

Effects of enzyme-ligand interactions on the
photoisomerization of a light-regulated
chemotherapeutic drug

Ruibin Liang and Amirhossein Bakhtiari*

Department of Chemistry and Biochemistry, Texas Tech University, Lubbock, TX, 79409, USA

AUTHOR INFORMATION

Corresponding Author

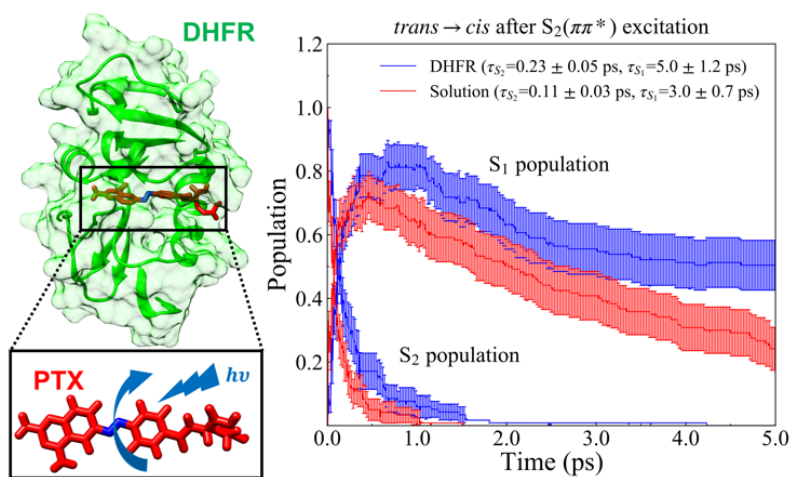
*Ruibin Liang

Email address: rliang@ttu.edu

Abstract

Molecular photoswitches permit using light to control protein activity with high spatiotemporal resolutions, thereby alleviating the side effects of conventional chemotherapy. However, due to the challenges in probing ultrafast photoisomerization reactions in biological environments, it remains elusive how the protein influences the photochemistry of the photoswitches, which hampers the rational design of light-regulated therapeutics. To overcome this challenge, we employed first-principles non-adiabatic dynamics simulations to characterize the photodynamics of the phototrexate (PTX), a recently developed photoswitchable anticancer chemotherapeutic that reversibly inhibits its target enzyme dihydrofolate reductase (DHFR). Our simulations show that the protein environment impedes the *trans* to *cis* photoisomerization of the PTX. The confinement in the ligand-binding cavity slows down the isomerization kinetics and quantum yield of the photoswitch by reshaping its conical intersection, increasing its excited-state free energy barrier, and quenching its local density fluctuations. Therefore, we predict that the PTX's *trans*→*cis* photoisomerization in solution precedes its binding with the protein, despite the favorable binding energy of the *trans* isomer. Our findings highlight the importance of the protein environment on the photochemical reactions of the molecular photoswitches. As such, our work represents an important step towards the rational design of light-regulated drugs in photopharmacology.

TOC GRAPHICS



Keywords

Photopharmacology, molecular photoswitches, non-adiabatic dynamics simulation, protein-ligand interaction, free energy simulation, first-principles electronic structure calculation

Introduction

Reversible, spatial and temporal control over protein's structure and function through non-native small molecules have significant implications in minimizing the side effects of chemotherapy.¹⁻⁵ Compared to traditional chemical tools, molecular photoswitches are unsurpassed in their ability to achieve this goal via light-controlled isomerization reactions (Figure 1), making them promising light-regulated drugs.^{3, 6-13} As such, the design and synthesis of novel molecular photoswitches has led to the rapid expansion of the field of photopharmacology.^{3, 6-13} For example, the dihydrofolate reductase (DHFR) has been an essential drug target to treat cancer because its inhibition interrupts the folate metabolism and leads to the death of cancer cells.¹⁴ Methotrexate (MTX) is an effective inhibitor of DHFR and thus a widely prescribed chemotherapeutic drug to treat cancer. However, the ubiquity of the DHFR in healthy cells often leads to severe side effects of MTX such as nausea, joint pain, vomiting and diarrhea. Therefore, to alleviate the side effect of chemotherapy, there is an urgent need to design new DHFR inhibitors with minimal side effects. The molecular photoswitch phototrexate (PTX) was recently designed as a photochromic analog of the MTX.¹³ The *trans* isomer of the PTX has a lower binding affinity with the DHFR than its *cis* isomer, and therefore the *trans*→*cis* photoisomerization promotes the binding of the ligand with the protein. Thus, the PTX can reversibly inhibit the DHFR enzyme with high spatial-temporal precision under the control of light (Figure 1),¹³ which can mitigate its side effect compared to the MTX.

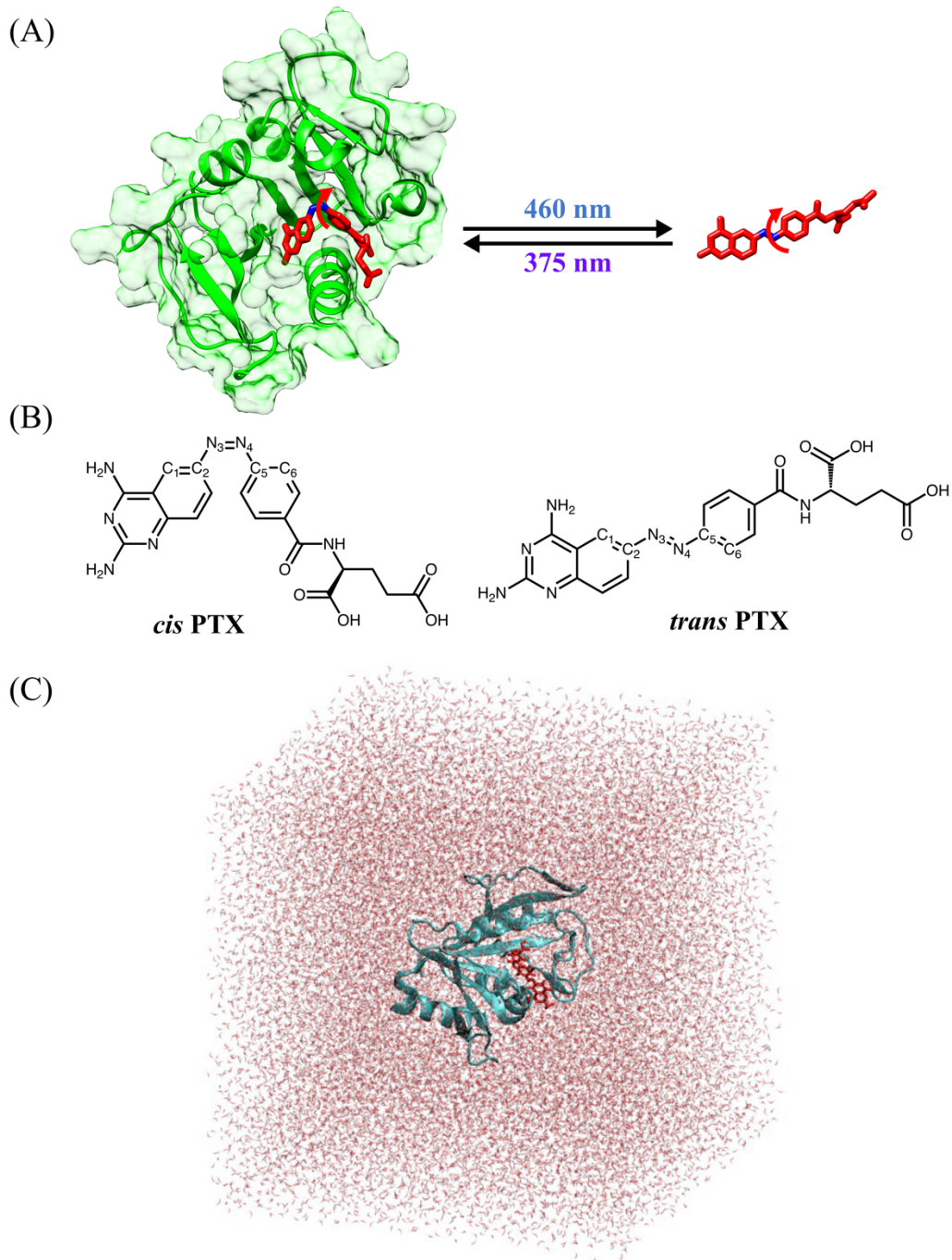


Figure 1. (A) The photoisomerization between the *cis* and *trans* PTX enables reversible inhibition of the DHFR under different wavelengths of light. (B) The key atoms that define the essential torsional modes in the photoisomerization are labeled in the chemical structures of the *cis* and

trans PTX. (C) Simulation setup for the *trans* PTX (red licorice) bound to the DHFR (green ribbons) solvated in the water (red dots).

