JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING



pubs.acs.org/jcim
Article

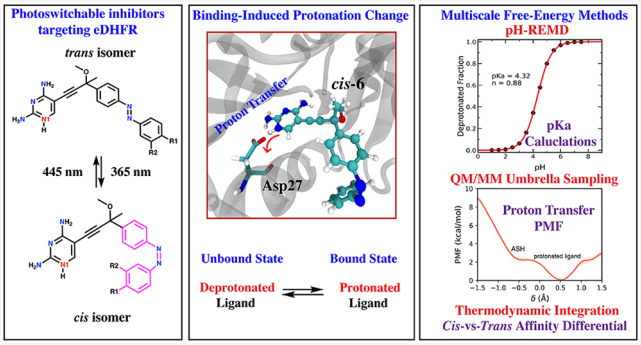
# Multiscale Free-Energy Methods for Protonation-Coupled Light-Responsive Binding of Ionizable Photoswitchable eDHFR Inhibitors

Mohammad Khavani,[†] Kambham Devendra Reddy,[†] Pauf Neupane, Gustavo J. Costa, Laleh Khalvati, and Ruibin Liang*

Cite This: https://doi.org/10.1021/acs.jcim.6c02133
Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

**ABSTRACT:** Photoswitchable ligands enable photocontrol of biomolecular activity by binding to targets in an isomer-dependent, light-responsive manner. Recent developments in ionizable photoswitchable ligands greatly expand their applications but introduce a major design challenge: light-responsive binding can depend on isomeric form, chemical substitution, and binding-induced shifts in protonation equilibria. These effects are tightly coupled, subtle in magnitude, and difficult to predict. Consequently, few computational methods have been developed and systematically benchmarked for quantitatively predicting them. Here, we establish a multiscale free energy method and benchmark it against experimental data for a series of recently developed photoswitchable inhibitors of *Escherichia coli* dihydrofolate reductase (eDHFR), a crucial target in photopharmacology. Constant pH replica-exchange molecular dynamics and quantum mechanics/molecular mechanics umbrella sampling simulations quantitatively characterize the ligand's protonation-state change upon binding to the eDHFR's active site. Thermodynamic integration (TI) simulations using alternative alchemical pathways, thermodynamic cycles, and protonation-state assignments were evaluated for predicting light-responsive affinity differentials and substituent effects. Direct *cis*-to-*trans* transformations with explicit treatment of environment-dependent protonation states best reproduce experimental trends. Compound-to-compound pathways are less reliable because force-field inaccuracies introduce large p$K_a$ errors that are difficult to correct when protonation/deprotonation processes implicitly enter the thermodynamic cycle. TI simulations that ignore binding-induced protonation-state changes fail to consistently reproduce experimental trends. Protein−ligand and ligand-water interaction analyses further reveal the energetic and structural origins of isomer-dependent binding. This study presents a multiscale free-energy framework applied to ionizable photoswitches targeting eDHFR, and highlights the importance of explicitly accounting for binding-induced changes in protonation state when predicting light-responsive binding behavior.




## INTRODUCTION

Molecular photoswitches provide a powerful toolset for the reversible control of biomolecular systems with light.[1] By switching between distinct isomeric forms, they enable noninvasive, remote, and selective photocontrol of target proteins with high spatial and temporal precision, thereby reducing off-target effects that remain a major challenge for conventional therapeutics.[2] These advantages have driven the rapid expansion of photopharmacology, in which molecular photoswitches are designed as light-regulated ligands that target specific biomolecules to minimize side effects.[1−4] The key to their successful design is achieving pronounced bioactivity contrast between the light-activated and dark-adapted states. To this end, first, the isomeric form with the highest bioactivity/affinity toward the target (referred to as "active isomer" below) needs to be correctly identified. For example, a "*cis*-on" effect indicates that the *cis* isomer of the ligand has higher bioactivity/affinity than the *trans* isomer, whereas a "*trans*-on" effect indicates the opposite. Second, the difference in bioactivity and/or binding affinity between distinct isomeric forms of the ligand needs to be maximized.

Therefore, the rational design of photoswitchable ligands can be greatly facilitated by accurate computational predictions of three quantities: the active isomer for a given biomolecular target, the light-responsive binding affinity differential between the active and inactive isomers, and the effect of chemical substitutions on this differential. Experimentally optimizing these properties often requires extensive trial and error, whereas all-atom simulations can, in principle, reveal the

Received: June 26, 2026
Revised: August 4, 2026
Accepted: August 6, 2026

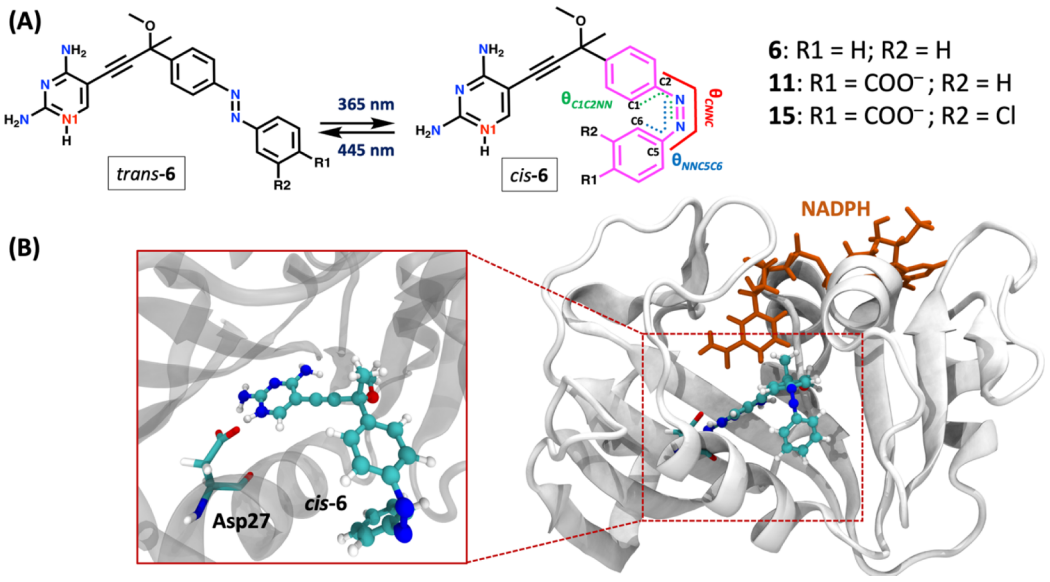



**Figure 1.** (A) Schematic illustration of the photoisomerization of TMP-derivatives (compounds **6**, **11** and **15**) under illumination at different wavelengths, along with their chemical structures. The N1 atom in the diaminopyrimidine ring (in red) can be either deprotonated or protonated. (B) The eDHFR (gray ribbons) in complex with the *cis*-**6** (sticks and balls), highlighting the Asp27 residue (sticks) in the binding pocket and binding-induced protonation of the N1 atom in the diaminopyrimidine ring. The NADPH (sticks) is shown in orange.

protein−ligand interactions that underlie isomer-specific binding. However, quantitative prediction of these free energy properties remains challenging. The structural differences between isomers are often localized near the isomerizing bond, and the resulting binding free-energy differences are typically small, often below a few kilocalories per mole.[5] Errors of similar magnitude can therefore lead to incorrect prediction of the active isomer and misguide ligand design.[6] Standard molecular docking, unbiased molecular dynamics, and end-point free-energy methods often lack the accuracy required to resolve these subtle light-responsive affinity differentials and substituent effects on them, necessitating the use of rigorous and accurate free-energy methods.[7,8]

An additional design challenge arises for ionizable photoswitchable ligands containing titratable groups, which have recently become increasingly important in photopharmacology because their photochemical and biochemical properties are tunable by pH in addition to illumination. Examples are photoswitchable acylhydrazones that can tautomerize between multiple protonation states.[9] Their dominant protonation state may differ between the bulk solution and the protein's binding pocket, where local electrostatic environments can strongly shift protonation equilibria.[10] In such cases, light-responsive binding depends not only on the geometric difference between isomers but also on how protein−ligand interactions shift protonation equilibria. Thus, accurate prediction of light-responsive binding for ionizable photoswitches requires free-energy methods that consistently treat the coupled effects of isomerization, chemical substitution, and protonation-state changes on protein−ligand interactions.

Here, we address these challenges using recently developed photoswitchable inhibitors of *Escherichia coli* dihydrofolate reductase (eDHFR) as a model system. DHFR catalyzes the NADPH-dependent conversion of dihydrofolate to tetrahydrofolate, a key cofactor required for nucleotide and amino acid biosynthesis,[11,12] and has long served as a platform for structure-based inhibitor design. Well-established DHFR inhibitors include methotrexate, which is widely used in cancer chemotherapy to target human DHFR,[11] as well as trimethoprim (TMP), a clinically important antibacterial agent commonly prescribed for urinary tract infections.[13] Recent work has demonstrated that DHFR activity can be effectively modulated using light[4,6,14] via TMP-derived photoswitchable inhibitors, in which the drug scaffold is conjugated to an azobenzene moiety. In a previous experimental design campaign,[6] a series of such compounds, including compounds **6**, **11**, and **15**, was synthesized and characterized (Figure 1A). Compound **11** introduces a *para*-carboxyl group to compound **6**, whereas compound **15** introduces both *meta*-chloro and *para*-carboxyl groups (Figure 1B). These modifications substantially alter both absolute inhibitory potency and the light-responsive bioactivity differential. In particular, the introduction of the carboxyl group (**6** → **11**) increases absolute affinity but unexpectedly reduces the "*cis*-on" effect, whereas further adding the chlorine substituent (**11** → **15**) partially restores light-responsive activity.

