

RESEARCH ARTICLE | JUNE 23 2022

Multiscale simulation unravels the light-regulated reversible inhibition of dihydrofolate reductase by phototrexate

Special Collection: [2022 JCP Emerging Investigators Special Collection](#)

Ruibin Liang  ; Amirhossein Bakhtiari

 Check for updates

J. Chem. Phys. 156, 245102 (2022)

<https://doi.org/10.1063/5.0096349>


View
Online


Export
Citation

 CrossMark

Articles You May Be Interested In

3D-QSAR Study of 7,8-Dialkyl-1,3-diaminopyrrolo-[3,2-f] Quinazolines with Anticancer Activity as DHFR Inhibitors

Chinese Journal of Chemical Physics (June 2009)

Conformational stability of Met20 loop of DHFR: A molecular dynamics study

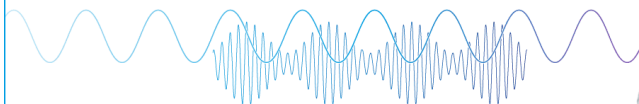
AIP Conference Proceedings (February 2013)

Screening of polymorphisms for MTHFR and DHFR genes in spina bifida children and their mothers

AIP Conference Proceedings (November 2013)

Webinar

Boost Your Signal-to-Noise
Ratio with Lock-in Detection



Sep. 7th – Register now

 Zurich
Instruments

Multiscale simulation unravels the light-regulated reversible inhibition of dihydrofolate reductase by phototrexate

Cite as: J. Chem. Phys. 156, 245102 (2022); doi: 10.1063/5.0096349

Submitted: 18 April 2022 • Accepted: 31 May 2022 •

Published Online: 23 June 2022



View Online



Export Citation



CrossMark

Ruibin Liang^{a)}  and Amirhossein Bakhtiari

AFFILIATIONS

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

Note: This paper is part of the 2022 JCP Emerging Investigators Special Collection.

^{a)} Author to whom correspondence should be addressed: rliang@ttu.edu

ABSTRACT

Molecular photoswitches are widely used in photopharmacology, where the biomolecular functions are photo-controlled reversibly with high spatiotemporal precision. Despite the success of this field, it remains elusive how the protein environment modulates the photochemical properties of photoswitches. Understanding this fundamental question is critical for designing more effective light-regulated drugs with mitigated side effects. In our recent work, we employed first-principles non-adiabatic dynamics simulations to probe the effects of protein on the *trans* to *cis* photoisomerization of phototrexate (PTX), a photochromic analog of the anticancer therapeutic methotrexate that inhibits the target enzyme dihydrofolate reductase (DHFR). Building upon this study, in this work, we employ multiscale simulations to unravel the full photocycle underlying the light-regulated reversible inhibition of DHFR by PTX, which remains elusive until now. First-principles non-adiabatic dynamics simulations reveal that the *cis* to *trans* photoisomerization quantum yield is hindered in the protein due to backward isomerization on the ground-state following non-adiabatic transition, which arises from the favorable binding of the *cis* isomer with the protein. However, free energy simulations indicate that *cis* to *trans* photoisomerization significantly decreases the binding affinity of the PTX. Thus, the *cis* to *trans* photoisomerization most likely precedes the ligand unbinding from the protein. We propose the most probable photocycle of the PTX-DHFR system. Our comprehensive simulations highlight the trade-offs among the binding affinity, photoisomerization quantum yield, and the thermal stability of the ligand's different isomeric forms. As such, our work reveals new design principles of light-regulated drugs in photopharmacology.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0096349>

INTRODUCTION

Molecular photoswitches are playing an ever-increasing role in biomedical applications.^{1–10} They typically undergo photoisomerization in opposite directions (e.g., *cis* to *trans* isomer and vice versa) when induced by different wavelengths of light, allowing reversible control of their conformations. Taking advantage of this unique property, a variety of photoswitches have been designed to interact with and control the function of biomolecules in a light-regulated and reversible fashion.^{1–10} The essence of their design principle is to maximize the difference in the interaction with the target biomolecule between different isomers of the photoswitch such that photoisomerization results in a significant change in the biomolecular activity. The key advantage of this design strategy is

the reversible control of the biomolecular activity with high spatial and temporal resolution using light, which opens new ways to design light-regulated drugs that can mitigate the side effect of conventional chemotherapeutics. For instance, dihydrofolate reductase (DHFR) is an important drug target to treat cancer.¹¹ Methotrexate (MTX) is a widely prescribed anticancer drug that inhibits the DHFR in the tumor cells. However, the ubiquity of the DHFR in all cell types makes the off-target toxicity of MTX unavoidable, which causes unbearable side effects, such as ulcerative stomatitis, nausea, and abdominal distress. To overcome this issue, phototrexate (PTX) was designed as a photoswitchable analog of MTX to reversibly inhibit DHFR under the control of light (Fig. 1).⁹ The *cis* isomer of PTX binds DHFR with a higher affinity than the *trans* isomer. The *cis* → *trans* and *trans* → *cis* photoisomerization can be triggered by

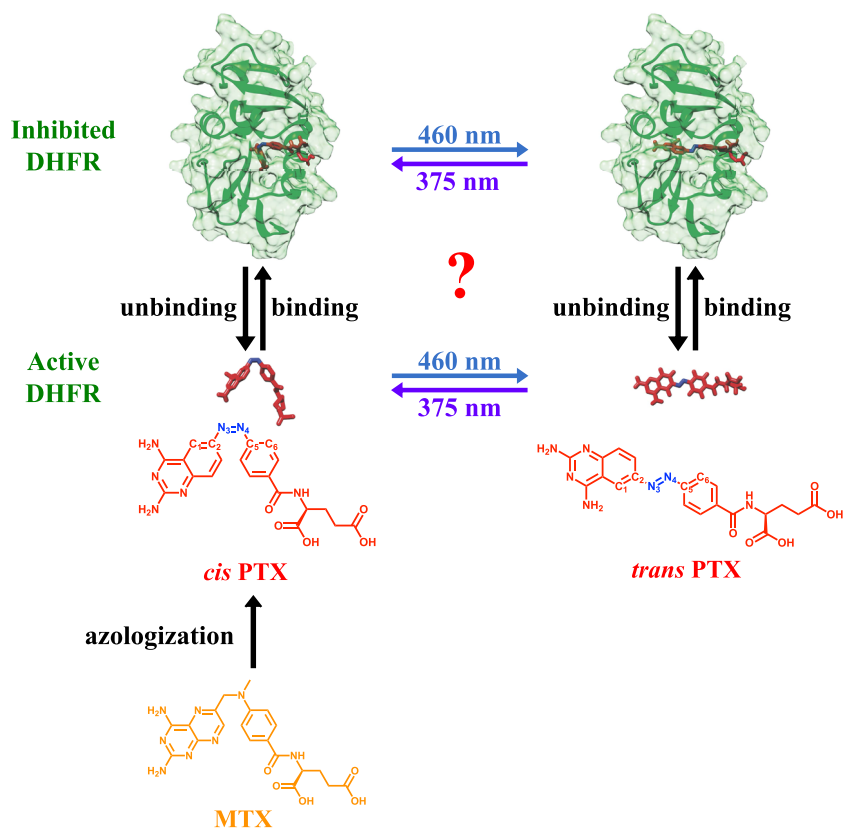


FIG. 1. The photoisomerization of phototrexate (PTX, red) enables its reversible inhibition of dihydrofolate reductase (DHFR) under illumination by different wavelengths of light. The chemical structures of the *cis* and *trans* PTX are designed based on the clinically prescribed anticancer drug methotrexate MTX (orange) via the azologization strategy, which introduces the N=N double bond (blue). The atoms involved in the torsional motion during the photoisomerization of PTX are labeled in the chemical structures, which are followed throughout the main text. The photocycle of the PTX-DHFR system remains elusive with regard to how the ligand binding/unbinding events are coupled with the bidirectional photoisomerization reactions, which is the focus of this work.