Despite the molecular photoswitches' successful applications in photopharmacology, their design often involves expensive trial-and-error experiments.¹³ Therefore, the rational design of more effective photoswitchable drugs is necessary. However, several fundamental questions remain unanswered regarding the interactions among the protein, the photoswitchable ligand and the light. For example, it was reported that the *trans* PTX's binding affinity with the DHFR is only slightly less favorable than the *cis* PTX (-9.3 kcal/mol vs. -10.8 kcal/mol).¹³ The non-negligible binding affinity of the *trans* PTX raises a critical question: does the *trans* isomer first bind with the protein and then undergo the *trans*→*cis* photoisomerization, or is the reverse sequence of events more probable? Additionally, it remains unclear how the protein environment affects the photoisomerization kinetics and quantum yield (QY) of the photoswitches. Answering these questions at molecular-level detail is crucial to the successful design of next-generation light-activatable drugs for cancer treatment.

Molecular simulation is indispensable to achieving this goal because it can reveal the energetics and kinetics of biochemical reactions at atomistic-level detail, which is often beyond the limit of current experimental techniques. However, it is challenging to use computational methods to study photochemical reactions in proteins. Firstly, traditional classical dynamics cannot describe the non-adiabaticity of photodynamics involving population transfer among different electronic states. Hence, accurate non-adiabatic dynamics methods are necessary to propagate the coupled nuclear and electronic degrees of freedom according to the time-dependent Schrödinger's equation.

However, such simulations often demand expensive computational costs. Secondly, the accuracy of non-adiabatic dynamics simulations also depends on the calculations of the potential energy surfaces (PES) and non-adiabatic couplings (NAC) of the electronic states, and it is essential to employ first-principles electronic structure methods that include both static and dynamic correlations.¹⁵⁻¹⁶ However, most first-principles electronic structure methods fulfilling these requirements are usually too expensive to afford “on-the-fly” non-adiabatic dynamics simulations. These computational challenges result in a general lack of understanding of how the confinement imposed by the biomolecules impacts the photodynamics of molecular photoswitches.

As a first step to overcome these challenges, in this work, we employed first-principles non-adiabatic dynamics simulations to probe the effect of protein-ligand interactions on the *trans*→*cis* photoisomerization of the PTX. The non-adiabatic molecular dynamics simulations were performed using the *ab initio* multiple spawning (AIMS)¹⁷⁻¹⁹ algorithm coupled with “on-the-fly” first-principles electronic structure calculations using the hole-hole Tamm-Dancoff approximated density functional theory (hh-TDA-DFT).¹⁵ The hh-TDA-DFT method incorporates both the static and dynamic electron correlations that are essential for the accurate description of the PES and NACs.^{15-16, 20} The accuracy of the hh-TDA-DFT method has been systematically benchmarked against high-level *ab initio* calculations and experiments.^{15, 20-22} The AIMS algorithm is a computationally efficient and accurate non-adiabatic dynamics method for evolving the nuclear wave-packets on coupled electronic states according to time-dependent Schrödinger’s equation.¹⁷⁻¹⁹ It uses trajectory basis functions (TBF) to represent the evolution of the nuclear wavefunctions, which significantly boosts its computational efficiency.¹⁷⁻¹⁹ Although a few simulation studies have investigated the photodynamics of azobenzene in biological systems,²³⁻²⁵ they employed

more approximate simulation methods than the current study. Specifically, the electronic structure was treated with either semiempirical methods²³⁻²⁴ or *ab initio* methods that lack dynamic electron correlation.²⁵ In addition, the non-adiabatic dynamics was simulated by the surface-hopping method,²⁶ which suffers from the overcoherence issue²⁷ that has not been fully resolved by the decoherence corrections without introducing new problems.²⁸⁻²⁹ By comparison, the combination of the hh-TDA-DFT and AIMS methods in the current study offers a more accurate description of the photodynamics of the azobenzene-derivatives. These state-of-the-art computational approaches enable us to gain deep insights into the effects of the protein environment on the photochemical reactions of the molecular photoswitches, which will be discussed in detail below.

Method

This section details the procedures for (1) system setup and molecular docking simulation; (2) classical molecular dynamics (MD) simulation; (3) ground state quantum mechanics/molecular mechanics (QM/MM) MD simulation; (4) absorption spectra calculation; (5) AIMS simulation; (6) characterization of the minimum energy conical intersections and (7) the excited-state umbrella sampling simulations.

System setup and molecular docking simulations

Starting from the crystal structure of DHFR (PDB code: 1U72)³⁰, the co-crystallized ligands, ions and water molecules were first removed. The missing protein hydrogen atoms were added, and the partial charges of all protein atoms were assigned based on the Amber ff14SB force field.³¹⁻³² To prepare the coordinates and partial charges of the ligand PTX, the gas-phase minimum corresponding to the *trans* isomer was first optimized on the ground state using the density

functional theory (DFT) with the B3LYP functional.³³ At the optimized geometry, the Hartree-Fock/6-31G* calculations were performed, followed by fitting the partial charges of the PTX using the restrained electrostatic potential (RESP)³⁴ approach. Then the Autodock Vina software³⁵ was used to identify the binding poses of the *trans* PTX in the DHFR. The structure with the best score in a cluster of top-score structures was selected. The selected *trans* PTX structure partially overlaps with the methotrexate in the crystal structure³⁰ for the functional groups on one side of the N=N bond. Then, the protein-ligand complex was solvated in a box of water molecules with a periodic boundary condition of $\sim 85 \times 91 \times 96 \text{ \AA}^3$ (Figure 1 C). The force field parameters of the PTX were generated using the general Amber force field (GAFF) procedure.³⁶⁻³⁷ The water and protein were modeled using the SPFF/Fw³⁸ and Amber ff14SB force fields,³¹⁻³² respectively.

In order to probe the effects of the protein on the photodynamics of PTX, an aqueous solution simulation system was also constructed by solvating the *trans* PTX in a box of water molecules, resulting in a periodic condition of $\sim 51 \times 53 \times 59 \text{ \AA}^3$.

