria* green fluorescent protein," Proc. Natl. Acad. Sci. U. S. A. **94**(6), 2306–2311 (1997).

[4] M. Chattoraj, B. A. King, G. U. Bublitz, and S. G. Boxer, "Ultra-fast excited state dynamics in green fluorescent protein: Multiple states and proton transfer," Proc. Natl. Acad. Sci. U. S. A. **93**(16), 8362–8367 (1996).

[5] S. R. Meech, "Excited state reactions in fluorescent proteins," Chem. Soc. Rev. **38**(10), 2922 (2009).

[6] H. Morise, O. Shimomura, F. H. Johnson, and J. Winant, "Intermolecular energy transfer in the bioluminescent system of *Aequorea*," Biochemistry **13**(12), 2656–2662 (1974).

[7] P. Armengol, R. Gelabert, M. Moreno, and J. M. Lluch, "Unveiling how an archetypal fluorescent protein operates: Theoretical perspective on the ultrafast excited state dynamics of GFP variant S65T/H148D," J. Phys. Chem. B **119**(6), 2274–2291 (2015).

[8] B. Salna, A. Benabbas, J. T. Sage, J. van Thor, and P. M. Champion, "Wide-dynamic-range kinetic investigations of deep proton tunnelling in proteins," Nat. Chem. **8**(9), 874–880 (2016).

[9] J. T. M. Kennis, D. S. Larsen, I. H. M. van Stokkum, M. Vengris, J. J. van Thor, and R. van Grondelle, "Uncovering the hidden ground state of green fluorescent protein," Proc. Natl. Acad. Sci. U. S. A. **101**(52), 17988–17993 (2004).

[10] M. Di Donato, L. J. G. W. van Wilderen, I. H. M. Van Stokkum, T. C. Stuart, J. T. M. Kennis, K. J. Hellingwerf, R. van Grondelle, and M. L. Groot, "Proton transfer events in GFP," Phys. Chem. Chem. Phys. **13**(36), 16295 (2011).

[11] J. J. van Thor, K. L. Ronayne, M. Towrie, and J. T. Sage, "Balance between ultrafast parallel reactions in the green fluorescent protein has a structural origin," Biophys. J. **95**(4), 1902–1912 (2008).

[12] B. L. Grigorenko, A. V. Nemukhin, I. V. Polyakov, D. I. Morozov, and A. I. Krylov, "First-principles characterization of the energy landscape and optical spectra of green fluorescent protein along the A → I → B proton transfer route," J. Am. Chem. Soc. **135**(31), 11541–11549 (2013).

[13] S. Wang and S. C. Smith, "Mechanistic aspects of proton chain transfer in the green fluorescent protein: Part II. A comparison of minimal quantum chemical models," Phys. Chem. Chem. Phys. **9**(4), 452–458 (2007).

[14] R. Zhang, M. T. Nguyen, and A. Ceulemans, "A concerted mechanism of proton transfer in green fluorescent protein. A theoretical study," Chem. Phys. Lett. **404**(4-6), 250–256 (2005).

[15] T. Fujisawa, H. Kuramochi, H. Hosoi, S. Takeuchi, and T. Tahara, "Role of coherent low-frequency motion in excited-state proton transfer of green fluorescent protein studied by time-resolved impulsive stimulated Raman spectroscopy," J. Am. Chem. Soc. **138**(12), 3942–3945 (2016).

[16] J. J. van Thor, "Photoreactions and dynamics of the green fluorescent protein," Chem. Soc. Rev. **38**(10), 2935 (2009).

[17] D. Stoner-Ma, A. A. Jaye, P. Matousek, M. Towrie, S. R. Meech, and P. J. Tonge, "Observation of excited-state proton transfer in green fluorescent protein using ultrafast vibrational spectroscopy," J. Am. Chem. Soc. **127**(9), 2864–2865 (2005).

[18] J. J. van Thor, G. Zanetti, K. L. Ronayne, and M. Towrie, "Structural events in the photocycle of green fluorescent protein," J. Phys. Chem. B **109**, 16099–16108 (2005).

[19] O. Vendrell, R. Gelabert, M. Moreno, and J. M. Lluch, "A potential energy function for heterogeneous proton-wires. Ground and photoactive states of the proton-wire in the green fluorescent protein," J. Chem. Theory Comput. **4**(7), 1138–1150 (2008).

[20] L. Zhang, D. Silva, Y. Yan, and X. Huang, "Force field development for cofactors in the photosystem II," J. Comput. Chem. **33**(25), 1969–1980 (2012).

[21] O. Vendrell, R. Gelabert, M. Moreno, and J. M. Lluch, "Operation of the proton wire in green fluorescent protein. A quantum dynamics simulation," J. Phys. Chem. B **112**(17), 5500–5511 (2008).

[22] G. Donati, A. Petrone, P. Caruso, and N. Rega, "The mechanism of a green fluorescent protein proton shuttle unveiled in the time-resolved frequency domain by excited state *ab initio* dynamics," Chem. Sci. **9**(5), 1126–1135 (2018).