460 and 375 nm light, respectively. Thus, the difference in the binding affinities and excitation wavelengths of the two isomers enables light-regulated reversible inhibition of DHFR (Fig. 1).

Despite the recent progress in photopharmacology, the intricate interplay between the ligand–protein and light–matter interactions remains largely elusive in general. For instance, if the active state of a photoswitchable ligand is the *cis* isomer, the ligand binding could, in principle, occur before or after the *trans* → *cis* photoisomerization. The same applies to the ligand unbinding event and the photoisomerization in the reverse direction. Understanding the sequence of ligand binding/unbinding and the bi-directional photoisomerization events will depict the full photocycle of photopharmacological systems (Fig. 1), which will benefit the rational design of effective light-regulated therapeutics. To the best of our knowledge, little effort has been devoted to understanding the fundamental molecular mechanism underlying these complicated light-induced processes in these artificial photoactive biomolecular systems, and their photocycles remain largely unknown (Fig. 1). Consequently, the experimental design of light-regulated drugs still requires extensive trials and errors.

As our first attempt to fill the gap in this field, in our recent simulation study,¹² we employed first-principles non-adiabatic dynamics simulations to probe the effect of protein on the *trans* → *cis* photoisomerization of PTX induced by 375 nm light. Our study highlighted the effects of the local molecular environment on the kinetics and quantum yield (QY) of this photochemical reaction.¹²

The DHFR was found to slow down the photoisomerization and decrease its QY.¹² As a result, we proposed that the PTX is more likely to bind with the enzyme as the *cis* isomer after the *trans* → *cis* photoisomerization takes place in the aqueous solution.¹² However, what remains elusive is the effect of protein on the *cis* → *trans* photoisomerization induced by 460 nm light. In addition, the previous estimation of PTX's binding affinity with the DHFR⁹ was based on simple docking simulations with the crude approximations that the protein structure remains static and the solvent molecules are absent.⁹ As mentioned above, it is crucial to accurately estimate the difference in the protein–ligand interactions between the two isomers, and therefore, it is necessary to incorporate the entropic and solvent effects in the binding free energy calculations. These additional pieces of information are indispensable to unraveling the photocycle of the PTX-DHFR system, which will reveal new design principles for light-regulated and reversible enzyme inhibition in general.

To this end, in this work, we combine multiscale simulation tools to (1) probe the effects of protein on the mechanism of *cis* → *trans* photoisomerization, (2) quantify the binding affinities of the *cis* and *trans* isomers of the PTX with the enzyme, and (3) depict the full photocycle of the PTX-DHFR system building upon our previous study.¹² Our multiscale simulations integrate the *ab initio* multiple spawning (AIMS) simulations^{13–15} and free energy calculations using enhanced sampling techniques. The AIMS simulations characterize the photodynamics of PTX in solution and

protein, while the free energy calculations quantify PTX's binding affinities in different isomer forms. Of particular importance in the AIMS simulation is the on-the-fly first-principles electronic structure calculation using the hole-hole Tamm-Dancoff Approximated Density Functional Theory (hh-TDA-DFT) during the propagation of the multistate nuclear wave-packets. The hh-TDA-DFT method incorporates both static and dynamic electron correlations that are necessary for an accurate description of potential energy surfaces (PES) and non-adiabatic couplings (NAC). The main advantage of this new method is the capability to efficiently and accurately predict both the absorption/emission spectra and the topography of the conical intersections (CI) between the ground and excited states.^{12,16–20} These advantages of the hh-TDA-DFT method are crucial for a correct description of the photodynamics of the azobenzene-derived photoswitches. Moreover, the interaction between the PTX with its environment is treated with the quantum mechanics/molecular mechanics (QM/MM) approach, which enables efficient “on-the-fly” calculation of the PES of large biomolecular systems during the non-adiabatic dynamics simulation. In addition, the AIMS algorithm^{13–15} efficiently propagates the coupled nuclear–electronic degrees of freedom on multiple electronic states according to the time-dependent Schrödinger's equation. Thus, the coupling of the hh-TDA-DFT/MM and AIMS approaches offers a powerful state-of-the-art computational tool to investigate the photodynamics of PTX in DHFR.

Our results reveal that the protein environment slightly accelerates the non-radiative decay induced by the *cis* → *trans* photoisomerization but significantly reduces the isomerization QY by reshaping the ground-state PES near the transition state. Also, our free energy calculations indicate a much larger binding affinity of the *cis* isomer than the *trans* isomer of the PTX. These new insights complete the photocycle of the PTX-DHFR system and highlight the trade-offs among the ligand-binding affinity, photoisomerization QY, and thermal stability. As such, our work provides new insights into the design principles of molecular photoswitches as light-regulated therapeutics.

METHOD

Except for a few differences, the simulation methods we employed for studying the *cis* → *trans* photoisomerization of PTX

are closely parallel to those in our recent study for its *trans* → *cis* photoisomerization. Thus, here, we only briefly summarize our methods with a focus on the new aspects of the current work and refer readers to Ref. 12 for more detailed information.

System setup, force field parameterization, benchmark calculation, and classical MD equilibration

First, a docking simulation using the Autodock Vina software²¹ was performed to identify the binding pose of the *cis* PTX in the crystal structure of DHFR (PDB code: 1U72).²² The top score structure of *cis* PTX was selected, and its binding mode in the DHFR largely overlaps with that of the MTX in the original crystal structure [Fig. 2(b)]. Following the docking simulation, the protein–ligand complex was solvated in a box of water molecules, resulting in a periodic boundary condition (PBC) of $96 \times 101 \times 106 \text{ \AA}^3$ [Fig. 2(a)]. We also set up an addition simulation where the PTX is solvated in a PBC box of water molecules ($55 \times 58 \times 55 \text{ \AA}^3$) to investigate the effect of protein on the *cis* → *trans* photoisomerization as well as ground state isomerization.

The Amber ff14SB^{23,24} and SPC/Fw²⁵ force fields were used to model the protein and water molecules, respectively. For the PTX, point charges and other parameters were derived using the general Amber force field (GAFF) procedure,^{26,27} as detailed in Ref. 12, except for the torsional terms of three dihedral angles: $C_2-N_3=N_4-C_5$ (θ_{CNNC}), $C_1=C_2-N_3=N_4$ (θ_{CCNN}), and $N_3=N_4-C_5=C_6$ (θ_{NNCC}) (see Fig. 1 for the definition of atom labels and Fig. S1 for torsion labels). The parameterization of these three dihedral angles needs special attention for the ground-state free energy calculations of the *cis* to *trans* isomerization process (see below) because it was found in Ref. 12 that, during the MD equilibration, the *trans* PTX sometimes deviates significantly from the planar geometry, an obvious artifact created by the default force field.

To parameterize each of the three torsions, the hh-TDA-DFT method using the B3LYP functional²⁸ and 6-31G* basis set (abbreviated as hh-TDA-B3LYP/6-31G* below) was used to optimize the PTX in the vacuum on the ground state while constraining the torsion at different values, generating a relaxed scan of the PES as a function of the torsion. Then, the force field parameters for this torsion were fitted to reproduce the barrier heights and the energy

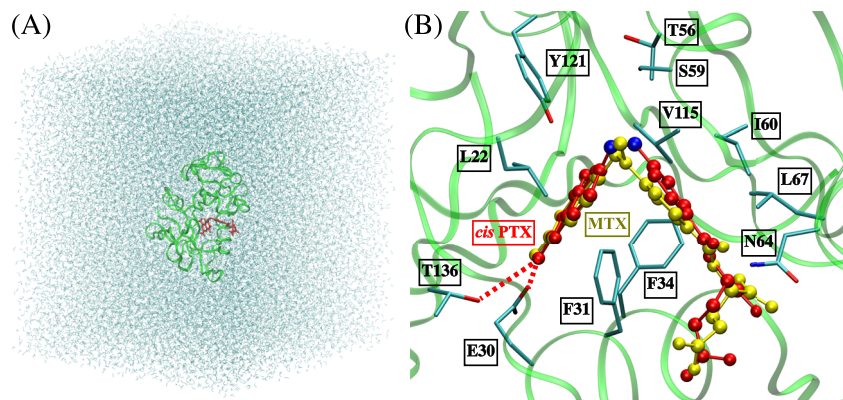


FIG. 2. (a) System setup of the *cis* PTX-DHFR complex. The PTX, protein, and water molecules are colored in red, green, and blue, respectively. (b) The binding mode of the PTX (red stick and balls) at the active site of the DHFR, compared to the binding mode of MTX (yellow stick and balls) in the crystal structure (PDB code: 1U72). The protein is rendered as green ribbons. The key residues surrounding the PTX are rendered as sticks with black labels. The PTX forms hydrogen bonds with the E30 and T136 residues, which are labeled as red dashed lines.

difference between the *cis* and *trans* minima in the hh-TDA-B3LYP PES (Fig. S1).