Despite these qualitative insights from the experiment,[6] it remains challenging to computationally quantify substituent effects on the light-responsive affinity differential of TMP-based photoswitchable inhibitors. A major challenge is that their ligand-binding affinity may be influenced by changes in protonation state upon binding to eDHFR.[15] In particular, the p$K_a$ of the diaminopyrimidine ring (Figure 1A), which is shared across all TMP-based photoswitchable inhibitors, can be significantly elevated through interactions with ionizable residues such as Asp27 (Figure 1B).[6,15,16] As a result, ligand binding can induce the protonation of the N1 atom on the diaminopyrimidine ring (Figure 1A). The shift in this protonation equilibrium could qualitatively change the active isomer form of the photoswitchable ligand when the *cis*-vs-*trans* affinity differential is small. When rationalizing the molecular basis of *cis*-vs-*trans* affinity differential via conventional MD simulations, the previous study[6] did not consider the potential change in protonation state upon shifting from aqueous solution to the active site. Additionally, there were no rigorous free energy calculations for the *cis*-vs-*trans* affinity


https://doi.org/10.1021/acs.jcim.6c02133
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

differential or the substituent effects on this quantity. As a result, these MD simulations[6] have not consistently predicted the trends of light-responsive activity differential among the photoswitchable ligands. We reason that these critical missing components, if implemented in a quantitative, multiscale free-energy framework, would have benefited quantitative prediction in the previous experimental design.[6]

Therefore, here, we establish a multiscale free-energy method to quantify protonation-coupled, light-responsive binding of these photoswitchable eDHFR inhibitors. Constant-pH replica-exchange molecular dynamics (pH-REMD) and quantum mechanics/molecular mechanics (QM/MM) umbrella sampling simulations were employed to characterize ligand protonation in solution and in the eDHFR's active site. These calculations define the protonation-state assignments used in subsequent thermodynamic integration (TI) simulations,[17,18] which were benchmarked against experimental data for the *cis*-vs-*trans* affinity differential and substituent effects. We compare alternative alchemical transformation pathways, including *cis*-to-*trans* transformations for each ligand and compound-to-compound transformations within fixed isomeric forms, and evaluate the sensitivity of these strategies to different treatments of protonation states. The predictive accuracy of end point methods, such as Molecular Mechanics Poisson−Boltzmann/Generalized-Born Surface Area (MM-PB(GB)SA)[19,20] was also benchmarked. Additionally, we analyzed protein−ligand and ligand-water interactions to interpret the energetic and structural origins of isomer-dependent binding.

It is worth mentioning that, unlike our previous benchmark study[7] on nonionizable propranolol-based photoswitches, where the main focus was evaluating the performance of different free energy methods in predicting substituent and receptor-subtype effects, the present work focuses on a different mechanistic problem: how binding-induced changes in the ligands' protonation state couple with the chemical substitutions to affect the light-responsive affinity differential of ionizable photoswitches.

## METHODS

### Overview and Rationale for the Multiscale Computational Framework

It is worth noting that one of the main goals of this study is to elucidate how binding-induced changes in the protonation state and chemical substitutions couple together to influence the light-responsive affinity differential of ionizable photoswitches. This is a nontrivial challenge that involves careful descriptions of protonation equilibria, protein−ligand binding, and isomer-dependent free energy changes. Therefore, a multiscale computational framework is required.

The pH-REMD simulations were used to determine the dominant protonation states of the ligands in aqueous solution and in the protein-bound state. These calculations allowed us to quantify p$K_a$ shifts and identify whether ligand binding alters the protonation equilibrium. Since the central point of this work involves the role of the protonation-state changes upon binding, establishing the preferred protonation state in each environment is a necessary first step.

Although pH-REMD simulations provide valuable information on p$K_a$ shifts, these calculations are based on molecular mechanics (MM) force fields and may be inaccurate due to parametrizations of the ligand's fixed point charges. Furthermore, MM force fields cannot explicitly describe the charge delocalization associated with strong hydrogen-bond interactions between the ligand and Asp27, or proton-transfer reactions between them. Therefore, it is necessary to validate the conclusions drawn from force-field-based pH-REMD simulations at the QM level of theory. For this reason, first-principles QM/MM umbrella sampling simulations were also performed to examine proton transfer between the Asp27 residue and the ligand in the eDHFR binding site. These first-principles QM/MM simulations provide a more accurate description of the hydrogen-bonding interactions and free-energy profile associated with proton transfer than the MM force field, allowing us to independently verify the relative stability of the protonated and deprotonated states predicted by the pH-REMD simulations.

Building upon these results, the TI simulations with different thermodynamic cycles and protonation states were then designed to investigate the coupled effects of changes in protonation state and chemical substitutions on light-responsive affinity. The TI method was used as the primary approach since it provided the highest accuracy in our previous study.[21] MM-GBSA and MM-PBSA calculations were included to assess whether these computationally less demanding methods could reproduce the trends observed in the experiments and thermodynamic integration calculations. Since these approaches require substantially less computational resources than TI, they could be useful tools for screening and lead optimization. Benchmarking their performance against TI and experimental data allowed us to evaluate their suitability for high-throughput computational design of ionizable photoswitches.

Finally, conventional molecular dynamics simulations and trajectory analyses were used to identify the molecular interactions that underlie the trend in light-responsive affinity differences observed in the free-energy calculations and experiments. Protein−ligand interactions, hydrogen-bonding patterns, and ligand-water interactions were analyzed to connect the energetic results with the underlying structural features.

### System Setup

The crystal structure of eDHFR (PDB ID: 3DAU)[22] was selected as the starting model for all molecular docking and molecular dynamics (MD) simulations. For each TMP-based ligand in either the *cis* or *trans* isomer, AutoDock Vina[23] was used to place it in the substrate-binding pocket of eDHFR. To assess the reliability of the docking poses, the resulting binding poses were inspected to ensure that key interactions with binding-site residues were, whenever possible, consistent with the methotrexate ligand in the crystal structure (PDB ID: 3DAU), as well as the docked structures reported in the original study of these TMP-derived photoswitchable inhibitors.[6] The selected poses reproduced the overall ligand orientation within the binding pocket and preserved the key interactions with important active-site residues, particularly those involving the diaminopyrimidine ring and Asp27 (Figure 7).

The NADPH molecule in the crystal structure was retained, while all crystal water molecules were removed because none were buried deep within the substrate-binding pocket. The initial protonation states of ionizable amino acid residues at physiological pH were assigned using the H++ server.[24] Both protonation states of the ligand were set up and tested. The




https://doi.org/10.1021/acs.jcim.6c02133
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

protein−ligand complexes were then solvated in a box of explicit water molecules under periodic boundary conditions (PBC) of ∼ 85 × 84 × 90 Å$^3$. Water molecules were modeled using the TIP3P model,[25] while the protein was described using the AMBER ff14SB force field.[26] The parameters in ref 27 were used for NADPH. Ligand parameters were first generated using the general AMBER Force Field (GAFF) approach.[28] For each isomer of each ligand, a separate aqueous solution set up where the ligand was solvated in a PBC box of water molecules of ∼ 45 × 45 × 45 Å$^3$.

The torsional parameters in the force field for three dihedral angles, C2−N3 = N4−C5 ($\theta_{CNNC}$), C1−C2−N3 = N4 ($\theta_{C1C2NN}$), and N3 = N4−C5−C6 ($\theta_{NNC5C6}$) (Figure 1A), were reparametrized using relaxed scans of their corresponding potential energy surface (PES) with a multireference electronic structure method. This step was necessary because our previous studies[7,8,29] indicated that unmodified GAFF parameters have artificially low torsional barriers, leading to unphysically fast *trans*↔*cis* thermal isomerization on nanosecond time scales during MD simulations. Due to the multireference character of the electronic wave function near the transition state of the $\theta_{CNNC}$ torsion,[30] all three torsional terms were refined based on relaxed PES scans of the N1-deprotonated ligands (Figure S1) using the floating occupation molecular orbital hole−hole Tamm-Dancoff approximated density functional theory (FOMO-hh-TDA-DFT) electronic structure method,[31,32] with the BH&HLYP functional and the 6−31 + G(d) basis set. This multireference electronic structure method has the advantage of incorporating both static and dynamic electron correlation and has been extensively validated for both the ground and excited-state PESs and dynamics of azobenzene-based photoswitches.[8,29,30,32,33] The torsional parameters were adjusted until the relaxed-scanned PES using the force field closely matched the relaxed-scanned PES using the FOMO-hh-TDA-DFT method in terms of *cis*-to-*trans* energy differences and isomerization barriers (Figure S1). All QM calculations used for force field refinement were performed with the TeraChem software package.[34,35]