Classical MD equilibration

The coordinates of the system were first optimized while restraining the heavy atoms of the protein and ligand using harmonic potential with $100 \text{ kcal/mol/\AA}^2$ force constant. Next, the system was equilibrated for 10 ps in the constant NVT ensemble at a temperature of 300K, followed by 5 ns equilibration in the NPT ensemble at the same temperature and 1 atm pressure, where the force constants of the restraints are gradually reduced to zero. Then, 50 ns simulation was performed in the constant NPT ensemble without any restraints. The first 10 ns of the trajectory was discarded as equilibration, and the last 40 ns of the trajectory was treated as a production run. A Langevin thermostat with a friction coefficient of 1 ps^{-1} and a Monte Carlo barostat with an attempt

frequency of 100 steps was used to maintain the temperature and pressure, respectively. A time step of 1 fs was used to propagate the dynamics. A 12 Å cutoff was used for non-bonded interactions, and electrostatic interactions were calculated using the particle mesh Ewald (PME) method. The classical MD simulations of the *trans* PTX in the aqueous solution followed the same procedures as the protein-ligand systems, except that no restraint on the protein was involved. All the classical MD simulations were performed with the OPENMM software package.³⁹

Ground state QM/MM dynamics

Due to the inaccuracies in the force-field description of the PTX ligand, we employed QM/MM MD simulations to adjust the geometries of the PTX sampled from the classical MD simulation. From the classical MD trajectory, 40 snapshots were extracted evenly with 1 ns time interval. For each snapshot, a subsystem with an open boundary condition was created and used as starting structures for the QM/MM simulations. This is due to the current limitation of the TeraChem software package⁴⁰⁻⁴³ in treating the PBC in the QM/MM simulations. The subsystem includes the protein, ligand and every water molecule that has at least one atom within 10 Å of any atom in the protein-ligand complex. Then, for each subsystem, a QM/MM MD trajectory was initiated and propagated on the ground state in the constant NVT ensemble at 300K temperature using a Langevin thermostat. Each QM/MM trajectory was propagated for ~7 ps on the ground state with a timestep of 0.5 fs. As a result, ~350 ps QM/MM dynamics were propagated in total, which originated from 40 ns MM conformational sampling of the ligand-protein complex. The QM region included only the PTX, while the MM region included the rest of the system. The QM and MM regions were coupled by the electrostatic embedding scheme, where the electronic density of the QM atoms interacts with the fixed point charges on the MM atoms. The QM region was treated

with B3LYP/6-31G*, and the MM region was treated with the same force fields as described in the previous sections. The QM/MM simulations for the *trans* PTX in the aqueous solution were also performed following the same procedures. The TERACHEM⁴⁰⁻⁴³ interfaced with the OPENMM packages³⁹ were used to perform the ground-state QM/MM simulations.

Absorption spectra calculation

From each of the 40 QM/MM trajectories, snapshots were extracted from the last 5 ps with a time interval of 10 fs, resulting in 20,000 snapshots in total. For each snapshot, the excitation energies and oscillator strengths for the $S_0 \rightarrow S_1$ ($n\pi^*$) and $S_0 \rightarrow S_2$ ($\pi\pi^*$) transitions were calculated using the QM/MM approach with the same partitioning scheme as the MD simulation. The QM region was treated with the hh-TDA-DFT method with B3LYP functional and 6-31G* basis set (hh-TDA-B3LYP/6-31G*), and the MM region with the same force fields as mentioned above. The absorption spectra were then calculated following ref⁴⁴, and Lorentzian functions were used to convolve the cross-section in energy with a full width at half maximum (FWHM) of 0.15 eV.

AIMS simulation

The PTX's *trans* \rightarrow *cis* isomerization was induced by illumination at 375 nm.¹³ The wavelength corresponds to the $S_0 \rightarrow S_2$ ($\pi\pi^*$) initial excitation ("Results and Discussion"). The AIMS simulations were prepared accordingly. In particular, 40 initial conditions (ICs, i.e., the initial coordinates and momenta of the TBFs' centroids) were taken from the last snapshots of the 40 ground-state QM/MM trajectories. According to our calculated absorption spectra (Figure 2), the TBFs were initially placed on the S_2 state. The TBFs were then propagated and spawned according to the full multiple spawning (FMS) algorithm.¹⁷⁻¹⁹ The quantities required for the FMS simulation,

such as energies, gradients and non-adiabatic couplings of the S_0 , S_1 and S_2 states, were calculated on the fly using the hh-TDA-B3LYP/6-31G* method. Stochastic-Selection AIMS (SS-AIMS) method⁴⁵ was employed to speed up the simulation while maintaining accuracy.²⁸ The TBF overlap threshold to trigger stochastic selection was set to $\sim 10^{-3}$.²¹ In the SS-AIMS simulations, three reruns with different random seeds were performed for each IC. For each IC, the simulation was terminated after more than 95% of the TBFs' population had decayed to the ground state or when the simulation time had reached 5 ps. The S_1 and S_2 states' relaxation time constants (τ_{S_1} and τ_{S_2}) were fitted using the following first-order kinetics model:

$$P_{S_2}(t) = \exp\left(-\frac{t}{\tau_{S_2}}\right) \quad \text{Eq 1}$$

$$P_{S_1}(t) = \frac{\tau_{S_1}}{\tau_{S_1} - \tau_{S_2}} \left[\exp\left(-\frac{t}{\tau_{S_1}}\right) - \exp\left(-\frac{t}{\tau_{S_2}}\right) \right] \quad \text{Eq 2}$$

The photoisomerization QY was calculated based on the centroid geometries and populations of the ground state TBFs. In particular, each ground state TBFs was propagated for ~ 100 fs after it was decoupled from the S_1 state. The decoupling was defined as when the dot product of the centroid's velocity with the S_1/S_0 derivative coupling vector decreased below 0.01 a.u. Then, the centroids of the TBFs were minimized on the ground state. If the absolute value of the final C2-N3=N4-C5 torsion angle (Figure 1 B) was below 15° , the population of this TBF was assigned to *cis* isomer. If the torsion angle is between 165° and 195° , the population was assigned to *trans* isomer. The QY for the *trans* \rightarrow *cis* isomerization was defined as the ratio of the total population of the *cis* isomer to the total population of all ground state TBFs. For the SS-AIMS simulations, the

QYs were averaged over three reruns. The error bars in the QYs were obtained by bootstrapping over 1000 random samples with replacement.

The AIMS simulations were performed using the FMS90 code interfaced with the TERACHEM/OPENMM software packages.³⁹⁻⁴³

Characterization of the minimum energy conical intersections (CI)

In the AIMS simulations of the DHFR and aqueous solution systems, for each ground state TBF successfully spawned through a reactive S_0/S_1 CI (see “Result and Discussion” for definition), the *minimum energy conical intersections* (MECI) were searched starting from the spawning geometry. During the CI optimization, the atoms that were more than 4 Å away from the PTX were kept frozen. Then, the following quantities of the MECI were calculated following the definition introduced by Yarkony:⁴⁶

$$\mathbf{g} = \frac{1}{2} \left(\frac{\partial E_1}{\partial \mathbf{R}} - \frac{\partial E_0}{\partial \mathbf{R}} \right) \quad \text{Eq 3}$$

$$\mathbf{h} = \left\langle \phi_1 \left| \frac{\partial H}{\partial \mathbf{R}} \right| \phi_0 \right\rangle \quad \text{Eq 4}$$

$$\mathbf{s} = \frac{1}{2} \left(\frac{\partial E_1}{\partial \mathbf{R}} + \frac{\partial E_0}{\partial \mathbf{R}} \right) \quad \text{Eq 5}$$

$$s^x = \frac{\mathbf{s} \cdot \mathbf{g}}{g^2} \quad \text{Eq 6}$$

$$s^y = \frac{\mathbf{s} \cdot \mathbf{h}}{h^2} \quad \text{Eq 7}$$

The E_1 and E_0 are the potential energies of the S_1 and S_0 states, respectively. The \mathbf{g} is the difference of the gradient vectors of the two states. The \mathbf{h} is the non-adiabatic coupling vector between the two states, and because the mixing of the two electronic wavefunctions (ϕ_1 and ϕ_0) is arbitrary at the CI, the \mathbf{g} and \mathbf{h} can be constructed as orthogonal vectors, with g and h denoting

their lengths. These two vectors span the branching plane of the MECI, meaning that displacements from the MECI along these two vectors lifts the energy degeneracy of the two states (Figure 5). The \mathbf{s} is the seam vector describing the topography, and the s^x and s^y are the scaled tilt parameters quantifying the sloped/peaked character along each of the two branching plane vectors. The s^x and s^y are zero for perfectly peaked (hourglass-shaped) CI and are larger for more sloped CI along the corresponding branching plane vectors.