The accuracy of the hh-TDA-B3LYP method was benchmarked against the multistate complete active space second-order perturbation theory (MS-CASPT2) method,²⁹ which is a high-level post-Hartree-Fock method that includes both static and dynamic electron correlations (Fig. S2). First, a relaxed scan of the PTX's ground-state PES was performed in the vacuum around the θ_{CNC} torsion using the state-averaged complete active space self-consistent field (SA-CASSCF) method. The SA-CASSCF calculation included the S_0 , S_1 , and S_2 states using an active space of 12 electrons and 12 orbitals with the 6-31G* basis set [i.e., SA-3-CASSCF(12e, 12o)/6-31G*]. The active space includes 5 π orbitals, 1 n orbital, and 6 π^* orbitals (Fig. S3). Then, on the SA-CASSCF optimized geometries, the MS-CASPT2 calculations were performed using the reference wavefunction obtained from CASSCF calculation with a larger active space (14 electrons and 13 orbitals) and basis set (ANO-L-VDZP). The active space includes 6 π orbitals, 1 n orbital, and 6 π^* orbitals (Fig. S4). The MS-CASPT2 included the S_0 , S_1 , and S_2 states with an IPEA shift of 0 a.u. and imaginary shift of 0.2 a.u. The hh-TDA-B3LYP/6-31G* calculations were performed on the same SA-CASSCF optimized geometries. The MS-CASPT2 calculations were performed with the OPENMOLCAS package,³⁰ and the gas-phase hh-TDA-DFT calculations were performed with the TERACHEM^{31–34} software package.

After the parameterization of the force fields, the following equilibration procedure was applied to both the DHFR and aqueous solution systems. First, geometry optimization was performed for 10 000 steps. Then, a 100 ps MD simulation was performed in the constant NVT ensemble at 300 K temperature. Subsequently, a 110 ns MD simulation was performed in the constant NPT ensemble at 300 K temperature and 1 atm pressure. A Langevin thermostat with a friction coefficient of 1 ps⁻¹ and a Monte Carlo barostat with a frequency of 0.1 ps⁻¹ were used to maintain the temperature and pressure of the system, respectively. A time step of 1 fs was used to integrate the equation of motion of the system. The first 10 ns trajectory was treated as equilibration and the last 100 ns trajectory as a production run.^{23–27}

Ground state QM/MM MD simulation and absorption spectra calculation

For each system (protein or solution), 100 snapshots were selected from the 100 ns classical MD trajectory with a 1 ns time interval. For each snapshot of the protein system, the protein, PTX ligand, and water molecules having at least one atom within 10 Å of any protein atoms were retained, and the rest of the system was discarded, resulting in 100 subsystems with open boundary conditions. For the snapshots of the solution system, the PTX and water molecules having at least one atom within 10 Å of any PTX atoms were retained. The construction of these subsystems was due to the limitations of the TERACHEM software package in treating the PBC. Then, 100 ground-state QM/MM MD simulations were initiated from the coordinates of these snapshots at 300 K temperature. Each of them was propagated in the constant NVT ensemble at the same temperature for 5 ps in order to sample the conformations of the *cis* PTX. Thus, the QM/MM sampling incorporated the conformational changes of the PTX and its surrounding environment

over 100 ns. The QM region included the PTX molecule and was treated with DFT using B3LYP/6-31G*.²⁸ The MM region included the rest of the system and was treated with the same force fields as the MD equilibration. Based on the snapshots from the 100 ground-state QM/MM MD trajectories, the absorption spectra of the *cis* PTX in the aqueous solution were calculated using the hh-TDA-B3LYP/MM method, following the procedure detailed in Ref. 12.

AIMS simulations

For each system (DHFR or solution), the AIMS simulations were initiated from the 100 initial conditions (ICs) taken from the last snapshots of the QM/MM MD trajectories, with all trajectory basis functions (TBFs) initially placed on the S_1 state. For each IC, an AIMS simulation was propagated for 250 fs or until the depletion of all excited-state populations. The potential energies, gradients, and non-adiabatic couplings were calculated on-the-fly using the hh-TDA-B3LYP/MM method with the same aforementioned partitioning of the QM and MM regions. The S_1 state's relaxation time constants (τ_0 and τ_1) were obtained by fitting the time evolution of the S_1 state's population $P_{S_1}(t)$ with the delayed exponential function,

$$P_{S_1}(t) = \begin{cases} 1, & t < \tau_0, \\ \exp\left(-\frac{t - \tau_0}{\tau_1}\right), & t \geq \tau_0. \end{cases} \quad (1)$$

The photoisomerization quantum yield (QY) was calculated based on the centroid geometries and populations of the final S_0 state TBFs, which were obtained by the ground-state optimization following the AIMS simulation (see Ref. 12 for details). If the absolute value of the final θ_{CNC} torsion was in the range of 0–45°, the population of this TBF was assigned to *cis* isomer. If the torsion angle was in the range of 135–225°, the population was assigned to *trans* isomer. The QY for the *cis* → *trans* isomerization was defined as the ratio of the total population of the *cis* isomer to the total population of all S_0 state TBFs. The error bars of the $P_{S_1}(t)$, τ_0 , τ_1 , and QYs were estimated by bootstrapping over 1000 random samples with replacement.

Excited-state QM/MM free energy calculation

First-principles QM/MM umbrella sampling simulations were performed to characterize the potential of mean force (PMF) of the PTX's *cis* → *trans* photoisomerization on the S_1 excited state for each of the two systems (protein and solution). The central torsion θ_{CNC} was chosen as the collective variable (CV) of the PMF. The umbrella windows spanned the CV range of 0°–100° with a 10° interval, and in each window, the θ_{CNC} was restrained to the window center using a harmonic potential with 100 kcal mol⁻¹/rad² force constant. For each window, the conformational space of the system was sampled using constant NVT MD simulation at 300 K temperature for ~5 ps, resulting in ~110 ps total QM/MM sampling time. The QM/MM partitioning and the level of theory were the same as the AIMS simulation. The PMFs were calculated by unbiasing the distributions of the CV using the WHAM algorithm.^{35,36} The error bars of the PMFs were estimated by block-average analysis.

Ground-state QM/MM PES scan near the transition state

To test the effect of the environment on the local ground-state PES near the transition state, we performed QM/MM relaxed scan of the PES starting from spawning geometries of 5 S_0 state TBFs in both the DHFR and the aqueous solution. The QM/MM partitioning and the level of theory were the same as the AIMS simulation. In the relaxed PES scan, only the PTX was allowed to relax with the θ_{CNNC} constrained at different values ranging from 70° to 110° , and the rest of the system was kept frozen at the spawning geometry for both the DHFR and aqueous solution. Freezing the environment during the optimization was justifiable in this case because the success or failure of photoisomerization was largely determined by the 20–40 fs ground-state dynamics following the spawning event. During this short period, the kinetic energy from the non-radiative decay had not fully dissipated to the environment, and the environment had not fully relaxed in response to the fast torsional dynamics of the PTX.