## Force-Field-Based MD Equilibration Simulations

Each system was initially relaxed through 100,000 steps of energy minimization, followed by gradual equilibration in the constant *NVT* ensemble at 300 K for 2 ns using a 1 fs time step. During this equilibration phase, positional restraints were applied to all non-hydrogen atoms of the protein using harmonic potentials with a force constant of 50.0 kcal/mol/Å$^2$. Subsequently, restrained MD simulations were carried out in the constant *NPT* ensemble at 300 K and 1 atm, with a 1 fs time step, gradually reducing the positional restraint to 0.1 kcal/mol/Å$^2$ over 5 ns. Then, removing all positional restraints, constant *NPT* simulations in the same ensemble were carried out for a total of 200 ns. The initial 50 ns of each trajectory were discarded as equilibration stage, and the remaining 150 ns were used for interaction energy analysis. The Langevin thermostat[36] with a collision frequency of 1 ps$^{-1}$ was applied to maintain the temperature at 300 K, while the pressure was regulated by a Monte Carlo barostat.[37] Electrostatic interactions were evaluated using the Particle Mesh Ewald (PME) method,[38] and a cutoff distance of 12 Å was employed for van der Waals interactions. All bonds involving hydrogen atoms were constrained using the SHAKE algorithm,[39] allowing the use of a 2 fs time step throughout the simulations. All minimizations and MD simulations were performed with the Amber24 software package.[40]

## pH-REMD Simulations

To identify the protonation state of the photoswitchable TMP derivatives in both the aqueous solution and eDHFR, pH-REMD simulations[41] were carried out for compound **6** as a representative example. To determine the deprotonation free energy ($\Delta G_{deprotonation}$) of **6** in aqueous solution, a thermodynamic integration (TI)[18] simulation was first performed to alchemically transform protonated **6** to deprotonated **6**. Details for this TI simulation are described in the TI section below. Then, the reference p$K_a$ value in the aqueous solution was calculated using DFT with the M06-2X functional[42] and TZVP basis set in a conductor-like polarizable continuum model (M06-2X/TZVP/CPCM, $\varepsilon$ = 80.15),[43] following similar computational protocols in previous studies.[44] The reference $\Delta G_{deprotonation}$ and p$K_a$ values were then used as input parameters for the pH-REMD simulations in the aqueous solution and protein.

Prior to performing pH-REMD simulations, a multistep equilibration protocol was applied to both the aqueous solution and protein−ligand complex systems. Each system was first subjected to energy minimization to remove unfavorable contacts. Minimization was carried out using a combination of steepest descent and conjugate gradient methods for a total of maximum 50,000 steps, with the first 10,000 steps performed using steepest descent. During this stage, harmonic positional restraints (10 kcal/mol/Å$^2$) were applied to the ligand while allowing the environment to relax. A cutoff of 12 Å was used for nonbonded interactions, and the systems were treated under PBC with explicit solvent. Following minimization, the systems were gradually heated to 300 K using the constant *NVT* ensemble for 1 ns. Afterward, a 1 ns equilibration simulation was performed in the constant *NPT* ensemble at 300 K and 1 atm. Positional restraints (5 kcal/mol/Å$^2$) were applied to the ligand during this step. A second equilibration stage was then carried out using constant pH MD (CpHMD) at pH = 4.5 without positional restraints at 300 K. In this step, the system was simulated for 1 ns for the ligand in water and 6 ns for the protein-ligand complex. Importantly, in the protein−ligand complex systems, both the Asp27 and the ligand were treated as ionizable in the CpHMD simulations.

Next, preliminary CpHMD calculations were performed as independent replicas to estimate the ligand's p$K_a$ values and identify the pH region where protonation-state transitions occur. These simulations were carried out for 10 ns at pH values separated by one pH unit. Based on the resulting preliminary titration curve, pH ranges encompassing the expected protonation-state transitions were selected for the pH-REMD calculations.

Following the pre-equilibration steps, pH-REMD simulations were then performed over a pH range of 1.0−7.5 for the aqueous solution and 7.0−17.5 for the protein−ligand complexes. *A* spacing of 0.5 pH units between neighboring replicas was employed to ensure a sufficient acceptance ratio for replica exchange between neighboring windows while avoiding unnecessary replicas outside the relevant pH range. This resulted in 14 replicas for the aqueous solution simulations and 22 for the protein-bound systems. The average pairwise acceptance ratio across all pairs of exchanged windows is 90.7%, with a minimum value of 60.2%.




https://doi.org/10.1021/acs.jcim.6c02133
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

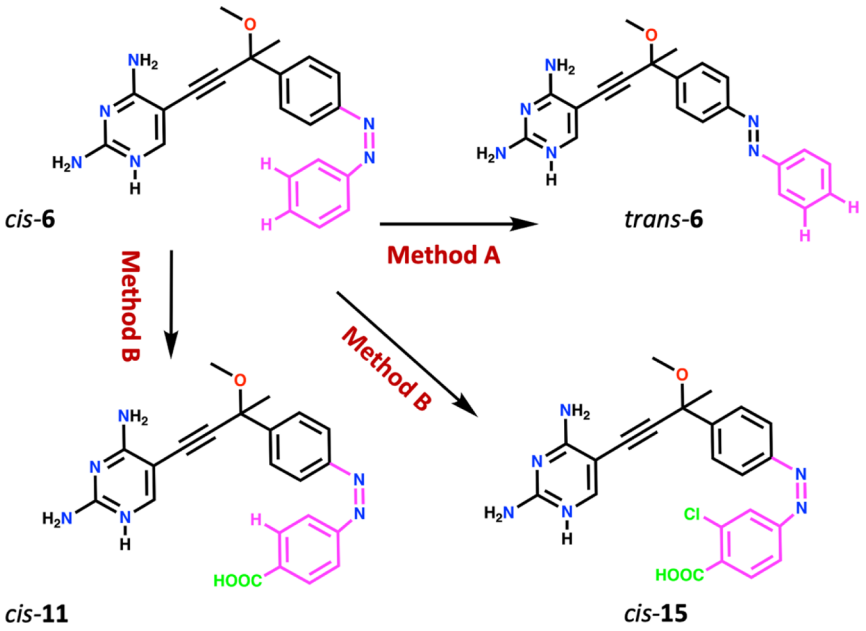



**Figure 2.** Schematic illustration of the two TI methods for quantifying the *cis*-vs-*trans* affinity differential and substituent effects. In Method A, the *cis*-vs-*trans* binding affinity differential is evaluated for a single compound by alchemically transforming from the *cis* to *trans* isomers. This transformation is achieved by relocating a phenyl group across the *N*═N bond. In Method B, the substituent effects on the binding affinity of each isomer are examined by transforming one compound into another while preserving the same isomeric form. In this case, modifications are introduced at specific positions on the benzyl ring, such as *para* or *meta* sites. For example, the *cis*-**6** is converted into the *cis*-**11** or *cis*-**15** by replacing a hydrogen atom with a COOH and/or a Cl group at the *para* and/or *meta* positions. Both methods can evaluate the substituent effects on the *cis*-vs-*trans* binding affinity differential. See Methods for details.

Each simulation was performed under constant volume conditions at 300 K for 10 ns. For each pH replica, protonation state sampling was enabled throughout the simulation, with Monte Carlo attempts to change protonation states performed every 500 steps. A short relaxation period of 500 steps was applied after each protonation-state change. The replica exchange is attempted every 1 ps between adjacent replicas. All CpHMD and pH-REMD simulations were performed using the explicit solvent implementation[41] in the AMBER24 software package.[40] All simulations were performed using the TIP3P water model[25] with a time step of 2 fs, with the SHAKE algorithm enabled.

## QM/MM MD and Umbrella Sampling Simulations

The potential of mean force (PMF) of the proton transfer (PT) between the eDHFR and compound **6** was calculated using QM/MM MD coupled with the umbrella sampling (US) technique.[45] The QM region comprised the ligand and the side chain of the Asp27 residue (Figures 5A and S2), which was treated at the B3LYP-D3/def2-SVP level of theory.[46,47] The MM region includes the rest of the system, which was treated with the same force field as the classical MD equilibration. The electrostatic embedding scheme was used for the electrostatic interaction between the QM and MM regions.

This QM/MM setting is supported by previous QM/MM studies of enzymatic reactions involving proton and hydrogen transfer.[48−50] In those studies, the B3LYP functional[46] with the D3 dispersion correction[47] were widely used to investigate these reactions in both static reaction-profile calculations and QM/MM umbrella sampling simulations, yielding reaction profiles and mechanistic trends consistent with available experimental observations. Notably, the B3LYP-D3/MM approaches have been benchmarked against first-principles correlated wave function methods for hydrogen- and proton-transfer reactions.[49,51] The QM/MM partition scheme was chosen to provide a computationally efficient and sufficiently sized model of the proton-transfer reaction. The QM region includes the *trans*-**6** ligand and the side chain of Asp27, thereby explicitly describing the electron delocalization among the atoms directly involved in the proton-transfer process. Thus, the use of B3LYP-D3 in the present QM/MM framework is consistent with previous applications to ground-state proton/hydrogen transfer reactions in biomolecular systems.[52]

The QM/MM MD simulations were performed at 300 K temperature using a Langevin thermostat and a time step of 0.5 fs. In the US simulation, a collective variable (CV) $\delta$ (eq 1) was chosen as the proton transfer coordinate (Figure 5) and biased in each umbrella window:

$$\delta = d(\mathrm{O}_{\mathrm{Asp27}} - \mathrm{H}_{\mathrm{compound\ 6}}) - d(\mathrm{N1}_{\mathrm{compound\ 6}} - \mathrm{H}_{\mathrm{compound\ 6}}) \quad (1)$$