Excited-state umbrella sampling simulations

Umbrella sampling simulations were performed on the S_1 state to calculate the free energy surfaces (FES) of the PTX in the DHFR and solution, resulting in two one-dimensional potential of mean forces (PMFs). The reaction coordinate of the PMF was the central $C_2-N_3=N_4-C_5$ dihedral angle ($\theta_{C_2N_3N_4C_5}$). For each PMF, ten umbrella windows were placed along the reaction path, with window centers ranging from 180° to 90° with a 10° interval, and a harmonic potential with $100 \text{ kcal/mol/radian}^2$ force constant was imposed on the reaction coordinate for each window. The first umbrella window started from the last snapshot of a ground-state QM/MM MD trajectory. The initial structure of each window was obtained from the structure of its adjacent window after equilibration for ~ 1 ps. The hh-TDA-B3LYP/Amber method was used to calculate the S_1 state forces on the atoms on the fly during the umbrella sampling simulations. For each umbrella window, the sampling was performed for at least 6 ps. The first 1 ps trajectory was discarded as equilibration, and the remaining 5 ps trajectory was treated as the production run. The total excited-state QM/MM umbrella sampling time was thus 100 ps. The PMFs were calculated by unbiasing the distributions of the reaction coordinates using the WHAM algorithm.⁴⁷⁻⁴⁸ The umbrella

sampling simulations were performed using the TERACHEM⁴⁰⁻⁴³ interfaced with the OPENMM packages.³⁹

Results and Discussion

Excited-state population decay and quantum yield

In the experiment,¹³ the *trans*→*cis* photoisomerization was induced by 375 nm light. Our calculated absorption spectrum in the aqueous solution (Figure 2) quantitatively agreed with the experiment for the wavelength of maximum absorbance, with a deviation of only 0.06 eV (Figure 2). The absorption spectrum is dominated by the $S_0 \rightarrow S_2$ excitation induced by the $\pi \rightarrow \pi^*$ transition, which has much larger oscillator strength than the $S_0 \rightarrow S_1$ excitation induced by the $n \rightarrow \pi^*$ transition. Based on this, we chose the S_2 state as the initial electronic state in the AIMS simulation. The time evolutions of the excited states' populations and the relaxation time constants are illustrated in Figure 3 and Table 1. The QYs for the photoisomerization in the DHFR and aqueous solution are compared in Table 1. It is evident that the *trans*→*cis* QY is reduced from 0.31 ± 0.05 in solution to 0.15 ± 0.05 in DHFR. Moreover, the S_1 and S_2 excited states have longer lifetimes in the DHFR than the aqueous solutions (Figure 3 and Table 1). Overall, the above analysis indicates that the protein environment impedes the *trans*→*cis* photoisomerization of the PTX in terms of both QY and reaction kinetics.

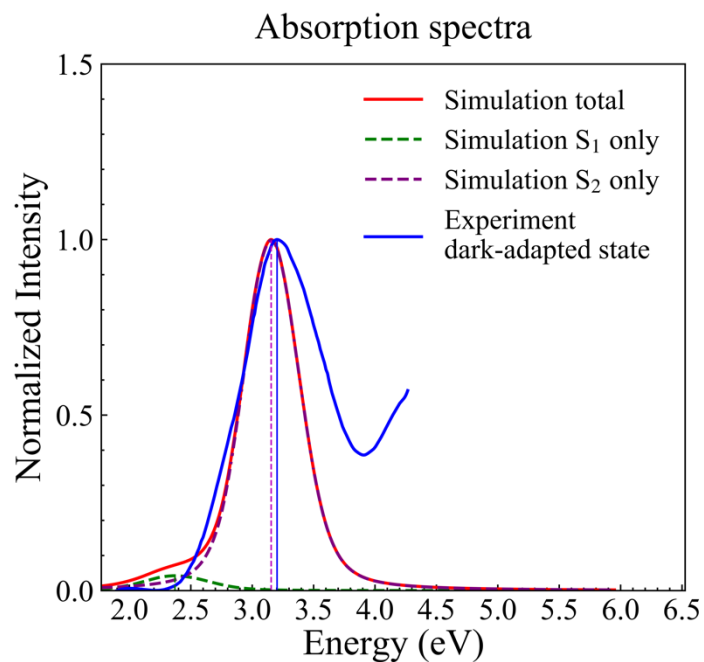


Figure 2. The calculated absorption spectrum (dashed red line) of the *trans* PTX in the aqueous solution agrees well with the experiment (solid blue line). The contributions from the $S_0 \rightarrow S_1$ and $S_0 \rightarrow S_2$ excitations to the calculated spectrum are in green and purple dashed lines, respectively. The peaks of the experimental and calculated spectra are labeled by the vertical lines. The experimental and calculated wavelengths of maximum absorption for the $S_0 \rightarrow S_2$ excitation are 386.7 nm (3.21 eV) and 393.6 nm (3.15 eV), respectively.

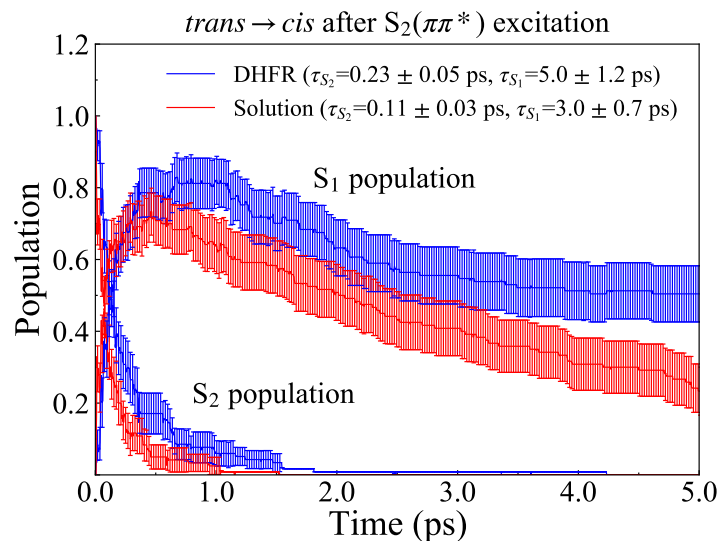


Figure 3. The excited-state population decay as a function of time for the *trans*→*cis* isomerization of the PTX following the S_2 ($\pi\pi^*$) excitation. The blue and red curves indicate the results for the PTX in the DHFR and aqueous solution, respectively. The relaxation time constants of the excited states (τ_{S_1} and τ_{S_2}) are labeled in the parentheses. The protein environment has a pronounced effect in slowing down the kinetics of the *trans*→*cis* isomerization.

Table 1. The isomerization QY, the excited-state lifetimes, the average and variance of the PTX's contact number (Eq 8) for the *trans*→*cis* photoisomerization after the S_2 excitation in the DHFR and aqueous solution. All quantities are averaged over 40 ICs.