Free energy calculation of ground-state isomerization and unbinding of *trans* PTX

Umbrella sampling simulations using molecular mechanics (MM) force fields were employed to characterize the ground-state PMFs of the *trans* to *cis* isomerization and the unbinding of *trans* PTX from the DHFR. The CV of the isomerization PMF was chosen as the θ_{CNNC} torsion. In total, 37 umbrella windows were simulated, covering the θ_{CNNC} from 0° to 180° with a 5° interval. The force constant of the harmonic biasing potential in each window was $100 \text{ kcal mol}^{-1}/\text{rad}^2$. The initial conformation of the umbrella sampling was chosen as the spawning geometry of an S_0 TBF, where the θ_{CNNC} torsion is closed to 90° . For each window, the initial structure was

generated by biasing the CV from the adjacent window, followed by 1 ns equilibration. The production run for each window was 20–30 ns, depending on the convergence of the CV distribution. The isomerization PMFs were calculated in both the DHFR and the aqueous solution. In total, there was 750–800 ns enhanced sampling for each system. The PMFs were calculated by unbiasing the distributions of the CV using the WHAM algorithm.^{35,36} The error bars of the PMFs were estimated by block-average analysis.

The initial structure for the unbinding PMF calculation was obtained from the equilibrated structure in an isomerization umbrella window centered at 180° (i.e., the *trans* PTX). The CV for the unbinding PMF was chosen as the distance between the center of mass (COM) of the PTX and that of the DHFR. In total, 27 umbrella windows were simulated, covering a CV range from 11 to 37 Å with 1 Å spacing. For each window, the force constant of the harmonic biasing potential was $10 \text{ kcal mol}^{-1}/\text{\AA}^2$, and the initial structure was generated by biasing the CV from the adjacent window followed by 1 ns equilibration. The production run for each window was 100–150 ns, depending on the convergence of the distribution of the CV. The total sampling time was thus $\sim 3 \mu\text{s}$. The PMFs were calculated by unbiasing the distributions of the CV using the WHAM algorithm.^{35,36} The error bars of the PMFs were estimated by block-average analysis.

All unbiased MM simulations were performed with the OPENMM software package.³⁷ The QM/MM simulations, including ground-state dynamics, absorption spectra, and umbrella sampling, were performed with the TERACHEM^{31–34} interfaced with the OPENMM packages.³⁷ The AIMS simulations were performed using the FMS90 code interfaced with the TERACHEM/OPENMM software packages.^{31–34,37} All MM umbrella sampling simulations were performed using the AMBER20 software package³⁸ interfaced with the PLUMED plugin version 2.6.^{39,40}

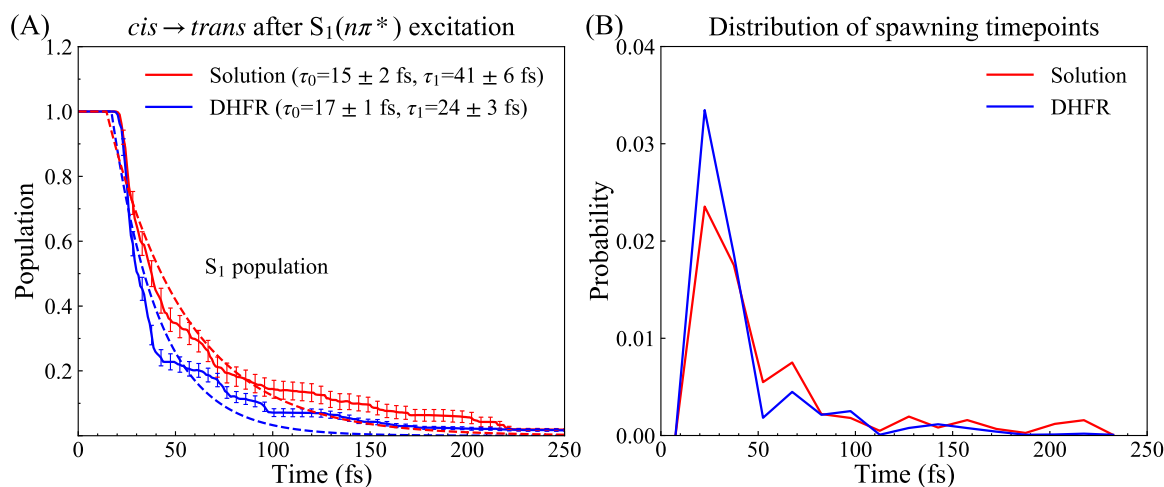


FIG. 3. (a) The excited-state population decay is faster in the DHFR than in the aqueous solution. The time evolutions of the S_1 state's population during the PTX's *cis* \rightarrow *trans* photoisomerization in the aqueous solution and DHFR are indicated by the red and blue lines, respectively. The photoisomerization is induced by $S_1(n\pi^*)$ excitation. The time constants for the population decay (τ_0 and τ_1) are obtained by fitting to a delayed exponential function [Eq. (1)], and the fits are indicated by the dashed lines. The error bars of the population decay curves and time constants are obtained by bootstrapping over 1000 random samples of the 100 initial conditions of the AIMS simulation. (b) The distribution of spawning timepoints of the S_0 TBFs weighted by their eventual populations. In the DHFR, the distribution before 50 fs is more peaked than the solution, indicating sooner non-radiative transition in the former environment than in the latter.

RESULTS AND DISCUSSION

Fast excited-state population decay driven by steep free energy surface

In the experiment, the *cis* → *trans* photoisomerization of PTX was induced by 460 nm light.⁹ This wavelength is in close agreement with the simulated maximum absorption wavelength of the $S_0 \rightarrow S_1$ excitation dominated by $n\pi^*$ transition (Fig. S5). Therefore, the AIMS simulations of this photoisomerization reaction were initiated on the S_1 state. As illustrated in Fig. 3(a), more than 95% S_1 population has decayed to the ground state within 250 fs in both DHFR and aqueous solution. Fitting the time evolution of the S_1 population to a delayed exponential function [Eq. (1)], the relaxation time constant $\tau = \tau_0 + \tau_1$ is 56 ± 8 fs in solution and 41 ± 4 fs in the DHFR. In the initial ~ 20 fs of the isomerization, the $\theta_{C_{NNC}}$ torsion starts to increase, but during this time, the PTX has not reached the CI where the $\theta_{C_{NNC}}$ is close to 90° . Consequently, the non-adiabatic coupling is too small to trigger non-radiative decay. This initial phase has a similar time constant τ_0 in the two systems (15 ± 2 fs in solution vs 17 ± 1 in DHFR). However, the τ_1 in the DHFR (24 ± 3 fs) is smaller than the solution (41 ± 6 fs), indicating a faster non-adiabatic transition near the CI. To further understand the difference in the τ_1 between the two systems, we analyzed the distribution of the spawning timepoints of the S_0 state TBFs weighted by their eventual populations [Fig. 3(b)]. This distribution measures how soon the S_1 TBFs start to significantly transfer their population to the S_0 state after photoexcitation. The distribution before 50 fs is narrower in the DHFR than in the solution. This indicates that in the DHFR the dominant spawning events contributing to the $S_1 \rightarrow S_0$ decay mostly finish before 50 fs, whereas in the solution, some important spawning events are slightly delayed.

The calculated S_1 state PMFs [Fig. 4(a)] as a function of the $\theta_{C_{NNC}}$ are in line with the kinetics of the non-radiative decay. Compared to the *trans* → *cis* isomerization,¹² the S_1 state PMF's slope is much larger for the *cis* → *trans* isomerization. Starting from the Frank–Condon (FC) region of the *cis* isomer, the free energy decreases by more than 21 kcal/mol upon reaching the CI, where the $\theta_{C_{NNC}}$ is close to 90° . In comparison, starting from the FC region of the *trans* isomer, the free energy decreases by no more than 8 kcal/mol when reaching the same type of CI.¹² The different slopes of the S_1 PMFs between the two isomerization directions lead to a large difference in the excited-state lifetime. The S_1 state lifetime is ~ 3 to 5 ps for *trans* → *cis* isomerization induced by S_2 ($\pi\pi^*$) excitation,¹² and it would be longer if the isomerization was induced by S_1 ($n\pi^*$) excitation, as observed for the azobenzene photodynamics.¹⁹ In contrast, the S_1 state lifetime is only 40–60 fs for *cis* → *trans* isomerization, significantly shorter than the *trans* → *cis* isomerization. Thus, a larger slope of the PMF leads to faster non-radiative decay.