This CV captures the concerted breaking and formation of the N−H and O−H bonds of the donor and acceptor during the PT event. A negative CV value represents the proton being closer to the carboxylic oxygen atom of the Asp27 residue than to the N1 atom on the diaminopyrimidine ring of compound **6**, and *vice versa*. In total, 41 umbrella windows whose harmonic centers were evenly distributed in a range of −1.5 to 1.5 Å of the CV were simulated. For each window, a harmonic potential with a force constant of 50 kcal/mol/Å$^2$ was applied and the QM/MM MD trajectory was propagated for approximately 6 ps, removing the first 1 ps as equilibration, resulting in a total QM/MM US simulation time of approximately 205 ps. The unbiased PMF along the above-defined CV was reconstructed using the weighted histogram analysis method (WHAM)[53] from the biased CV distributions


https://doi.org/10.1021/acs.jcim.6c02133
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

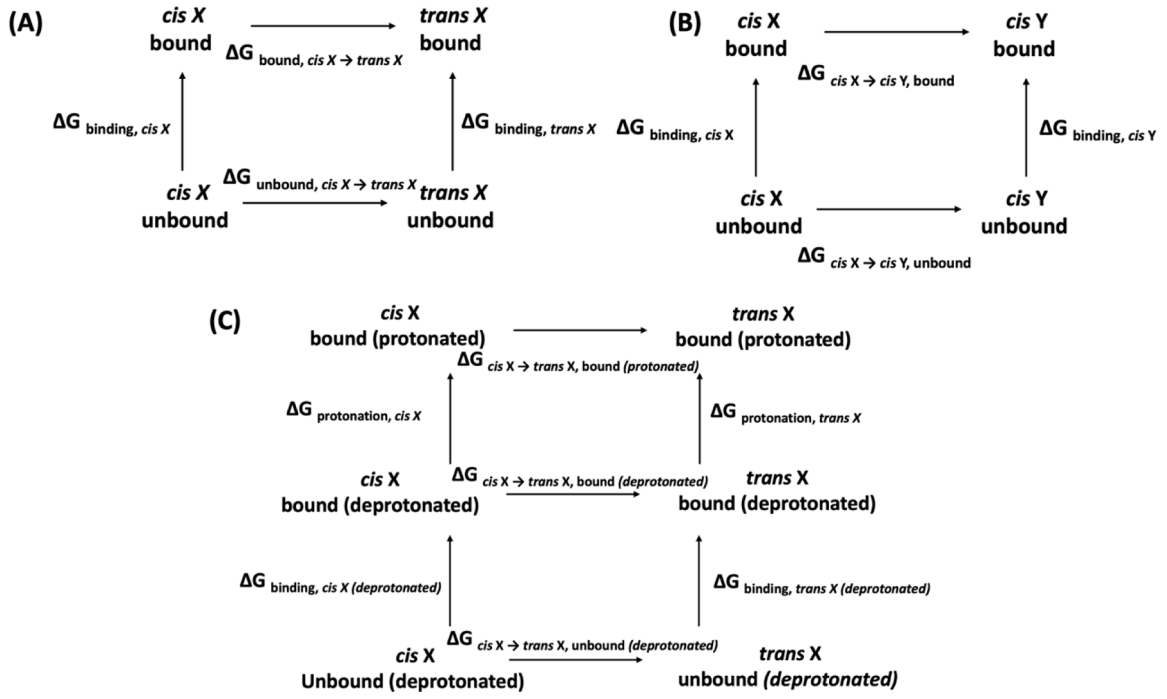



**Figure 3.** Thermodynamic cycles used for quantifying *cis*-vs-*trans* affinity differential and substituent effects. (A) In Method A, the *cis*-X is converted into *trans*-X through an alchemical transformation performed separately in the protein-bound state ($\Delta G_{\text{bound},cisX\rightarrow transX}$) and in aqueous solution ($\Delta G_{\text{unbound},cisX\rightarrow transX}$). The corresponding free energy changes obtained from TI simulations are combined to yield the *cis*-vs-*trans* affinity differential ($\Delta\Delta G_{X,cis\rightarrow trans}$) according to eq 2. (B) In Method B, X is alchemically transformed into Y while maintaining the same isomeric form and protonation state. The associated free energy changes are computed in both the protein ($\Delta G_{cisX\rightarrow cisY,\text{bound}}$) and aqueous solution ($\Delta G_{cisX\rightarrow cisY,\text{unbound}}$). These values are combined to calculate the relative binding affinities between X and Y for each isomeric form ($\Delta\Delta G_{\text{binding},cisX\rightarrow cisY}$, $\Delta\Delta G_{\text{binding},transX\rightarrow transY}$), according to eqs 3 and 4. For both Methods A & B, the combined results from these calculations allow direct evaluation of substituent effects on the *cis*-vs-*trans* affinity differential using eq 5. (C) For Protocol 1 of Method B, which describes protonation state change of the ligand upon protein binding, the thermodynamic cycle shown in B can be thought of as a combination of two subcycles. The upper cycle describes the protonation state change of the ligand in the same environment, while the lower cycle describes the ligand's binding process in the same protonation state. The overall $\Delta\Delta G_{\text{binding},cisX\rightarrow cisY}$ and $\Delta\Delta G_{\text{binding},transX\rightarrow transY}$ calculated using this protocol adds up both the relative p$K_a$'s of the two ligands (upper-half cycle, two vertical arrows) and the relative binding affinity between the two ligands in the same protonation states (lower-half cycle, two vertical arrows). For Protocol 2 of Method B, only the lower subcycle is involved, and no protonation state change is considered.

in each window, and statistical uncertainties were estimated via block average analysis. All QM/MM MD simulations were carried out at B3LYP-D3/def2-SVP/AMBER using TeraChem[34,35] interfaced with OpenMM,[54] and the external harmonic bias was applied via interface with the PLUMED plugin.[55]

## Thermodynamic Integration (TI) Simulations

TI simulations and thermodynamic cycles were employed to quantify two types of relative binding free energies. The first type is the relative binding free energies between the *cis* and *trans* isomers of each photoswitchable compound, referred to as the *cis*-vs-*trans* affinity differential ($\Delta\Delta G_{X,\ cis\rightarrow trans}$). To calculate this quantity in the TI simulations, *cis* X was converted to *trans* X via an alchemical transformation in which one phenyl ring gradually disappeared from one side of the N═N double bond and reappeared on the opposite side (Figure 2), while retaining the same protonation state. The alchemical transformations of the ligand were carried out in both the protein and aqueous solution, yielding the free energy changes $\Delta G_{\text{bound},cisX\rightarrow transX}$ and $\Delta G_{\text{unbound},cisX\rightarrow transX}$, respectively. The protonation state of the ligands can differ between the two environments. These values were then combined using a thermodynamic cycle (Figure 3A, eq 2) to obtain the *cis*-vs-*trans* binding affinity difference for compound X ($\Delta\Delta G_{X,cis\rightarrow trans}$). TI simulations employing this alchemical pathway and thermodynamic cycle are hereafter referred to as Method A.

The second type of relative binding affinities is between a pair of compounds in the same isomer, yielding $\Delta\Delta G_{\text{binding},cisX\rightarrow cisY}$ and $\Delta\Delta G_{\text{binding},transX\rightarrow transY}$. To calculate this quantity, compound X was alchemically transformed into Y by gradually changing the substituent groups that differ between the two molecules, while retaining the isomeric form and protonation state (Figure 2). The free energy changes for these transformations were computed in both the protein ($\Delta G_{cisX\rightarrow cisY,\text{bound}}$ and $\Delta G_{transX\rightarrow transY,\text{bound}}$) and the aqueous solution ($\Delta G_{cisX\rightarrow cisY,\text{unbound}}$ and $\Delta G_{transX\rightarrow transY,\text{unbound}}$). The protonation state of the ligands can differ between the two environments. These values were combined using a thermodynamic cycle (Figure 3B, eqs 3 & 4) to predict the relative binding free energies, $\Delta\Delta G_{\text{binding},cisX\rightarrow cisY}$ and $\Delta\Delta G_{\text{binding},transX\rightarrow transY}$. TI simulations employing this alchemical pathway and thermodynamic cycle are hereafter referred to as Method B.




https://doi.org/10.1021/acs.jcim.6c02133
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Finally, the impact of substituent modification on the *cis*-vs-*trans* affinity difference, denoted as $\Delta\Delta\Delta G_{X\rightarrow Y,cis\text{-vs-}trans\ \mathrm{binding}}$, was derived from the relative binding free energies obtained from either method using eq 5. This quantity indicates how the chemical substitutions that convert *X* to *Y* modulate the light-responsive affinity differential.