[23] M. Nadal-Ferret, R. Gelabert, M. Moreno, and J. M. Lluch, "Transient low-barrier hydrogen bond in the photoactive state of green fluorescent protein," Phys. Chem. Chem. Phys. **17**(46), 30876–30888 (2015).

[24] L. Masgrau, A. Roujeinikova, L. O. Johannissen, P. Hothi, J. Basran, K. E. Ranaghan, A. J. Mulholland, M. J. Sutcliffe, N. S. Scrutton, and D. Leys, "Atomic description of an enzyme reaction dominated by proton tunneling," Science **312**(5771), 237–241 (2006).

[25] M. . Wikström, V. Sharma, V. R. I. Kaila, J. P. Hosler, and G. Hummer, "New perspectives on proton pumping in cellular respiration," Chem. Rev. **115**(5), 2196–2221 (2015).

[26] M. Y. Okamura, M. L. Paddock, M. Graige, and G. Feher, "Proton and electron transfer in bacterial reaction centers," Biochim. Biophys. Acta, Bioenerg. **1458**(1), 148–163 (2000).

[27] P. Leiderman, R. Gepshtein, I. Tsimberov, and D. Huppert, "Effect of temperature on excited-state proton tunneling in wt-green fluorescent protein," J. Phys. Chem. B **112**(4), 1232–1239 (2008).

[28] A. Petrone, P. Cimino, G. Donati, H. P. Hratchian, M. J. Frisch, and N. Rega, "On the driving force of the excited-state proton shuttle in the green fluorescent protein: A time-dependent density functional theory (TD-DFT) study of the intrinsic reaction path," J. Chem. Theory Comput. **12**, 4925–4933 (2016).

[29] H. Zhang, Q. Sun, Z. Li, S. Nanbu, and S. S. Smith, "First principle study of proton transfer in the green fluorescent protein (GFP): Ab initio PES in a cluster model," Comput. Theor. Chem. **990**, 185–193 (2012).

[30] G. A. Worth, "QUANTICS: A general purpose package for quantum molecular dynamics simulations," Comput. Phys. Commun. **248**, 107040 (2020).

[31] G. A. Worth, K. Giri, G. W. Richings, M. H. Beck, A. Jäckle, and H.-D. Meyer, QUANTICS Package, Version 2.1, 2021.

[32] G. W. Richings, I. Polyak, K. E. Spinlove, G. A. Worth, I. Burghardt, and B. Lasorne, "Quantum dynamics simulations using Gaussian wavepackets: The vMCG method," Int. Rev. Phys. Chem. **34**(2), 269–308 (2015).

[33] G. Christopoulou, A. Freibert, and G. A. Worth, "Improved algorithm for the direct dynamics variational multi-configurational Gaussian method," J. Chem. Phys. **154**(12), 124127 (2021).

19 February 2024 16:50:08

© Author(s) 2024

[34]G. A. Worth, M. A. Robb, and I. Burghardt, "A novel algorithm for non-adiabatic direct dynamics using variational Gaussian wavepackets," Faraday Discuss. **127**, 307 (2004).
[35]G. W. Richings and G. A. Worth, "Multi-state non-adiabatic direct-dynamics on propagated diabatic potential energy surfaces," Chem. Phys. Lett. **683**, 606–612 (2017).
[36]G. A. Worth and L. S. Cederbaum, "Beyond Born-Oppenheimer: Molecular dynamics through a conical intersection," Annu. Rev. Phys. Chem. **55**(1), 127–158 (2004).
[37]K. E. Spinlove, M. Vacher, M. Bearpark, M. A. Robb, and G. A. Worth, "Using quantum dynamics simulations to follow the competition between charge migration and charge transfer in polyatomic molecules," Chem. Phys. **482**, 52–63 (2017).
[38]I. Burghardt, H.-D. Meyer, and L. S. Cederbaum, "Approaches to the approximate treatment of complex molecular systems by the multiconfiguration time-dependent Hartree method," J. Chem. Phys. **111**(7), 2927–2939 (1999).
[39]K. C. Thompson, M. J. T. Jordan, and M. A. Collins, "Polyatomic molecular potential energy surfaces by interpolation in local internal coordinates," J. Chem. Phys. **108**, 8302–8316 (1998).
[40]T. J. Frankcombe, "Using hessian update formulae to construct modified Shepard interpolated potential energy surfaces: Application to vibrating surface atoms," J. Chem. Phys. **140**, 114106–114108 (2014).
[41]C. Fang and B. Durbeej, "Calculation of free-energy barriers with TD-DFT: A case study on excited-state proton transfer in indigo," J. Phys. Chem. A **123**(40), 8485–8495 (2019).
[42]J.-D. Chai and M. Head-Gordon, "Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections," Phys. Chem. Chem. Phys. **10**(44), 6615 (2008).
[43]C. S. M. Allan, B. Lasorne, G. A. Worth, and M. A. Robb, "A straightforward method of analysis for direct quantum dynamics: Application to the photochemistry of a model cyanine," J. Phys. Chem. A **114**, 8713–8729 (2010).
[44]M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and J. D. Fox, Gaussian09, 2013.

19 February 2024 16:50:08

© Author(s) 2024