Quantities	Solution	DHFR
QY	0.31 ± 0.05	0.15 ± 0.05
τ_{S_1}	3.0 ± 0.7 ps	5.0 ± 1.2 ps
τ_{S_2}	0.11 ± 0.03 ps	0.23 ± 0.05 ps
\bar{N}	148 ± 1	134 ± 1
$\overline{\sigma_N^2}$	119 ± 5	73 ± 5

Origin of the reduced isomerization quantum yield in the DHFR

In order to understand the origin of the lowered QY in the DHFR, we categorized the $S_1 \rightarrow S_0$ non-radiative decay by the type of the S_0/S_1 conical intersection (CI) through which the population transfer occurs (Figure 4). Similar to the azobenzene's photodynamics,^{15, 20-21} the PTX's S_1 state population can decay to the S_0 state through either the unreactive or reactive S_0/S_1 CIs (Figure 4 C&D), which belong to the same crossing seam.⁴⁹⁻⁵¹ The unreactive CI features a near planar geometry with the central $C_2-N_3=N_4-C_5$ dihedral angle ($\theta_{C_{NNC}}$) close to 180° and large $C_2-N_3=N_4$ and $N_3=N_4-C_5$ bending angles (α_{CNN} and α_{NNC}) beyond 140° (Figure 4 C). The non-radiative decay through this type of CI exclusively generates the *trans* isomer because there is no rotation around the $\theta_{C_{NNC}}$. In contrast, the reactive CI features a near perpendicular geometry with the $\theta_{C_{NNC}}$ close to 90° (Figure 4 D). The non-radiative decay through this type of CI will have some possibility to successfully generate the *cis* isomer product due to the rotation of the $\theta_{C_{NNC}}$. The unreactive CI is energetically higher than the reactive CI by 0.78 eV in the vacuum. As a result, the former can only be accessed in the initial stage of the photodynamics when the high kinetic energy released from the $S_2 \rightarrow S_1$ relaxation has not been fully dissipated to the environment.²¹ Thus, the $S_1 \rightarrow S_0$ decay through the reactive CI on average occurs at a longer timescale than the unreactive CI (1.9 ± 0.2 ps vs. 0.9 ± 0.1 ps in DHFR and 2.2 ± 0.2 ps vs. 0.4 ± 0.1 ps in solution, estimated from AIMS simulations up to 5 ps). Indeed, after 2.5 ps of photoexcitation, the $S_1 \rightarrow S_0$ decay is exclusively through the reactive CI.

Our population transfer analysis (Figure 4 A & B) indicates that the molecular environment around the PTX impacts the *cis:trans* branching ratio after the system decays through the reactive S_0/S_1 CI. Specifically, in the DHFR, ~23% of the population decaying through the reactive CI eventually generates the *cis* isomer, whereas in the aqueous solution, it is ~46%. Thus, the difference in the

branching ratio through the reactive S_0/S_1 CI contributes to the different QYs in these two systems. Although our AIMS simulations were not propagated until full depletion of the excited-state populations due to resource limitations, the conclusion from the above analysis will not be changed much by further extending our simulations. This is because it is reasonable to assume that for each system, after 5 ps, all the remaining S_1 populations (50% for DHFR and 24% for aqueous solution) will decay through the reactive CI with the same branching ratio as the first 5 ps. Based on this assumption, we estimate that the final QY will be 0.19 and 0.35 in the DHFR and the aqueous solution, respectively, which confirms our observation that a lower branching ratio of the reactive CI leads to a reduced QY in the DHFR as compared to the aqueous solution.

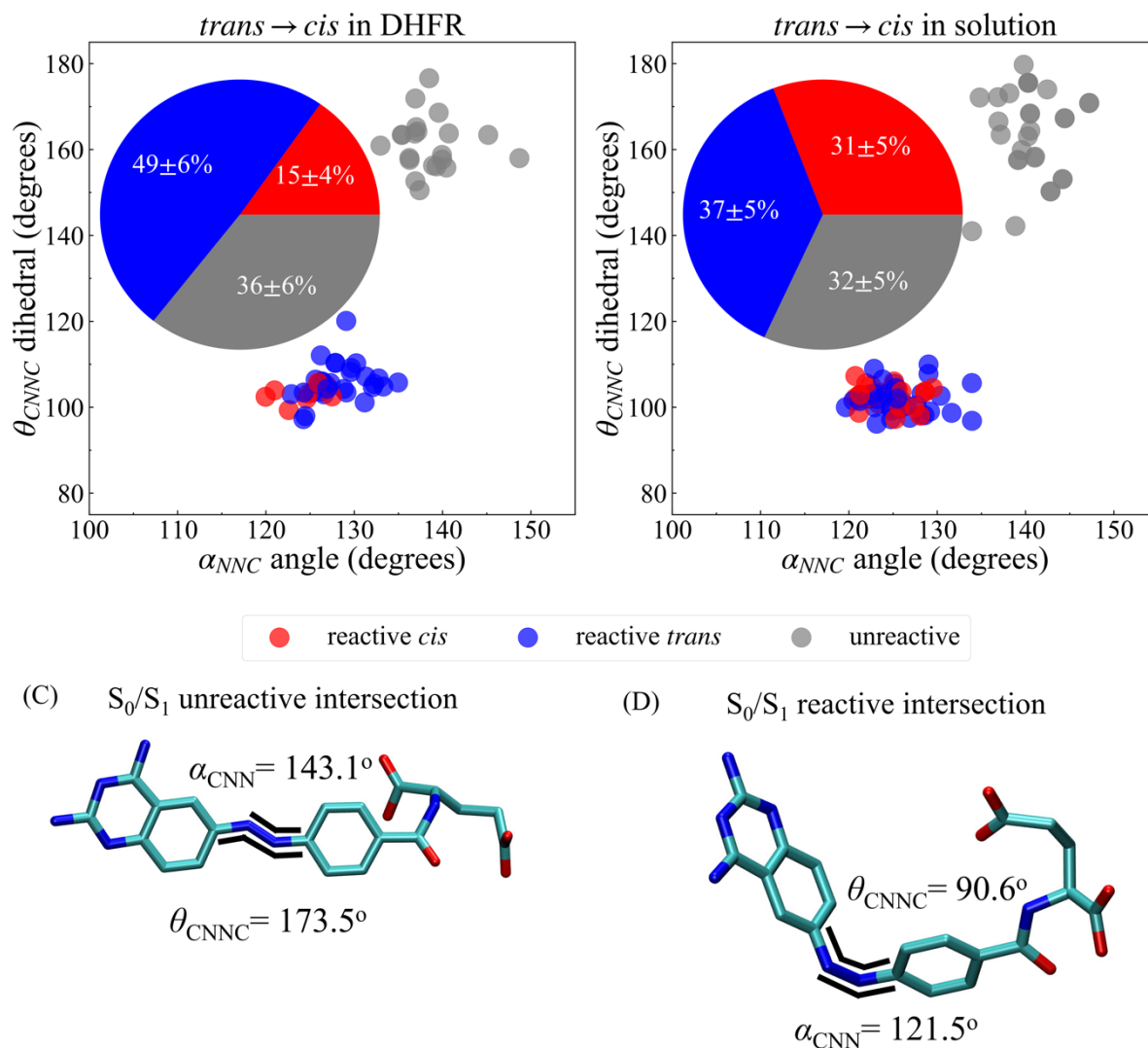


Figure 4. (A, B) Decomposition of population decay through the unreactive and reactive S_0/S_1 CIs for the PTX *trans*→*cis* photoisomerization following the S_2 excitation in the DHFR and aqueous solution, respectively. Each circle represents a S_0 state TBF, with its location determined by the TBF's centroid geometry at the spawning time, i.e., the larger one of the two C-N=N angles (α_{CNN}) and the C-N=N-C torsion (θ_{CNMC}). The circle is grey if the TBF is spawned through a S_1/S_0 unreactive CI (C) and exclusively ends up as the *trans* isomer. The circle is colored red (blue) if the S_0 TBF is spawned through a reactive S_0/S_1 CI (D) and ends up as the *cis* (*trans*) isomer. The

pie charts in the insets indicate the relative populations of the S_0 TBFs in each of the three categories. The structures of the unreactive and reactive minimum energy CIs (MECI) were optimized in the gas phase and are similar to those encountered in the DHFR and the aqueous solution. The uncertainties are estimated as one standard error using bootstrapping with 1000 samples.

The different branching ratios of the reactive CIs in the DHFR and solution suggest that the protein environment shapes the topography of the reactive S_0/S_1 CI to make it less productive for generating the *cis* isomer. To test this hypothesis, we further characterized the topography of the reactive S_0/S_1 MECIs in the DHFR and the solution (Table 2 and Figure 5). As shown in Table 2, the MECIs in the DHFR environment have larger s^x on average (Eq 6) than in the aqueous solution. The larger s^x indicates that the MECIs in the DHFR are more sloped along the \mathbf{g} vector than those in the aqueous solution. A visual inspection of the MECIs in the two environments (Figure 5) confirms such observation. The \mathbf{g} vector is approximately in line with the direction of the N=N bond rotation (Figure 5 C). The more sloped MECI will generally lead to a reduced QY.⁵² Therefore, the DHFR's protein environment reduces the photoisomerization QY by changing the topography of the reactive S_0/S_1 CI.