$$\Delta\Delta G_{X,cis\rightarrow trans} = \Delta G_{\mathrm{binding},trans\,X} - \Delta G_{\mathrm{binding},cis\,X} = \Delta G_{\mathrm{bound},cis\,X\rightarrow trans\,X} - \Delta G_{\mathrm{unbound},cis\,X\rightarrow trans\,X} \quad (2)$$

$$\Delta\Delta G_{\mathrm{binding},cis\,X\rightarrow cis\,Y} = \Delta G_{\mathrm{binding},cis\,Y} - \Delta G_{\mathrm{binding},cis\,X} = \Delta G_{cis\,X\rightarrow cis\,Y,\mathrm{bound}} - \Delta G_{cis\,X\rightarrow cis\,Y,\mathrm{unbound}} \quad (3)$$

$$\Delta\Delta G_{\mathrm{binding},trans\,X\rightarrow trans\,Y} = \Delta G_{\mathrm{binding},trans\,Y} - \Delta G_{\mathrm{binding},trans\,X} = \Delta G_{trans\,X\rightarrow trans\,Y,\mathrm{bound}} - \Delta G_{trans\,X\rightarrow trans\,Y,\mathrm{unbound}} \quad (4)$$

$$\begin{aligned}\Delta\Delta\Delta G_{X\rightarrow Y,cis\text{-vs-}trans\ \mathrm{binding}} &= \Delta\Delta G_{Y,cis\rightarrow trans} - \Delta\Delta G_{X,cis\rightarrow trans}\\ &= [\Delta G_{\mathrm{binding},trans\,Y} - \Delta G_{\mathrm{binding},cis\,Y}] - [\Delta G_{\mathrm{binding},trans\,X} - \Delta G_{\mathrm{binding},cis\,X}]\\ &= [\Delta G_{\mathrm{binding},trans\,Y} - \Delta G_{\mathrm{binding},trans\,X}] - [\Delta G_{\mathrm{binding},cis\,Y} - \Delta G_{\mathrm{binding},cis\,X}]\\ &= \Delta\Delta G_{\mathrm{binding},trans\,X\rightarrow trans\,Y} - \Delta\Delta G_{\mathrm{binding},cis\,X\rightarrow cis\,Y}\end{aligned} \quad (5)$$

For each alchemical transformation, the atoms disappearing from the initial molecule or appearing in the final one were first defined in a dual topology setup. Then, the free-energy calculation was carried out using a three-stage protocol designed for numerical stability. First, the charges of the disappearing atoms were gradually reduced to zero. Second, the nonbonded and bonded interactions associated with the disappearing atoms were reduced to zero, while those for the appearing atoms were increased to full strength. Both the disappearing and appearing atoms have zero point-charges in this step. Softcore vdW potentials were employed to smooth steric clashes and increase numeric stability. Third, the charges of the appearing atoms were gradually increased to their full values. Each of these three stages was simulated using 11 equally spaced windows ($\Delta\lambda = 0.1$), resulting in a total of 33 $\lambda$ windows for each transformation. This $\lambda$ window spacing scheme was chosen to ensure a smooth alchemical transition while maintaining computational efficiency. Similar TI studies have employed comparable schemes.[56,57] For example, Chen et al.[57] used a three-stage protocol with 12 $\lambda$ windows per stage to investigate the effects of protein residue mutations on inhibitor binding. Because the transformations considered in the present study involve relatively small ligand modifications, especially in Method B, rather than the mutation of an entire protein side chain, we found that 11 $\lambda$ windows per stage yielded stable and reproducible free-energy estimates. The same protocol was also successfully applied in our recent study of photoswitchable antagonists of $\beta$-adrenergic receptors, where the calculated free-energy differences showed good agreement with experimental measurements.[7]

The starting geometry for each TI simulation was selected from a representative snapshot of the production MD trajectory for the initial compound/isomer X. Then, it was equilibrated using a mixed Hamiltonian of the X and Y (final compound/isomer) connected by the alchemical transformation, at window $\lambda = 0.5$ that scaled both the charge and vdW interactions. Specifically, with this mixed Hamiltonian, the standard-MD-equilibrated structures were first minimized for 15,000 steps with the steepest descent algorithm, followed by gradual heating to 300 K over 1 ns using a Langevin thermostat[36] with a collision frequency of 1 $ps^{-1}$. The systems were then equilibrated for 20 ns in the constant *NVT* ensemble at 300 K and 1 atm. The final equilibrated structure in the $\lambda = 0.5$ mixed Hamiltonian served as the starting point for all $\lambda$ windows across the three stages. For each $\lambda$ value in each stage, the system was equilibrated for 500 ps in the constant *NVT* ensemble, followed by at least 10 ns of production sampling in the constant *NPT* ensemble. During all TI simulations, bonds involving atoms undergoing alchemical changes were excluded from the SHAKE constraints,[39] and a 1 fs time step was used consistently throughout. The error bars in the free energy changes were evaluated using block-averaging analysis.

### MM-PBSA and MM-GBSA Simulations

Binding free energies were also estimated using the MM-PB(GB)SA methods[19] based on the final 50 ns of the production MD trajectories. In these calculations, the dielectric constants for the aqueous solution and the protein were set to 80 and 4, respectively. A solvent probe radius of 1.4 Å was used for defining the solvent-accessible surface, and standard force-field atomic radii were applied.[58] The binding free energies were computed for both isomeric forms of each compound ($\Delta G_{\mathrm{binding},trans\ X}$, $\Delta G_{\mathrm{binding},cis\ X}$). Then, the *cis*-vs-*trans* binding affinity differential, $\Delta\Delta G_{X,cis\rightarrow trans}$, was obtained from these values using eq 2. The $\Delta\Delta\Delta G_{X\rightarrow Y,cis\text{-vs-}trans\ \mathrm{binding}}$ values was subsequently derived according to eq 5.

## RESULTS

This section is organized as follows. First, we investigate the change in the protonation state of the ligands' diaminopyrimidine ring upon protein binding using pH-REMD and QM/MM umbrella-sampling simulations. Second, we benchmark the performance of TI and end point free-energy methods to characterize the *cis*-vs-*trans* affinity differential and the substituent effects against the experimental data. The importance of explicitly accounting for the ligand's binding-induced change in protonation state was highlighted. Also, the accuracy of TI simulations employing two different alchemical transformation pathways was compared and discussed. Third, based on MD simulations, we interpret the protein−ligand interactions that lead to the *cis*-vs-*trans* affinity differential for each ligand under investigation.

### Ligand Protonation upon Binding to the Active Site of eDHFR

The diaminopyrimidine ring is shared in all three compounds investigated: compounds **6**, **11**, and **15**. To calculate the change in protonation state of the N1 atom on the diaminopyrimidine ring of the three ligands upon binding to the protein, pH-REMD simulations were performed to characterize the $pK_a$'s of compound **6** in both aqueous


https://doi.org/10.1021/acs.jcim.6c02133
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

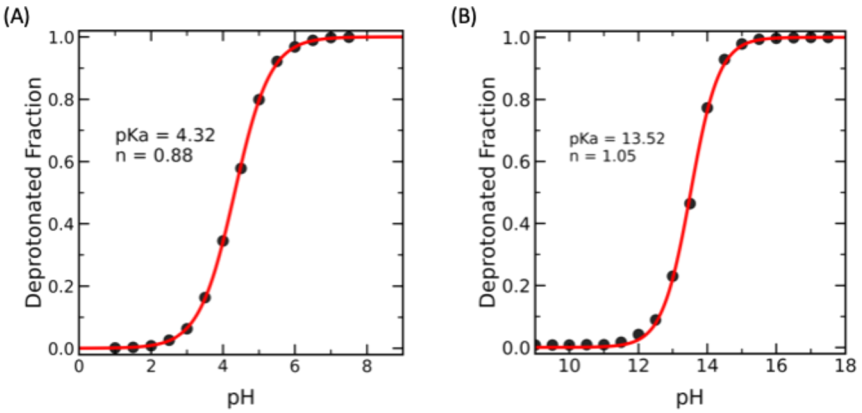



**Figure 4.** PH-titration curves of *cis*-**6** obtained from pH-REMD simulations. (A) Aqueous solution system. (B) eDHFR-ligand complex system. The calculated p$K_a$ values of the ligand substantially increase upon protein binding. In the pH-REMD simulations of the protein−ligand complex, Asp27 was treated as ionizable but remained deprotonated throughout the pH range covered. The Hill coefficients (n's) are also indicated.

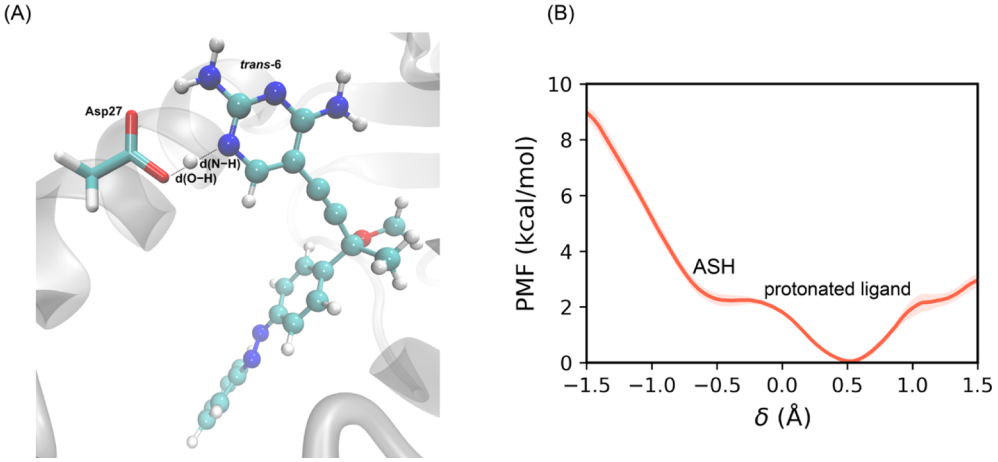



**Figure 5.** Free energy profile of proton transfer between the Asp27 residue and *trans*-**6**. (A) The salt bridge between the Asp27 and the ligand in the binding site, highlighting the transferring proton and the two key distances for this reaction: d(O−H) and d(N−H). (B) Potential of mean force (PMF) obtained from B3LYP-D3/MM US simulations along the reaction coordinate $\delta$ = d(O−H) − d(N−H). The free energy profile has the lowest minimum in the region where the proton resides on the ligand ($\delta$ = ∼ 0.5Å), indicating that the protonated ligand/deprotonated Asp27 is thermodynamically more favored than the deprotonated ligand/protonated Asp27. The shaded belt represents the error bars.

solution and the protein. Compound **6** was selected as a representative case of all three compounds, since the N1 atom is well-separated from the additional substituents in compounds **11** and **15** (Figure 1), which are unlikely to significantly interfere with the p$K_a$ of the diaminopyrimidine ring atom. Based on these simulations, the fraction of the deprotonated state for the ionizable group is plotted as a function of pH, yielding the titration curves (Figure 4). These curves were then fitted to the Hill equation to estimate the p$K_a$ of the ionizable group.