In addition, for the *cis* → *trans* isomerization, the PMF is slightly steeper in the DHFR than in the solution, leading to slightly faster isomerization in the former environment. The distribution of the velocity of $\theta_{C_{NNC}}$ at the $S_1 \rightarrow S_0$ spawning timepoint [Fig. 4(b)] is in line with the slopes of PMFs in the two environments. The DHFR shifts the velocity distribution toward the positive direction relative to the one in the solution, and the positive velocity of $\theta_{C_{NNC}}$ favors the formation of the *trans* isomer photoproduct. Moreover, the distributions of the initial values of $\theta_{C_{NNC}}$ are similar in the two

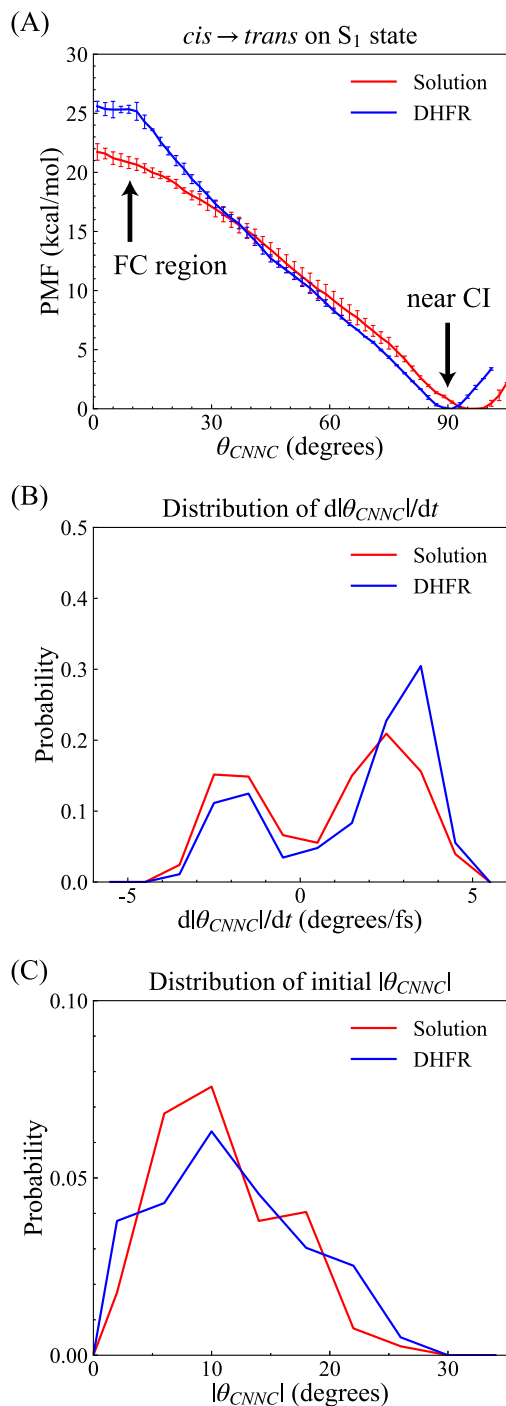


FIG. 4. The DHFR's protein environment facilitates the rotation of the $\theta_{C_{NNC}}$ on the S_1 state from the FC region ($\sim 0^\circ$) to the proximity of the CI ($\sim 90^\circ$). (a) The PMF for the isomerization of the $\theta_{C_{NNC}}$ in the DHFR (blue) and the aqueous solution (red). (b) The probability distribution of the velocity of the absolute value of $\theta_{C_{NNC}}$ at the spawning time of the S_0 TBFs. (c) The initial distributions of the absolute value of the $\theta_{C_{NNC}}$ are similar between the DHFR and solution. The steeper slope in the S_1 state PMF in the DHFR translates to a higher probability of more positive torsional velocity during the $S_1 \rightarrow S_0$ non-adiabatic transition than the solution.

TABLE I. Isomerization quantum yields (QY) for the *cis* → *trans* photoisomerization.

System	QY
DHFR	0.28 ± 0.04
Solution	0.51 ± 0.04

systems [Fig. 4(c)]. Thus, our analysis indicates that in the DHFR the steeper S_1 state PMF allows the θ_{CNNC} torsional mode to accumulate more momentum toward forming the *trans* isomer as the PTX travels from the FC region to the CI, compared to the case in the solution. Therefore, the protein environment slightly accelerates the *cis* → *trans* photoisomerization. This is in contrast to the *trans* → *cis* isomerization, which is hindered by the protein environment.¹²

Isomerization quantum yield influenced by ground-state dynamics

The *cis* → *trans* isomerization QYs are summarized in Table I. To our surprise, the QY is lower in the DHFR (0.28 ± 0.04) than the solution (0.51 ± 0.04), despite the more positive velocity of θ_{CNNC} at the spawning time. This interesting phenomenon can be explained by the ground-state dynamics immediately following the non-radiative decay. Compared to the solution [Fig. 5(a)], in the DHFR, it is more likely for an S_0 TBF spawned with a positive θ_{CNNC} torsional velocity to reverse its isomerizing direction and end up in the ground-state *cis* minimum [Fig. 5(b)].

The more pronounced reversal of the θ_{CNNC} 's velocity in the DHFR is due to the S_0 state PES of the PTX being reshaped near its intrinsic transition state by the protein environment, which is illustrated in Fig. 6. It is evident that the protein environment makes it energetically uphill for the θ_{CNNC} to continue increasing on the

S_0 state [Fig. 6(b)], whereas, in the solution, the increase of θ_{CNNC} beyond 100° starts to become energetically downhill. The reshaping of the local ground-state PES by the protein environment can also be understood from the perspective of ligand binding affinity. Compared to the solution, the *cis* isomer is more stabilized in the DHFR [Fig. 7(a)]. Because the *cis* → *trans* isomerization is ultrafast on a time scale of ~40 to 60 fs, the protein environment cannot respond to the changing geometry of the PTX in time before it gets close to the CI. Right after the non-radiative transition, the protein remains in the conformation that favorably binds the *cis* isomer. In the vacuum, the transition state of isomerization on the ground state features a θ_{CNNC} value near 90°, and the increase of θ_{CNNC} from 90° is energetically downhill (Figs. S1 and S2). However, the ligand-protein interactions reshape the local PES near the transition state, creating a non-negligible energy penalty for the θ_{CNNC} to continue increasing and yield the *trans* photoproduct. Eventually, the θ_{CNNC} 's velocity in some S_0 TBFs is reverted, aborting the photoisomerization and yielding the *cis* reactant. Thus, the local environment at the binding cavity influences the isomerization QY of the PTX.