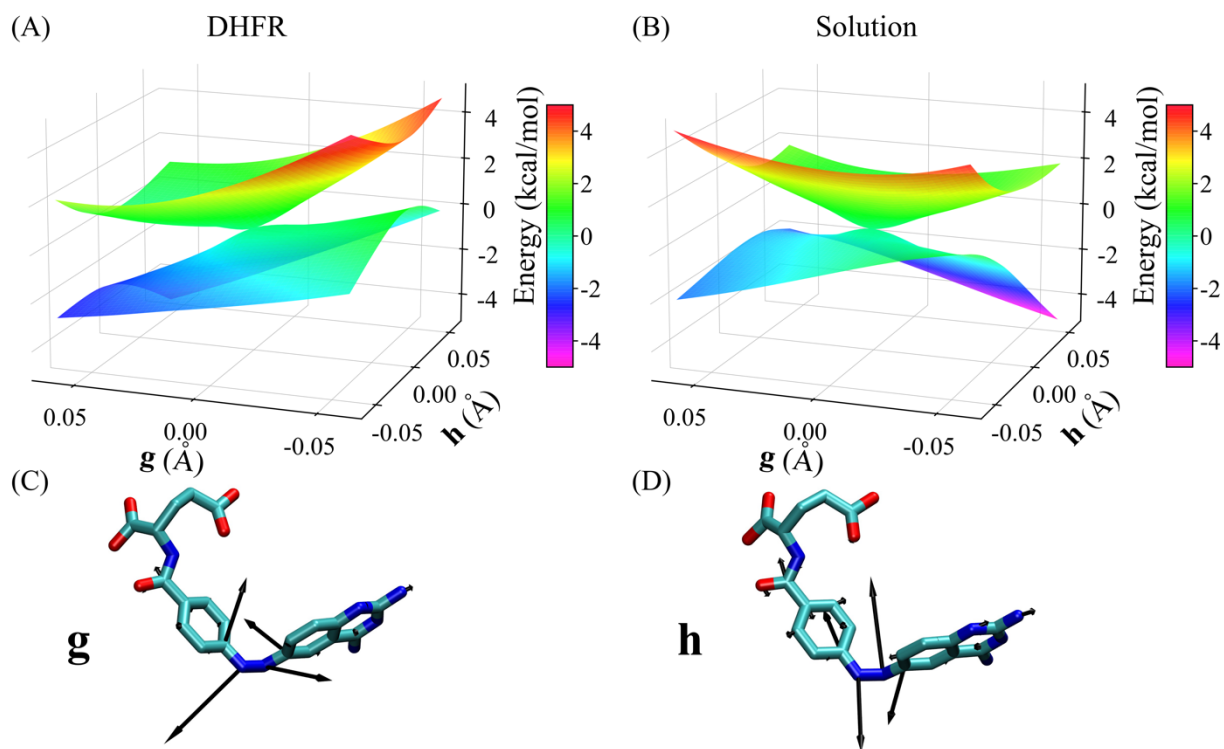


Figure 5. The topography of the reactive S_0/S_1 MECI in the (A) DHFR and (B) aqueous solution. The protein environment makes the MECI more sloped along the \mathbf{g} vector than the aqueous solution. (C) and (D): the \mathbf{g} and \mathbf{h} vectors (black arrows) spanning the CI plane for the reactive S_0/S_1 MECI. The Franck-Condon point is projected on the 2-dimensional CI plane at 0.13 Å and 0.20 Å along the \mathbf{g} and \mathbf{h} vectors, respectively.

Origin of the slowdown of isomerization kinetics in the DHFR

To further understand the effect of the environment on the kinetics of the non-radiative decay, we performed umbrella sampling simulations with the hh-TDA-B3LYP/Amber method and calculated the S_1 state PMFs for the PTX's isomerization in both the DHFR and the aqueous solution (Figure 6). The PMFs were calculated along the rotation of $\theta_{C_{NNC}}$ which is key to photoisomerization. In the aqueous solution, the rotation around the $\theta_{C_{NNC}}$ from the planar

geometry ($\sim 180^\circ$) to the near perpendicular geometry around the CI ($\sim 90^\circ$) only needs to overcome a small free energy barrier of ~ 0.5 kcal/mol (Figure 6), and the process is exergonic with a $\Delta G = -7.3$ kcal/mol. In contrast, in the DHFR, the rotation around the θ_{CNC} needs to overcome a larger barrier of 1.7 kcal/mol, and the process is endergonic with $\Delta G = 1.7$ kcal/mol (Figure 6). As mentioned above, the reactive CI is the major channel for the $S_1 \rightarrow S_0$ decay beyond 1 ps. Thus, the significant slowdown of the S_1 state population decay in the DHFR can be mainly attributed to the increased S_1 state free energy barrier that impedes the rotation of the θ_{CNC} .

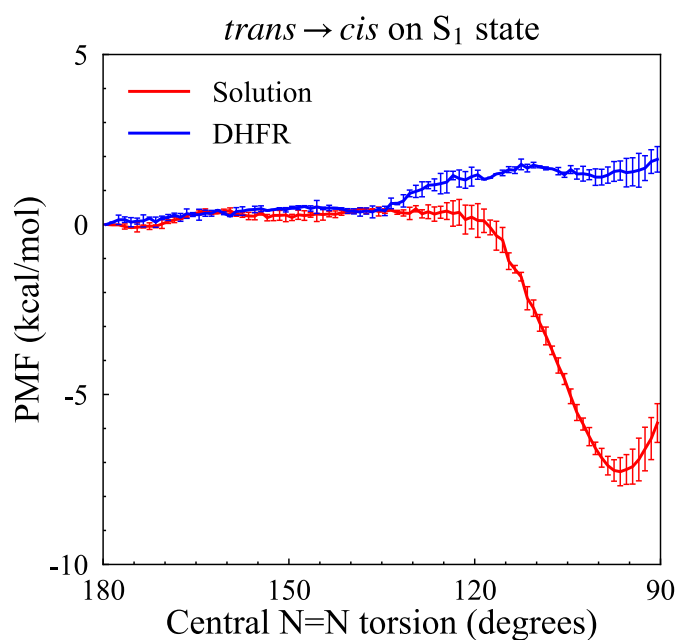


Figure 6. Potential of mean force for the rotation of the θ_{CNC} on the S_1 state in the aqueous solution (red) and the DHFR (blue). The protein environment increases the free energy barrier to access the reactive S_0/S_1 CI and slows down the $trans \rightarrow cis$ photoisomerization. The error bars were estimated by block averaging analysis of the QM/MM trajectories of the umbrella windows.

The slowdown of the isomerization kinetics in the protein can also be understood in terms of the PTX's local density fluctuation, which is characterized by the variance of the PTX's contact number N defined as below:⁵³

$$N = \sum_{i \in A} \sum_{j \in B} \frac{1 - [(r_{ij} - d_0)/r_0]^6}{1 - [(r_{ij} - d_0)/r_0]^{12}} \quad \text{Eq 8}$$

The A and B are the sets of atoms comprising the PTX molecule and the rest of the system, respectively. The r_{ij} is the distance between a particular pair of atoms from the PTX and its environment. The $d_0 = 2.5 \text{ \AA}$ and the $r_0 = 0.5 \text{ \AA}$. With these parameters, the N effectively counts the number of atoms between 2 to 3 \AA around the PTX.⁵³ For each S_1 state TBF that encountered a reactive S_0/S_1 CI and successfully spawned the final S_0 state TBF (see "Method"), the average and variance of N (\bar{N} , σ_N^2) were calculated for the trajectory of this TBF's centroid. These quantities were averaged over all 40 initial conditions (IC) and summarized in Table 1. It is noteworthy that the $\overline{\sigma_N^2}$ is lower in the DHFR than in water, indicating that the protein scaffold reduces the fluctuation of atom density around the PTX. In other words, the protein environment imposes more rigid confinement on the PTX than the aqueous solution. Thus, there is a stronger caging effect in the DHFR that slows down the PTX's photoisomerization kinetics, which is consistent with our free energy calculations.