In the pH-REMD simulations, the reference p$K_a$ value in aqueous solution was set to 4.3 for compound **6**. This reference p$K_a$ value was calculated using M06-2X/TZVP/CPCM ($\varepsilon$ = 80.15, see Method for details).[43] It is worth noting that the TMP, which served as the parent compound for the design of compound **6**, has an experimental p$K_a$ of 7.12 in bulk solution.[59] Using the same electronic structure method, the calculated p$K_a$ of TMP is 8.61. Thus, our DFT calculation predicts that the p$K_a$ of compound **6** is ∼ 4.3 units lower than TMP. Although this p$K_a$ shift could be overestimated, it suggests that the chemical modifications introduced in TMP reduce its basicity, likely due to changes in local charge distribution. Considering that the experimental p$K_a$ of 7.12 for the TMP[59] and our DFT calculated p$K_a$ shift, the compound **6** is most likely deprotonated in the aqueous solution in near-neutral pH conditions.

The p$K_a$ of the N1 atom in compound **6**, calculated by pH-REMD simulation, is increased to 13.5 from 4.3 upon binding to the active site of eDHFR (Figure 4B). Importantly, in the pH-REMD simulations of the protein−ligand complex, Asp27 was treated as ionizable but remained predominantly deprotonated (>97% of the population) throughout the covered pH range. This ∼9 pH unit upward shift in the ligand's p$K_a$ indicates strong stabilization of the ligand's protonated state within the protein environment, mostly due to forming a hydrogen bond with the deprotonated Asp27 residue. As a result, the pH-titration curve confirms that compound **6** remains predominantly N1-protonated near physiological pH range.

The large shift in the ligand's p$K_a$ upon protein binding highlights the strong influence of the local electrostatic environment on the ligand's protonation equilibrium. In this case, strong hydrogen-bond interactions with Asp27 favor the protonation of the N1 atom. Importantly, these pH-REMD simulation results are consistent with the proton transfer free energy profile obtained from the QM/MM umbrella sampling


https://doi.org/10.1021/acs.jcim.6c02133
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

simulations (see below), which also indicate that the protonated form of the diaminopyrimidine ring is stabilized in the protein-bound state. Furthermore, the results align with previous experimental studies, in which the diaminopyrimidine rings of the TMP and methotrexate in complex with eDHFR have p$K_a$ values above 10 and are considered protonated.[15,16]

To further validate the results from constant pH simulations, we performed QM/MM US simulations to characterize the potential of mean force (PMF) of proton transfer between the N1 atom of the compound **6** and the carboxylic group on the Asp27 residue in the active site of eDHFR (Figure 5A). The QM region includes the side chain of Asp27 and the whole ligand, which was treated at the B3LYP-D3/def2-SVP level of theory (Method). The PMF is constructed along the proton transfer coordinate ($\delta$) (Figure 5A and eq 1). The PMF (Figure 5B) features a lowest-energy minimum at $\delta = \sim 0.5$ Å, corresponding to the Asp27-deprotonated/N1-protonated state. Meanwhile, a shoulder in the PMF at $\delta = \sim -0.5$ Å corresponding to the Asp27-protonated/N1-deprotonated state has more than 2 kcal/mol higher free energy than the minimum at $\delta = \sim 0.5$ Å. Thus, the PMF demonstrates that proton transfer from Asp27 to the ligand is thermodynamically favorable in the protein environment. In other words, when the Asp27's carboxylic group and the ligand's N1 atom form a hydrogen bond, the proton involved in this hydrogen bond prefers to be localized on the ligand. The higher proton binding affinity of the ligand than that of Asp27 likely arises from the delocalization of the positive charge over the diaminopyrimidine ring, which helps stabilize its protonated state. The positively charged, N1-protonated diaminopyrimidine ring can thus form a strong salt bridge with the negatively charged, deprotonated Asp27. Removal of the proton from the ligand not only needs to overcome the intrinsic proton affinity of the diaminopyrimidine ring in isolation, but also break the strong salt−bridge interaction between the ligand and the Asp27. Therefore, this salt−bridge interaction further stabilizes the ligand's protonated state, greatly increasing its p$K_a$ in the active site of eDHFR. Adding an additional proton to the carboxyl group of Asp27 while the compound **6** is in the N1-protonated state also destabilizes this salt-bridge, leading to low pKa of the Asp27. This observation from the *ab initio* QM/MM PMF calculation is thus consistent with our MM-level pH-REMD simulations.

Importantly, the characterization of the ligand's protonation state has direct implications for the interpretation of binding free energy calculations discussed below. Without explicit consideration of the N1-atom's protonation state change upon protein binding, no free energy method tested could consistently predict the experimentally observed *cis*-vs-*trans* affinity differential or the substituent effects on it.

## METHOD BENCHMARK: *CIS*-VS-*TRANS* AFFINITY DIFFERENTIAL AND SUBSTITUENT EFFECTS

The experimental design campaign of the TMP-derived photoswitchable eDHFR inhibitor[6] has shown that chemical modifications can substantially alter the absolute and light-responsive affinity differential of these ligands. These substitutions were introduced at the *para* and *meta* positions of the phenyl ring in the azobenzene moiety. In particular, introducing a carboxylic acid group to compound **6**, yielding compound **11**, increases the ligand's inhibitory potency but eliminates the "*cis*-on" effect. Meanwhile, introducing an additional chlorine atom on the azobenzene moiety of compound **11**, yielding compound **15**, recovers the "*cis*-on" effect.

These prior experimental results provide a valuable data set for benchmarking free energy methods. Here, we evaluated the accuracy of TI and MM-PB/GBSA methods in predicting the *cis*-vs-*trans* affinity differential and its substituent effects. The experimental reference data set consists of relative binding affinities between different ligands (compounds **6**, **11**, and **15**) in the same isomeric form, as well as the *cis*-vs-*trans* affinity differential for each ligand. They were derived from the ratios of reported $IC_{50}$ values.[6]

1. TI simulations with different alchemical transformation pathways

Given its established accuracy for photoswitchable ligands targeting G-protein coupled receptor and tubulin,[7,8] the TI method was first tested. Due to finite sampling, the difficulty of converging the free energy changes often depends on the selected thermodynamic cycle and alchemical transformation pathways. Here, we compared two pathways. In the first pathway (Method A), the *cis* and *trans* isomers are interconverted by gradually making a phenyl ring attached to the N═N double bond disappear from one side and reappear on the other. Such an alchemical transformation was performed in both the aqueous solution and the eDHFR binding pocket. Based on the above-described p$K_a$ calculations, the N1 atom of all three ligands was first chosen to be deprotonated in the aqueous solution and protonated in the protein. This approach yields the *cis*-vs-*trans* affinity differential for each compound X through a thermodynamic cycle (Figure 3A, eq 2), i.e., $\Delta\Delta G_{X,cis\rightarrow trans}$. By taking the difference of this quantity for a pair of compounds, X and Y, the substituent effects on the *cis*-vs-*trans* affinity differential, $\Delta\Delta\Delta G_{X\rightarrow Y,cis\text{-vs-}trans\ \mathrm{binding}}$, can be obtained. *Accurately predicting this property is key to rational design aimed at increasing the ligand's light responsiveness via chemical substitution.* A positive value of this quantity indicates that changing the substituents from X to Y enhances the "*cis*-on" effect. The resulting $\Delta\Delta G_{X,cis\rightarrow trans}$ and $\Delta\Delta\Delta G_{X\rightarrow Y,cis\text{-vs-}trans\ \mathrm{binding}}$ obtained from Method A are summarized alongside experimental data in Tables 1 and 3.