It is noteworthy that in our interpretation of the ground-state dynamics, we did not carry out the calculation of the *ab initio* PMF. Instead, we used constrained optimization to analyze the local PES near several S_0 TBF's spawning geometries (see Method section). This is because the dynamics are intrinsically non-equilibrium right after the system decays to the ground state. Within 20–40 fs after the non-radiative decay, the protein has not fully relaxed to the change in the θ_{CNNC} , but we observed that during this time the ground-state dynamics of PTX largely determines the success/failure of its photoisomerization. Thus, the system may not closely follow its minimum free energy pathway on the ground state after the non-radiative decay. Also, here, we are mainly interested in the *qualitative* difference in the shapes of the local PES between the two systems, e.g., with or without turning points. Thus, we employed the less

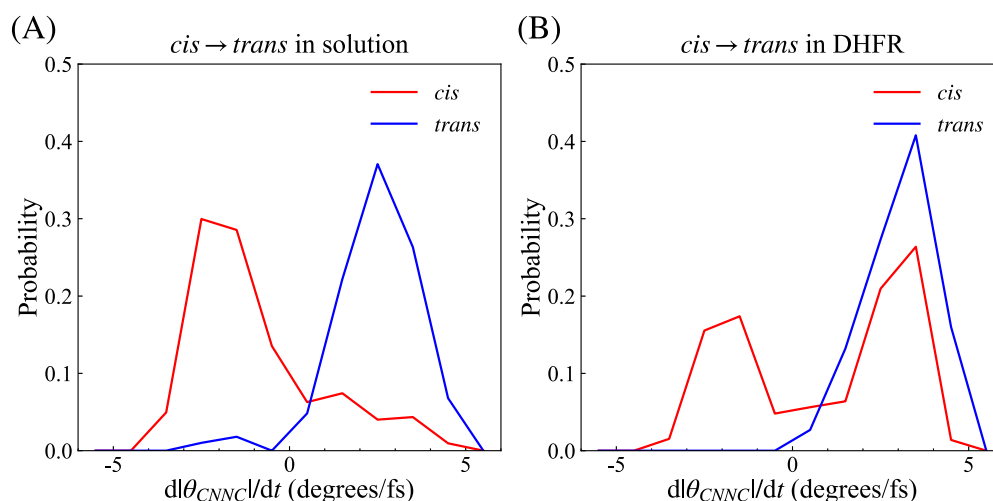


FIG. 5. The probability distribution of the velocity of the absolute value of θ_{CNNC} at the spawning time of the S_0 state TBF in (a) solution and (b) DHFR. The velocity distributions are classified by the final isomer of the S_0 state TBF (blue for *trans* and red for *cis*). The reversal of the isomerization back to the *cis* reactant is more likely in the DHFR than in the solution.

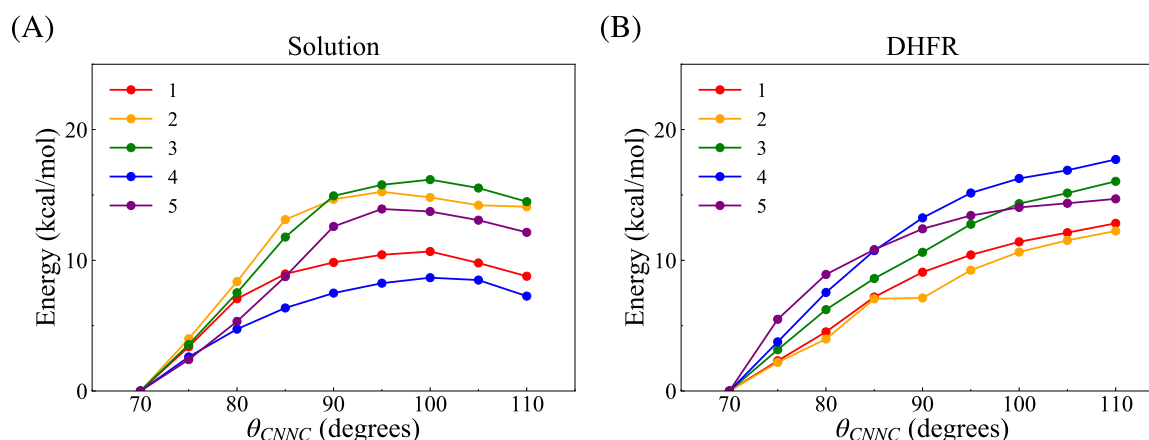


FIG. 6. The ground-state PES of the PTX's isomerization in (a) solution and (b) DHFR. For each system, the potential energy surfaces are obtained by constrained geometry optimization, freezing all protein and solvent atoms in the conformation at the spawning time of the S_0 TBFs as well as the θ_{CNNC} torsion, while allowing the rest degrees of freedoms in the PTX to fully relax. For each system, five different starting structures were used for the relaxed PES scan, resulting in five PES curves. The PES's have turning points in the solution but are monotonically increasing in the DHFR, which indicates that the DHFR imposes a larger energetic penalty for the PTX to continue its isomerization on the ground state after the non-radiative transition.

time-consuming PES analysis rather than the more expensive *ab initio* PMF calculations to interpret the difference in the ground-state dynamics between the two systems.

A careful inspection of the binding site of DHFR reveals the detailed intermolecular interactions between the protein and PTX [Fig. 2(b)]. The *cis* PTX forms hydrogen bonds with the E30 and T136 residues. Also, the rings of the diaminopteridine and benzamido moieties in the *cis* PTX surround the F31 and F34 residues. Thus, the PTX's hydrogen bonding with polar side chains and van der Waals interactions with the hydrophobic side chains of DHFR stabilizes its *cis* isomer, creating a local environment that kinetically disfavors the complete isomerization to yield the *trans* isomer.

Effects of photoisomerization on the binding affinity of PTX

To quantify how much the photoisomerization affects the binding affinity of the PTX, we constructed the thermodynamic cycle [Fig. 7(a)] and employed molecular mechanics force fields to calculate the PMF of the *cis* \leftrightarrow *trans* isomerization of the PTX on the ground state in both environments [Fig. 7(b)].

The difference in the binding affinity of the *trans* and *cis* isomers can be calculated as

$$\begin{aligned} \Delta\Delta G_{\text{binding}} &= \Delta G_{\text{binding,cis}} - \Delta G_{\text{binding,trans}} \\ &= \Delta G_{\text{cis-trans,unbound}} - \Delta G_{\text{cis-trans,bound}} \\ &= (-11.8 \text{ kcal mol}) - (-4.0 \text{ kcal mol}) \\ &= -7.8 \text{ kcal mol.} \end{aligned} \quad (2)$$

Thus, the binding affinity of the *cis* PTX is higher than the *trans* isomer by ~ 7.8 kcal/mol. The isomerization from the *cis* to *trans* isomer significantly destabilizes the PTX's binding with the protein and accelerates the ligand unbinding. To quantify the free energy barrier for the unbinding of the *trans* PTX, we estimated the unbinding

free energy profile as a function of the distance between the COMs of the protein and the ligand. The unbinding of *trans* PTX needs to overcome a free energy barrier of ~ 11.4 kcal/mol, much lower than the 20.1 kcal/mol barrier for the thermal backward isomerization on the ground state [Fig. 7(b)]. In other words, the unbinding of the *trans* PTX precedes its thermal backward isomerization in the protein. On the other hand, according to the thermodynamic cycle [Fig. 7(a)], the unbinding free energy of the *cis* PTX is 11.4 kcal/mol + 7.8 kcal/mol = 19.2 kcal/mol, higher than the ~ 16.1 kcal/mol barrier for the thermal *cis* to *trans* isomerization on the ground state [Fig. 7(b)]. It is worth noting that the calculated 16.1 kcal/mol thermal barrier for the *cis* \rightarrow *trans* isomerization is close to the ~ 16.4 kcal/mol activation energy estimated from the experimental half-life of thermal relaxation of the *cis* PTX⁹ (see the [supplementary material](#)), which justifies the force field used in our simulation. Thus, the *cis* to *trans* thermal isomerization will occur faster than the unbinding of the *cis* PTX. Consequently, the *cis* to *trans* isomerization, driven by either light or thermal fluctuation, precedes the unbinding of the *cis* PTX, and the PTX most likely unbinds from the protein as the *trans* isomer.