Our conclusions are consistent with the study of Mondal et al.,²³ which concludes that the DNA double helix significantly slows down the *trans*→*cis* photoisomerization of the covalently incorporated azobenzene. Our conclusions are also in line with the study of Kingsland et al.,⁵³ which reported a positive correlation between the density fluctuation around the azobenzene and its photoisomerization QY in the DNA double helix.⁵³ However, the study in ref ⁵³ was limited to the ground-state force field description of the azobenzene's dynamics. In contrast, we explicitly

simulated the non-radiative decay of the excited states using first-principles non-adiabatic dynamics simulations. Hence, our data offers a more comprehensive and accurate description of the photodynamics of the azobenzene-derived photoswitches and extends the previous findings to the protein environment, which has been less studied to date.

Table 2. The average values of s^x and s^y parameters (Eqs 6 & 7) depicting the topography of the reactive S_0/S_1 CI of PTX in the aqueous solution and the DHFR.

CI parameters	Solution	DHFR
$\overline{s^x}$	0.33 ± 0.03	0.47 ± 0.05
$\overline{s^y}$	$(6 \pm 5) \times 10^{-5}$	$(1.1 \pm 0.2) \times 10^{-6}$

Conclusion

In this work, we aimed to answer an essential question in photopharmacology: how do biomolecules perturb the photochemical properties of the photoswitchable drug molecules? We employed first-principles non-adiabatic dynamics simulations to understand how the DHFR affects the *trans*→*cis* photoisomerization of PTX, an azobenzene-derived molecular photoswitch recently designed as an anticancer chemotherapeutic agent. We found that the *trans*→*cis* photoisomerization is impeded by the protein environment compared to the aqueous solution. The protein environment makes the reactive S_0/S_1 CI more sloped and reduces the QY. Also, it quenches local density fluctuations around the PTX and increases the excited-state free energy barrier for the central torsion mode ($\theta_{C_{NN}C}$), which eventually slows down the photoisomerization kinetics. Thus, the PTX's *trans*→*cis* photoisomerization is more likely to occur in the aqueous solution before binding with the DHFR, despite the *trans* isomer's non-negligible binding affinity with the enzyme.¹³ Our first-principles non-adiabatic dynamics simulations thus deepen our

understanding of the delicate interplay between the photoswitchable inhibitor and the DHFR. Further studies of the *cis*→*trans* photoisomerization (work in progress) will still be necessary to unravel the whole picture of the photocycle underlying the interactions between the DHFR, PTX and light, which will shed light on the design principles of light-activatable therapeutics in photopharmacology.

ASSOCIATED CONTENT

AUTHOR INFORMATION

Corresponding Author

Ruibin Liang

ORCID: 0000-0001-8741-1520

Notes

The authors declare no competing financial interests.

ACKNOWLEDGMENT

This work was supported by the startup funds from Texas Tech University. The researcher used GPU computing facilities provided by the High-Performance Computing Center at Texas Tech University. The authors also acknowledge the helpful insights provided by Prof. Dirk Trauner in the Department of Chemistry at New York University.

REFERENCES

1. Fornasari, D. Pharmacotherapy for Neuropathic Pain: A Review. *Pain and Therapy* **2017**, *6*, 25-33.
2. Kalinichenko, L. S.; Gulbins, E.; Kornhuber, J.; Müller, C. P. The role of sphingolipids in psychoactive drug use and addiction. *J. Neural Transm.* **2018**, *125*, 651-672.
3. Klippenstein, V.; Mony, L.; Paoletti, P. Probing Ion Channel Structure and Function Using Light-Sensitive Amino Acids. *Trends Biochem. Sci.* **2018**, *43*, 436-451.
4. Rexer, B. N.; Arteaga, C. L. Intrinsic and Acquired Resistance to HER2-Targeted Therapies in HER2 Gene-Amplified Breast Cancer: Mechanisms and Clinical Implications. **2012**, *17*, 1-16.
5. van Beijnum, J. R.; Nowak-Sliwinska, P.; Huijbers, E. J. M.; Thijssen, V. L.; Griffioen, A. W. The Great Escape; the Hallmarks of Resistance to Antiangiogenic Therapy. *Pharmacol. Rev.* **2015**, *67*, 441.
6. Hüll, K.; Morstein, J.; Trauner, D. In Vivo Photopharmacology. *Chem. Rev.* **2018**.
7. Broichhagen, J.; Frank, J. A.; Trauner, D. A Roadmap to Success in Photopharmacology. *Acc. Chem. Res.* **2015**, *48*, 1947-1960.
8. Tochitsky, I.; Kienzler, M. A.; Isacoff, E.; Kramer, R. H. Restoring Vision to the Blind with Chemical Photoswitches. *Chem. Rev.* **2018**, *118*, 10748-10773.
9. Brechun, K. E.; Arndt, K. M.; Woolley, G. A. Strategies for the photo-control of endogenous protein activity. *Curr. Opin. Struct. Biol.* **2017**, *45*, 53-58.
10. Szymański, W.; Beierle, J. M.; Kistemaker, H. A. V.; Velema, W. A.; Feringa, B. L. Reversible Photocontrol of Biological Systems by the Incorporation of Molecular Photoswitches. *Chem. Rev.* **2013**, *113*, 6114-6178.

11. Mart, R. J.; Allemann, R. K. Azobenzene photocontrol of peptides and proteins. *Chem. Commun.* **2016**, *52*, 12262-12277.
12. Gautier, A.; Gauron, C.; Volovitch, M.; Bensimon, D.; Jullien, L.; Vriza, S. How to control proteins with light in living systems. *Nat. Chem. Biol.* **2014**, *10*, 533.
13. Matera, C.; Gomila, A. M. J.; Camarero, N.; Libergoli, M.; Soler, C.; Gorostiza, P. Photoswitchable Antimetabolite for Targeted Photoactivated Chemotherapy. *J. Am. Chem. Soc.* **2018**, *140*, 15764-15773.
14. Vander Heiden, M. G. Targeting cancer metabolism: a therapeutic window opens. *Nature Reviews Drug Discovery* **2011**, *10*, 671-684.
15. Yu, J. K.; Bannwarth, C.; Hohenstein, E. G.; Martínez, T. J. Ab Initio Nonadiabatic Molecular Dynamics with Hole–Hole Tamm–Dancoff Approximated Density Functional Theory. *J. Chem. Theory Comput.* **2020**, *16*, 5499-5511.
16. Bannwarth, C.; Yu, J. K.; Hohenstein, E. G.; Martínez, T. J. Hole–hole Tamm–Dancoff-approximated density functional theory: A highly efficient electronic structure method incorporating dynamic and static correlation. *J. Chem. Phys.* **2020**, *153*, 024110.
17. Ben-Nun, M.; Quenneville, J.; Martínez, T. J. Ab Initio Multiple Spawning: Photochemistry from First Principles Quantum Molecular Dynamics. *J. Phys. Chem. A* **2000**, *104*, 5161-5175.
18. Curchod, B. F. E.; Martínez, T. J. Ab Initio Nonadiabatic Quantum Molecular Dynamics. *Chem. Rev.* **2018**, *118*, 3305-3336.
19. Ben-Nun, M.; Martínez, T. J., Ab Initio Quantum Molecular Dynamics. In *Adv. Chem. Phys.*, John Wiley & Sons, Inc.: 2002; pp 439-512.