In the second pathway (Method B), compound X is converted to Y in the same isomeric form (*cis* or *trans*) in both the aqueous solution and the protein. Through a thermody-

**Table 1. Calculated *cis*-vs-*trans* Affinity Differentials ($\Delta\Delta G_{X,cis\rightarrow trans}$, in kcal/mol) for Compounds 6, 11, and 15, Calculated Using TI (Method A)[a]**

| Quantity | Compounds | | |
|---|---|---|---|
| | 6 | 11 | 15 |
| $\Delta\Delta G_{X,cis\rightarrow trans}$ (Exp.) | 0.5 ± 0.5 | −0.1 ± 0.4 | 0.4 ± 0.2 |
| $\Delta G_{\mathrm{unbound},cis\ X\rightarrow trans\ X}$ (TI) | −11.9 ± 0.1 | −11.9 ± 0.1 | −11.6 ± 0.1 |
| $\Delta G_{\mathrm{bound},cis\ X\rightarrow trans\ X}$ (TI) | −11.3 ± 0.5 | −13.0 ± 0.1 | −11.3 ± 0.1 |
| $\Delta\Delta G_{X,cis\rightarrow trans}$ (TI) | 0.6 ± 0.5 | −1.1 ± 0.1 | 0.3 ± 0.1 |

[a]Experimental values[6] are included for comparison. Positive $\Delta\Delta G_{X,cis\rightarrow trans}$ values indicate "*cis*-on" effect, and *vice versa*. In these calculations, the ligands are N1-deprotonated and N1-protonated in the aqueous solution and protein, respectively.



https://doi.org/10.1021/acs.jcim.6c02133
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Table 2. Calculated Binding Free Energy Changes ($\Delta\Delta G_{\text{binding,X}\rightarrow\text{Y}}$, in kcal/mol) upon Converting Compound X to Y, Calculated Using TI (Method B)[a]

| Transformations | Quantities | | | |
|---|---|---|---|---|
| | $\Delta G_{\text{X}\rightarrow\text{Y, unbound}}$ | $\Delta G_{\text{X}\rightarrow\text{Y, bound}}$ | $\Delta\Delta G_{\text{binding,X}\rightarrow\text{Y}}$ (TI) | $\Delta\Delta G_{\text{binding,X}\rightarrow\text{Y}}$ (exp.) |
| | | Protocol 1 | | |
| **11** → **15** (*cis*) | −26.8 ± 0.5 | −4.2 ± 0.6 | 22.6 ± 0.8 | −0.6 ± 0.3 |
| **11** → **15** (*trans*) | −27.2 ± 0.3 | −4.3 ± 0.5 | 22.9 ± 0.6 | −0.1 ± 0.2 |
| **6** → **11** (*cis*) | −57.8 ± 0.0 | −8.2 ± 0.0 | 49.6 ± 0.0 | −0.5 ± 0.4 |
| **6** → **11** (*trans*) | −57.9 ± 0.1 | −5.8 ± 0.0 | 52.1 ± 0.1 | −1.1 ± 0.3 |
| | | Protocol 2 | | |
| **N1-protonated** | | | | |
| **11** → **15** (*cis*) | −3.1 ± 0.2 | −4.2 ± 0.6 | −1.1 ± 0.6 | −0.6 ± 0.3 |
| **11** → **15** (*trans*) | −3.8 ± 0.4 | −4.3 ± 0.5 | −0.5 ± 0.6 | −0.1 ± 0.2 |
| **6** → **11** (*cis*) | −7.6 ± 0.0 | −8.2 ± 0.0 | −0.6 ± 0.0 | −0.5 ± 0.4 |
| **6** → **11** (*trans*) | −7.4 ± 0.1 | −5.8 ± 0.0 | 1.6 ± 0.1 | −1.1 ± 0.3 |
| **N1-deprotonated** | | | | |
| **11** → **15** (*cis*) | −26.8 ± 0.5 | −27.0 ± 0.7 | −0.2 ± 0.9 | −0.6 ± 0.3 |
| **11** → **15** (*trans*) | −27.2 ± 0.3 | −26.4 ± 0.4 | 0.8 ± 0.5 | −0.1 ± 0.2 |
| **6** → **11** (*cis*) | −57.8 ± 0.0 | −55.4 ± 0.5 | 2.4 ± 0.5 | −0.5 ± 0.4 |
| **6** → **11** (*trans*) | −57.9 ± 0.1 | −55.8 ± 0.1 | 2.1 ± 0.1 | −1.1 ± 0.3 |

[a]Experimental values[6] are included for comparison. Negative values of $\Delta\Delta G_{\text{binding,X}\rightarrow\text{Y}}$ correspond to an increase in binding affinity upon converting X to Y, and *vice versa*. Calculations were performed using two different protocols. In Protocol 1, the ligands were in the N1-protonated state in the protein and the N1-deprotonated state in the aqueous solution. In Protocol 2, the ligands were in the same protonation state (either N1-protonated or N1-deprotonated) in both environments.

Table 3. Substituent Effects on the Light-Responsive Affinity Differential ($\Delta\Delta\Delta G_{\text{X}\rightarrow\text{Y},cis\text{-vs-}trans\text{ binding}}$ in kcal/mol) upon Converting X to Y, Evaluated Using TI (Methods a and B), and Compared with Experimental Values[6].[a]

| Approach | $\Delta\Delta\Delta G_{\text{X}\rightarrow\text{Y},cis\text{-vs-}trans\text{ binding}}$ | | |
|---|---|---|---|
| | **6** → **15** | **6** → **11** | **11** → **15** |
| Experiment | −0.1 ± 0.5 | −0.6 ± 0.6 | 0.5 ± 0.4 |
| TI (Method A) | −0.3 ± 0.5 | −1.7 ± 0.5 | 1.4 ± 0.1 |
| TI (Method B, N1-protonated) | 2.8 ± 0.8 | 2.2 ± 0.1 | 0.6 ± 0.8 |
| TI (Method B, N1-deprotonated) | 0.7 ± 1.1 | −0.3 ± 0.5 | 1.0 ± 1.0 |

[a]Positive values indicate that X → Y conversion enhances the "*cis*-on" effects, and *vice versa*.

namic cycle (Figure 3B), this approach estimates the relative binding free energy between the two compounds in identical isomeric form and protonation state, i.e., $\Delta\Delta G_{\text{binding},cis\text{X}\rightarrow cis\text{Y}}$ and $\Delta\Delta G_{\text{binding},trans\text{X}\rightarrow trans\text{Y}}$ (Figure 3B and eqs 3 &4). The difference between these two $\Delta\Delta G_{\text{binding}}$ values corresponds to the substituent effects on the *cis*-vs-*trans* affinity differential, $\Delta\Delta\Delta G_{\text{X}\rightarrow\text{Y},cis\text{-vs-}trans\text{ binding}}$, which can also be evaluated by Method A. The ligand's protonation state in solution and in the protein can be the same or different. Tables 2 and 3 summarize the results obtained from this approach with the same protonation state in solution and protein (reasons for this choice are discussed below), and experimental values. In our previous studies, this method has provided the best quantitative interpretation of experimentally observed trends for a series of photoswitchable ligands targeting tubulin and membrane receptors.[7,8]

Using Method A, the *cis*-vs-*trans* affinity differentials ($\Delta\Delta G_{\text{X},cis\rightarrow trans}$) for all three compounds **6**, **11** and **15** reproduced the experiment with near-quantitative accuracy (Table 1). The TI simulations correctly predicted the signs of *cis*-vs-*trans* affinity differentials ($\Delta\Delta G_{\text{X},cis\rightarrow trans}$), i.e., the "*cis*-on" or "*trans*-on" effects. The positive $\Delta\Delta G_{\text{X},cis\rightarrow trans}$ values of compounds **6** and **15** indicate "*cis*-on" effect, whereas the small negative $\Delta\Delta G_{\text{X},cis\rightarrow trans}$ of compound **11** indicates a slight "*trans*-on" effect. Beyond this, Method A also performs well in predicting how chemical substitution alters the *cis*-vs-*trans* affinity differential (Table 3). The calculated $\Delta\Delta\Delta G_{\text{X}\rightarrow\text{Y},cis\text{-vs-}trans\text{ binding}}$ values indicate that the **6** → **11** transformation (introducing a carboxylate group) neutralizes the "*cis*-on" effect and results in a slight "*trans*-on" effect. In contrast, the **11** → **15** transformation (introducing a chlorine atom in addition to the carboxylate group) recovers the "*cis*-on" effect of compound **6**. These calculated results are all consistent with experiments, indicating that Method A is a robust framework for predicting both active isomer and substituent effects on their light-responsive affinity differential. The ability of Method A to quantitatively predict whether a ligand is *cis*-on or *trans*-on, and how chemical modifications magnify, mitigate, or reverse these effects, is critical for rational design of photoswitchable ligands in photopharmacology.

Method A does not perform as well if the protonation states are set to be the same in both the aqueous solution and protein environments, as shown in Table S1. Specifically, when the ligand is set as N1-deprotonated in both environments, the calculated $\Delta\Delta G_{\text{X},cis\rightarrow trans}$ for compounds **11** and **15** are largely overestimated, and the $\Delta\Delta\Delta G_{\text{X}\rightarrow\text{Y},cis\text{-vs-}trans\text{ binding}}$ values fail to predict the sign of experimental data for any of the three transformations, generating the worst prediction. When the ligand is set as N1-protonated in both environments, the calculated $\Delta\Delta G_{\text{X},cis\rightarrow trans}$ for compound **11** is largely underestimated, so is the $\Delta\Delta\Delta G_{\text{X}\rightarrow\text{Y},cis\text{-vs-}trans\text{ binding}}$ for **6** → **11** transformation. Also, the calculated $\Delta\Delta\Delta G_{\text{X}\rightarrow\text{Y},cis\text{-vs-}trans\text{ binding}}$ for the **6** → **15** has the incorrect sign. These negative results




https://doi.org/10.1021/acs.jcim.6c02133
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

highlight that the binding-induced changes in the ligand's protonation state must be explicitly accounted for in TI simulations to obtain accurate predictions of light-responsive affinity differential and substituent effects.