It is worth noting that the free energy calculations here differ from the estimation from the docking simulations, which predicted binding energies of ~ -10 kcal/mol and ~ -9 kcal/mol for the *cis* and *trans* isomers, respectively.^{9,12} Based on the docking simulations, there is only a 1–1.5 kcal/mol difference in the binding affinity between the two isomers, whereas in this work the difference of 7.8 kcal/mol is much larger. The free energy simulation in this work samples the thermal fluctuations of the system solvated in explicit solvent molecules, which were not captured by the docking simulations.^{9,12} Also, the calculation of relative binding affinity using the thermodynamic cycle [Figs. 7(a) and 7(b)] is typically more accurate than subtracting the absolute binding energies,⁴¹ the latter of which is often subject to large statistical uncertainty unless extensive sampling of the phase space is performed (as done in this work with

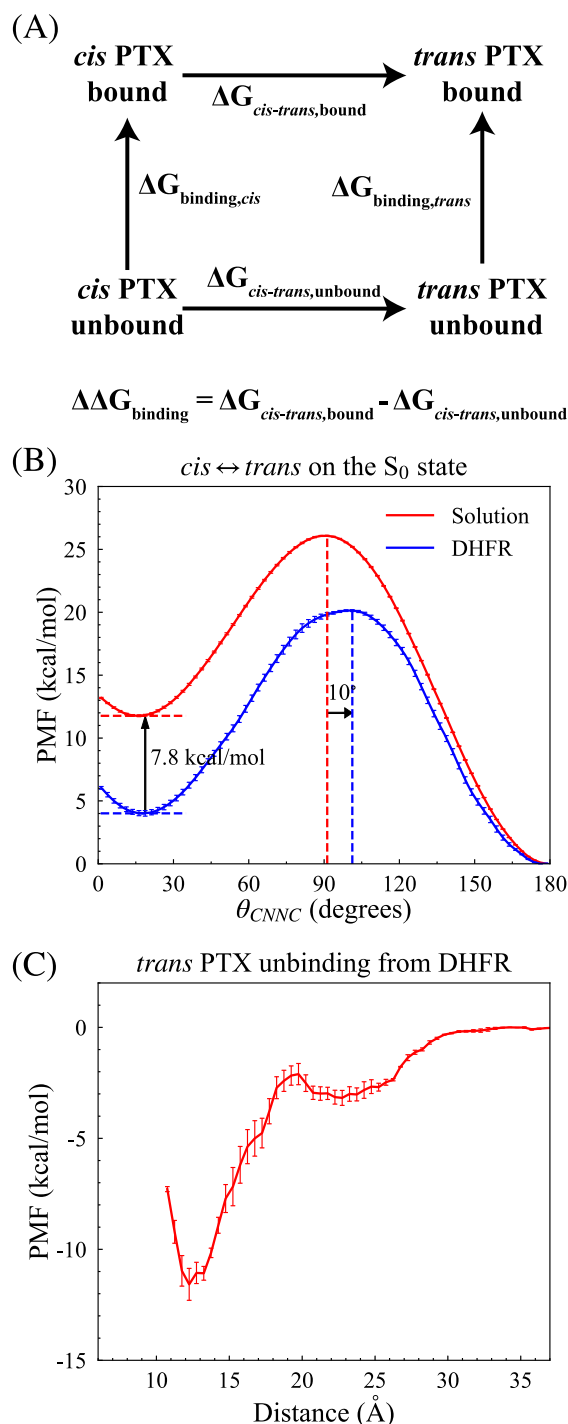


FIG. 7. (a) The thermodynamic cycle used to calculate the relative binding affinity between the *cis* and *trans* isomers. (b) PMF of the ground-state *trans* to *cis* isomerization of the PTX in the solution (red) and DHFR (blue). The peak position of the PMF is shifted to a larger value of the θ_{CNNC} by 10° in the DHFR compared to the solution, indicating more likelihood for ground-state backward isomerization. The *cis* isomer's binding affinity is 7.8 kcal/mol higher than the *trans* isomer. (c) The PMF for the unbinding of the *trans* PTX from the DHFR.

$\sim 3 \mu\text{s}$ enhanced sampling simulations). Thus, we argue that the estimation of the relative binding affinities in this work is more accurate than the docking simulations in previous studies.^{9,12}

Photocycle of the PTX-DHFR system

Building upon our previous work¹² and based on the discussion above, we propose the following photocycle of the PTX-DHFR system (Fig. 8): (1) *trans* → *cis* photoisomerization in solution under 375 nm light, (2) the *cis* PTX binding with the DHFR, (3) the *cis* → *trans* isomerization in DHFR under 460 nm light, and (4) the PTX unbinding from the protein in the *trans* isomer form. Based on our multiscale simulation, the proposed photocycle is the most probable one considering the following factors: (1) differences in the photoisomerization kinetics and quantum yields between solution and DHFR, (2) the binding affinities of the PTX in the two isomer forms, and (3) the ground-state thermal isomerization barriers connecting the two isomer forms. To the best of our knowledge, it is the first time that multiscale simulation techniques are integrated to reveal the full photocycle of a light-regulated photoactive biomolecular system in photopharmacology.

We note that our proposed photocycle is currently not justified quantitatively because several key kinetic parameters necessary for a complete quantitative analysis of the photocycle are missing. For example, we do not have access to the rate constants of the ligand binding and unbinding for both PTX isomers. Calculating the rates of these diffusive barrier crossing events would require not only the free energy barriers but also the position dependent diffusion coefficients along the binding/unbinding pathways,⁴² which necessitate extensive free energy calculations for the unbinding of both PTX isomers. Given the challenges to obtain these parameters accurately, we refrain from characterizing the photocycle by solving master equations, which would be a rigorous way to quantify the reaction flux through the photocycle.

Although a fully quantitative analysis of the photocycle is computationally infeasible now, the photocycle can be partly justified semi-quantitatively based on our current results. Our calculations in this work indicate that *cis* → *trans* photoisomerization completes

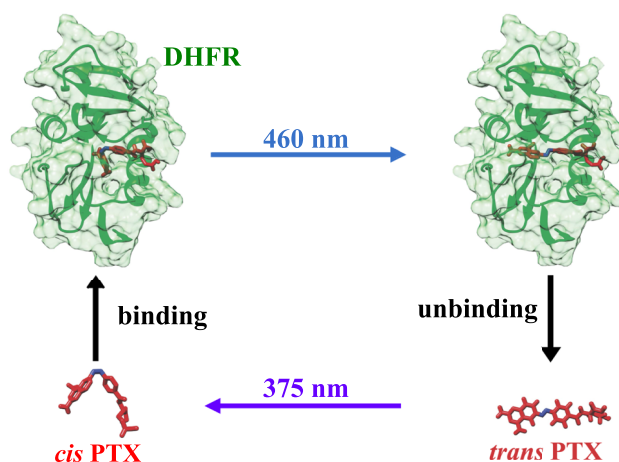


FIG. 8. The photocycle of the PTX-DHFR system.

within a picosecond. In contrast, the ligand unbinding of both the *cis* and *trans* isomers from DHFR needs to overcome large free energy barriers: at least 11.4 kcal/mol for *trans* PTX and 19.2 kcal/mol for *cis* PTX. Although the diffusion coefficients are unknown, these large free energy barriers typically correspond to activated processes occurring beyond microsecond timescale in physiological conditions, as estimated based on biological proton transport processes with similar barriers.^{43–46} Thus, it is clear that the ligand-unbinding is the rate-limiting step for the combined process of unbinding and *cis* → *trans* photoisomerization. In this regard, it is expected that the reaction flux will be dominant for the pathway where the *cis* → *trans* photoisomerization precedes the unbinding process due to the much lower unbinding free energy barrier of the *trans* isomer. Future work is needed to characterize the PTX ligand binding/unbinding kinetics for both isomers in order to fully justify our conclusion quantitatively.

CONCLUSIONS

In this work, we employed multiscale simulation to characterize the *cis* → *trans* photoisomerization of PTX in the DHFR. Building upon our previous study,¹² we propose the most likely photocycle of the PTX-DHFR system. Our results show that (1) the protein environment facilitates the excited-state population decay, (2) the protein environment reduces the isomerization QY by reshaping the ground-state PES near the transition state, (3) the binding affinity of the *cis* PTX is significantly higher than the *trans* PTX, and (4) the unbinding of the PTX most likely follows the *cis* → *trans* photoisomerization in the protein.