20. Yu, J. K.; Bannwarth, C.; Liang, R.; Hohenstein, E. G.; Martínez, T. J. Nonadiabatic Dynamics Simulation of the Wavelength-Dependent Photochemistry of Azobenzene Excited to the $n\pi^*$ and $\pi\pi^*$ Excited States. *J. Am. Chem. Soc.* **2020**, *142*, 20680-20690.
21. Liang, R. First-Principles Nonadiabatic Dynamics Simulation of Azobenzene Photodynamics in Solutions. *J. Chem. Theory Comput.* **2021**, *17*, 3019-3030.
22. Liang, R.; Das, D.; Bakhtiari, A. Protein confinement fine-tunes aggregation-induced emission in human serum albumin. *Phys. Chem. Chem. Phys.* **2021**, *23*, 26263-26272.
23. Mondal, P.; Granucci, G.; Rastädter, D.; Persico, M.; Burghardt, I. Azobenzene as a photoregulator covalently attached to RNA: a quantum mechanics/molecular mechanics-surface hopping dynamics study. *Chem. Sci.* **2018**, *9*, 4671-4681.
24. Mondal, P.; Biswas, M.; Goldau, T.; Heckel, A.; Burghardt, I. In Search of an Efficient Photoswitch for Functional RNA: Design Principles from a Microscopic Analysis of Azobenzene-linker-RNA Dynamics with Different Linkers. *J. Phys. Chem. B* **2015**, *119*, 11275-11286.
25. Xia, S.-H.; Cui, G.; Fang, W.-H.; Thiel, W. How Photoisomerization Drives Peptide Folding and Unfolding: Insights from QM/MM and MM Dynamics Simulations. *Angew. Chem. Int.* **2016**, *55*, 2067-2072.
26. Tully, J. C. Molecular dynamics with electronic transitions. *J. Chem. Phys.* **1990**, *93*, 1061-1071.
27. Bittner, E. R.; Rossky, P. J. Quantum decoherence in mixed quantum-classical systems: Nonadiabatic processes. *J. Chem. Phys.* **1995**, *103*, 8130-8143.

28. Ibele, L. M.; Lassmann, Y.; Martínez, T. J.; Curchod, B. F. E. Comparing (stochastic-selection) ab initio multiple spawning with trajectory surface hopping for the photodynamics of cyclopropanone, fulvene, and dithiane. *J. Chem. Phys.* **2021**, *154*, 104110.
29. Wang, L.; Qiu, J.; Bai, X.; Xu, J. Surface hopping methods for nonadiabatic dynamics in extended systems. *WIREs Computational Molecular Science* **2020**, *10*, e1435.
30. Cody, V.; Luft, J. R.; Pangborn, W. Understanding the role of Leu22 variants in methotrexate resistance: comparison of wild-type and Leu22Arg variant mouse and human dihydrofolate reductase ternary crystal complexes with methotrexate and NADPH. *Acta Crystallographica Section D* **2005**, *61*, 147-155.
31. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theo. Comp.* **2015**, *11*, 3696-3713.
32. Dickson, C. J.; Madej, B. D.; Skjevik, Å. A.; Betz, R. M.; Teigen, K.; Gould, I. R.; Walker, R. C. Lipid14: The Amber Lipid Force Field. *J. Chem. Theo. Comp.* **2014**, *10*, 865-879.
33. Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785-789.
34. Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **1993**, *97*, 10269-10280.
35. Trott, O.; Olson, A. J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31*, 455-461.

36. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157-1174.
37. Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.* **2006**, *25*, 247-260.
38. Wu, Y. J.; Tepper, H. L.; Voth, G. A. Flexible simple point-charge water model with improved liquid-state properties. *J. Chem. Phys.* **2006**, *124*.
39. Eastman, P.; Friedrichs, M. S.; Chodera, J. D.; Radmer, R. J.; Bruns, C. M.; Ku, J. P.; Beauchamp, K. A.; Lane, T. J.; Wang, L.-P.; Shukla, D.; et al. OpenMM 4: A Reusable, Extensible, Hardware Independent Library for High Performance Molecular Simulation. *J. Chem. Theory Comput.* **2013**, *9*, 461-469.
40. Ufimtsev, I. S.; Martinez, T. J. Quantum Chemistry on Graphical Processing Units. 3. Analytical Energy Gradients, Geometry Optimization, and First Principles Molecular Dynamics. *J. Chem. Theory Comput.* **2009**, *5*, 2619-2628.
41. Titov, A. V.; Ufimtsev, I. S.; Luehr, N.; Martinez, T. J. Generating Efficient Quantum Chemistry Codes for Novel Architectures. *J. Chem. Theory Comput.* **2013**, *9*, 213-221.
42. Seritan, S.; Bannwarth, C.; Fales, B. S.; Hohenstein, E. G.; Kokkila-Schumacher, S. I. L.; Luehr, N.; Snyder, J. W.; Song, C.; Titov, A. V.; Ufimtsev, I. S.; et al. TeraChem: Accelerating electronic structure and ab initio molecular dynamics with graphical processing units. *J. Chem. Phys.* **2020**, *152*, 224110.
43. Seritan, S.; Bannwarth, C.; Fales, B. S.; Hohenstein, E. G.; Isborn, C. M.; Kokkila-Schumacher, S. I. L.; Li, X.; Liu, F.; Luehr, N.; Snyder Jr, J. W.; et al. TeraChem: A graphical processing unit-accelerated electronic structure package for large-scale ab initio molecular dynamics. *Wiley Interdiscip. Rev. Comput. Mol. Sci* **2021**, *11*, e1494.

44. Barbatti, M.; Aquino, A. J. A.; Lischka, H. The UV absorption of nucleobases: semi-classical ab initio spectra simulations. *Phys. Chem. Chem. Phys.* **2010**, *12*, 4959-4967.
45. Curchod, B. F. E.; Glover, W. J.; Martínez, T. J. SSAIMS—Stochastic-Selection Ab Initio Multiple Spawning for Efficient Nonadiabatic Molecular Dynamics. *J. Phys. Chem. A* **2020**, *124*, 6133-6143.
46. Yarkony, D. R. Nuclear dynamics near conical intersections in the adiabatic representation: I. The effects of local topography on interstate transitions. *J. Chem. Phys.* **2001**, *114*, 2601-2613.
47. Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules .1. The Method. *J. Comput. Chem.* **1992**, *13*, 1011-1021.
48. Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. Multidimensional Free-Energy Calculations Using the Weighted Histogram Analysis Method. *J. Comput. Chem.* **1995**, *16*, 1339-1350.
49. Aleotti, F.; Soprani, L.; Nenov, A.; Berardi, R.; Arcioni, A.; Zannoni, C.; Garavelli, M. Multidimensional Potential Energy Surfaces Resolved at the RASPT2 Level for Accurate Photoinduced Isomerization Dynamics of Azobenzene. *J. Chem. Theory Comput.* **2019**, *15*, 6813-6823.
50. Granucci, G.; Persico, M. Excited state dynamics with the direct trajectory surface hopping method: azobenzene and its derivatives as a case study. *Theor. Chem. Acc.* **2007**, *117*, 1131-1143.

51. Ciminelli, C.; Granucci, G.; Persico, M. The Photoisomerization Mechanism of Azobenzene: A Semiclassical Simulation of Nonadiabatic Dynamics. *Chem. Eur. J* **2004**, *10*, 2327-2341.
52. Liu, L.; Liu, J.; Martinez, T. J. Dynamical Correlation Effects on Photoisomerization: Ab Initio Multiple Spawning Dynamics with MS-CASPT2 for a Model trans-Protonated Schiff Base. *J. Phys. Chem. B* **2016**, *120*, 1940-1949.
53. Kingsland, A.; Samai, S.; Yan, Y.; Ginger, D. S.; Maibaum, L. Local Density Fluctuations Predict Photoisomerization Quantum Yield of Azobenzene-Modified DNA. *J. Phys. Chem. Lett.* **2016**, *7*, 3027-3031.