Our results reveal new and potentially important design principles of light-regulated therapeutics in photopharmacology. The binding affinity of the photoswitchable ligand may affect the QY of the photoisomerization when it is embedded in the biomolecule. If the binding of one isomer (e.g., *cis* isomer) is too strong with the target biomolecule, the photoisomerization to the other isomer is likely to be hindered, which diminishes the reversibility of light-regulation and is counter-effective for the design goal. One solution to this issue would be to increase the ground-state isomerization barrier such that the slope of the PES is steeper near the perpendicular transition state. This way, after the reshaping of the PES by the protein, it will still be energetically downhill to continue the ground-state isomerization following the non-radiative decay, which reduces the backflow to the reactant minimum. However, the increase of the ground-state isomerization barrier also has its drawback, because once the photoswitch is converted to its active isomer by light (in the case of PTX, 375 nm light populates the *cis* isomer), the resting state of the photoswitch (e.g., *trans* PTX) cannot be easily recovered through thermal isomerization. This drawback will limit the application of the photoswitch as drugs with reduced side effects, because if the active isomer diffuses out of the illuminated area, it cannot easily reset itself to the inactive form under room temperature unless it absorbs light at another wavelength (in this case, 460 nm light populates the *trans* PTX). Thus, the rational design of the photoswitch remains a challenging field, and the design principle needs to further consider the impact of protein–ligand interactions on the photochemistry and effectiveness of the photoswitch.

SUPPLEMENTARY MATERIAL

The [supplementary material](#) contains Figs. S1–S5, which benchmark the accuracy of the force field and the hh-TDA-B3LYP method for describing the ground-state PES of PTX and its model compounds (Figs. S1 and S2), the active space orbitals in the CASSCF calculations (Figs. S3 and S4), the experimental and calculated absorption spectra of the *cis* PTX (Fig. S5). It also contains the estimation of the activation energy of *cis* to *trans* thermal isomerization from the experiment in Ref. 9.

ACKNOWLEDGMENTS

Ruibin Liang and Amirhossein Bakhtiiari acknowledge grant support from the Welch Foundation (Grant No. D-2108-20220331) and Texas Tech University. The researchers used GPU computing facilities provided by the High-Performance Computing Center at Texas Tech University. The authors also acknowledge helpful insights provided by Professor Dirk Trauner from the Department of Chemistry at New York University, New York, USA, and Dr. Umberto Raucci at the Italian Institute of Technology, Genova, Italy.

AUTHOR DECLARATIONS

Conflict of Interest

The authors declare that they have no conflicts of interest.

Author Contributions

Ruibin Liang: Conceptualization (lead); Data curation (lead); formal analysis (lead); funding acquisition (lead); investigation (lead); methodology (lead); project administration (lead); resources (lead); software (lead); supervision (lead); validation (lead); visualization (lead); writing-original-draft (lead); writing-review-editing (lead). **Amirhossein Bakhtiiari:** Writing-original-draft (supporting); writing – review & editing (supporting).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

- ¹K. Hüll, J. Morstein, and D. Trauner, *Chem. Rev.* **118**, 10710 (2018).
- ²J. Broichhagen, J. A. Frank, and D. Trauner, *Acc. Chem. Res.* **48**, 1947 (2015).
- ³I. Tochitsky *et al.*, *Chem. Rev.* **118**, 10748 (2018).
- ⁴K. E. Brechun, K. M. Arndt, and G. A. Woolley, *Curr. Opin. Struct. Biol.* **45**, 53 (2017).
- ⁵W. Szymański *et al.*, *Chem. Rev.* **113**, 6114 (2013).
- ⁶V. Klippenstein, L. Mony, and P. Paoletti, *Trends Biochem. Sci.* **43**, 436 (2018).
- ⁷R. J. Mart and R. K. Allemann, *Chem. Commun.* **52**, 12262 (2016).
- ⁸A. Gautier *et al.*, *Nat. Chem. Biol.* **10**, 533 (2014).
- ⁹C. Matera *et al.*, *J. Am. Chem. Soc.* **140**, 15764 (2018).
- ¹⁰M. Borowiak *et al.*, *Cell* **162**, 403 (2015).
- ¹¹M. G. Vander Heiden, *Nat. Rev. Drug Discovery* **10**, 671 (2011).
- ¹²R. Liang and A. Bakhtiiari, *J. Phys. Chem. B* **126**, 2382 (2022).

- ¹³M. Ben-Nun, J. Quenneville, and T. J. Martínez, *J. Phys. Chem. A* **104**, 5161 (2000).
- ¹⁴B. F. E. Curchod and T. J. Martínez, *Chem. Rev.* **118**, 3305 (2018).
- ¹⁵M. Ben-Nun and T. J. Martínez, *Advances in Chemical Physics* (John Wiley & Sons, Inc., 2002), p. 439.
- ¹⁶C. Bannwarth *et al.*, *J. Chem. Phys.* **153**, 024110 (2020).
- ¹⁷J. K. Yu *et al.*, *J. Am. Chem. Soc.* **142**, 20680 (2020).
- ¹⁸J. K. Yu *et al.*, *J. Chem. Theory Comput.* **16**, 5499 (2020).
- ¹⁹R. Liang, *J. Chem. Theory Comput.* **17**, 3019 (2021).
- ²⁰R. Liang, D. Das, and A. Bakhtiari, *Phys. Chem. Chem. Phys.* **23**, 26263 (2021).
- ²¹O. Trott and A. J. Olson, *J. Comput. Chem.* **31**, 455 (2010).
- ²²V. Cody, J. R. Luft, and W. Pangborn, *Acta Crystallogr., Sect. D* **61**, 147 (2005).
- ²³J. A. Maier *et al.*, *J. Chem. Theory Comput.* **11**, 3696 (2015).
- ²⁴C. J. Dickson *et al.*, *J. Chem. Theory Comput.* **10**, 865 (2014).
- ²⁵Y. J. Wu, H. L. Tepper, and G. A. Voth, *J. Chem. Phys.* **124**, 024503 (2006).
- ²⁶J. Wang *et al.*, *J. Comput. Chem.* **25**, 1157 (2004).
- ²⁷J. Wang *et al.*, *J. Mol. Graphics Modell.* **25**, 247 (2006).
- ²⁸C. Lee, W. Yang, and R. G. Parr, *Phys. Rev. B* **37**, 785 (1988).
- ²⁹J. Finley *et al.*, *Chem. Phys. Lett.* **288**, 299 (1998).
- ³⁰F. Aquilante *et al.*, *J. Comput. Chem.* **37**, 506 (2016).
- ³¹I. S. Ufimtsev and T. J. Martínez, *J. Chem. Theory Comput.* **5**, 2619 (2009).
- ³²A. V. Titov *et al.*, *J. Chem. Theory Comput.* **9**, 213 (2013).
- ³³S. Seritan *et al.*, *J. Chem. Phys.* **152**, 224110 (2020).
- ³⁴S. Seritan *et al.*, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **11**, e1494 (2021).
- ³⁵S. Kumar *et al.*, *J. Comput. Chem.* **13**, 1011 (1992).
- ³⁶S. Kumar *et al.*, *J. Comput. Chem.* **16**, 1339 (1995).
- ³⁷P. Eastman *et al.*, *J. Chem. Theory Comput.* **9**, 461 (2013).
- ³⁸D. A. Case *et al.*, *Computer Program Amber 2021* (University of California, San Francisco, 2021).
- ³⁹G. A. Tribello *et al.*, *Comput. Phys. Commun.* **185**, 604 (2014).
- ⁴⁰M. Bonomi *et al.*, *Nat. Methods* **16**, 670 (2019).
- ⁴¹Z. Cournia, B. Allen, and W. Sherman, *J. Chem. Inf. Model.* **57**, 2911 (2017).
- ⁴²G. Hummer, *New J. Phys.* **7**, 34 (2005).
- ⁴³R. Liang *et al.*, *Proc. Natl. Acad. Sci. U. S. A.* **113**, 7420 (2016).
- ⁴⁴R. Liang *et al.*, *Proc. Natl. Acad. Sci. U. S. A.* **111**, 9396 (2014).
- ⁴⁵R. Liang *et al.*, *Proc. Natl. Acad. Sci. U. S. A.* **114**, 5924 (2017).
- ⁴⁶R. Liang *et al.*, *Proc. Natl. Acad. Sci. U. S. A.* **113**, E6955 (2016).