Next, we benchmarked the accuracy of Method B. This method uses the free energy change of alchemical transformation from X to Y in the protein and aqueous solution (e.g., $\Delta G_{trans\mathrm{X}\rightarrow trans\mathrm{Y},\mathrm{bound}}$ and $\Delta G_{trans\mathrm{X}\rightarrow trans\mathrm{Y},\mathrm{unbound}}$) to estimate the relative binding affinity between a pair of compounds (e.g., $\Delta\Delta G_{\mathrm{binding},trans\mathrm{X}\rightarrow trans\mathrm{Y}}$). In principle, one should simulate the alchemical transformation in the protein-bound state (e.g., $\Delta G_{trans\mathrm{X}\rightarrow trans\mathrm{Y},\mathrm{bound}}$) with N1-protonated X and Y, and that in the aqueous solution state (e.g., $\Delta G_{trans\mathrm{X}\rightarrow trans\mathrm{Y},\mathrm{unbound}}$) with N1-deprotonated X and Y (Protocol 1). Such a choice would, ideally, reflect the binding-induced change in protonation state. Alternatively, one could make a suboptimal choice of setting the ligands to the same protonation state (either N1-protonated or N1-deprotonated) in both the solution and the protein (Protocol 2), which could introduce systematic errors in and $\Delta\Delta G_{\mathrm{binding},cis\mathrm{X}\rightarrow cis\mathrm{Y}}$. Note that binding-induced changes in the ligand's protonation state were naturally incorporated in the experimental data. Comparing the performance of Protocols 1 and 2 enables a systematic analysis of how different treatments of the ligand's protonation state affect the accuracy of free energy calculations.

The calculated $\Delta\Delta G_{\mathrm{binding},trans\mathrm{X}\rightarrow trans\mathrm{Y}}$ and $\Delta\Delta G_{\mathrm{binding},cis\mathrm{X}\rightarrow cis\mathrm{Y}}$ calculated using Method B (Protocol 2) are reported in the Table 2, and the *cis*-vs-*trans* affinity differential $\Delta\Delta\Delta G_{\mathrm{X}\rightarrow\mathrm{Y},cis\text{-vs-}trans\ \mathrm{binding}}$ in Table 3. For the **11** → **15** transformation in the N1-protonated state, Method B (Protocol 2) captures the experimental trend reasonably well. The calculated $\Delta\Delta G_{\mathrm{binding},cis\ \mathrm{X}\rightarrow cis\ \mathrm{Y}}$ (−1.1 ± 0.6 kcal/mol) indicates that introducing an additional chlorine substituent alongside the carboxylate group strengthens the binding affinity of *cis* **15** relative to *cis* **11**, in agreement with the experimental value of −0.6 ± 0.3 kcal/mol. A similar stabilization is observed for the *trans* isomer, where both simulation (−0.5 ± 0.6 kcal/mol) and experiment (−0.1 ± 0.2 kcal/mol) suggest a slight increase in affinity upon conversion from *trans* **11** to *trans* **15**. These results indicate that, for this specific modification, Method B reproduces the qualitative effect of substitution on binding affinity of both isomeric forms. For the *cis* **6**→ *cis* **11** transformation in the protonated system in the N1-protonated state, Method B predicts stronger binding upon introducing the carboxylic acid group for the *cis* isomer, in line with experimental data. This is mainly due to the formation of a favorable salt−bridge interaction between the negatively charged carboxylate on *cis* **11** and the Arg57 residue in eDHFR, which is absent in *cis* **6**. However, the performance of Method B (Protocol 2) deteriorates for the *trans* **6** → *trans* **11** transformation. The experimental data indicate an increase in the binding affinity after this transformation ($\Delta\Delta G_{\mathrm{binding},trans\mathrm{X}\rightarrow trans\mathrm{Y}}$ = −1.1 ± 0.3 kcal/mol), whereas the calculation indicates a decrease in the binding affinity (1.6 ± 0.1 kcal/mol).

For the N1-deprotonated state of the ligands, Method B (Protocol 2) cannot consistently capture the experimental trend. For the *trans* **11** → *trans* **15** transformation, the calculations predict a decrease in the binding affinity ($\Delta\Delta G_{\mathrm{binding},trans\mathrm{X}\rightarrow trans\mathrm{Y}}$ = 0.8 ± 0.5 kcal/mol). This result is qualitatively inconsistent with the experimental value (−0.1 ± 0.2 kcal/mol). A similar loss of accuracy is observed for the **6** → **11** transformation, where the calculations predict a decrease in binding affinity for both isomers, in contrast with the increase in binding affinity observed from experimental data (Table 2). The limitations of Method B (Protocol 2) become more obvious for $\Delta\Delta\Delta G_{\mathrm{X}\rightarrow\mathrm{Y},cis-\mathrm{vs}-\mathrm{transbinding}}$ (Table 3). For the **11** → **15** transformation, Method B correctly predicts an enhancement of the "*cis*-on" effect for both protonated and deprotonated ligands, consistent with experiment. However, for the **6** → **11** transformation, Method B (Protocol 2) fails to reproduce the negative $\Delta\Delta\Delta G_{\mathrm{X}\rightarrow\mathrm{Y},cis\text{-vs-}trans\ \mathrm{binding}}$ from the experiments. This sign error indicates that this method cannot reliably resolve the delicate shift in the *cis*-vs-*trans* affinity differential for this transformation. Since Method B (Protocol 2) uses the difference between $\Delta\Delta G_{\mathrm{binding},cis\mathrm{X}\rightarrow cis\mathrm{Y}}$ and $\Delta\Delta G_{\mathrm{binding},trans\mathrm{X}\rightarrow trans\mathrm{Y}}$ to calculate $\Delta\Delta\Delta G_{\mathrm{X}\rightarrow\mathrm{Y},cis\text{-vs-}trans\ \mathrm{binding}}$, small inaccuracies in either of the two $\Delta\Delta G_{\mathrm{binding}}$ quantities, as discussed above (Table 2), can propagate and lead to a qualitatively incorrect prediction of $\Delta\Delta\Delta G_{\mathrm{X}\rightarrow\mathrm{Y},cis\text{-vs-}trans\ \mathrm{binding}}$.

It is worth noting that a source of systematic error in Method B (Protocol 2)'s estimates of $\Delta\Delta G_{\mathrm{binding},trans\mathrm{X}\rightarrow trans\mathrm{Y}}$ and $\Delta\Delta G_{\mathrm{binding},cis\mathrm{X}\rightarrow cis\mathrm{Y}}$ is the erroneous assumption that the ligand's protonation state remains the same in both aqueous solution and protein. However, in practice, the suboptimal Protocol 2 had to be chosen deliberately because of the poor results from the conceptually optimal Protocol 1. When choosing Method B (Protocol 1), the $\Delta\Delta G_{\mathrm{binding},trans\mathrm{X}\rightarrow trans\mathrm{Y}}$ and $\Delta\Delta G_{\mathrm{binding},cis\mathrm{X}\rightarrow cis\mathrm{Y}}$ deviate much more from the experiment than Method B (Protocol 2) (Table 2), often by 1−2 orders of magnitude. For example, the $\Delta\Delta G_{\mathrm{binding},cis\mathrm{X}\rightarrow cis\mathrm{Y}}$ for *cis* **6**→ *cis* **11** transformation estimated by Protocol 1 is ∼ 50 kcal/mol, in sharp contrast to the experimental value of −0.5 ± 0.4 kcal/mol and the Protocol 2's estimated values of −0.6 ± 0.0 kcal/mol (protonated state) and 2.4 ± 0.5 kcal/mol (deprotonated state). The reason for the poor performance of Protocol 1 is as follows: ligands change protonation state upon binding with the protein. When calculating the relative binding free energies using Protocol 1, its associated thermodynamic cycle (Figure 3C) entails addressing two difficult tasks together: (1) estimating relative binding free energies in the same protonation state between X and Y, (2) estimating the relative p$K_a$ between X and Y. For task 2, when the force field parameters are not accurately parametrized for this specific purpose, small perturbations in the fitted point charges in both protonation states can readily lead to large errors in the predicted p$K_a$ without *ad hoc* reference correction (See Discussion). Although these p$K_a$ values are not explicitly calculated for a pair of compounds connected by alchemical transformation, their large errors must cancel out each other to predict correct relative binding free energies. However, such error cancellations are *not* guaranteed, leading to large errors in the $\Delta\Delta G_{\mathrm{binding},trans\mathrm{X}\rightarrow trans\mathrm{Y}}$ and $\Delta\Delta G_{\mathrm{binding},cis\mathrm{X}\rightarrow cis\mathrm{Y}}$ estimates. As a result, the prediction of $\Delta\Delta\Delta G_{\mathrm{X}\rightarrow\mathrm{Y},cis\text{-vs-}trans\ \mathrm{binding}}$, which is defined as the difference between these two quantities, becomes very unreliable.

In contrast, Protocol 2 involves only task 1 (Figure 3B). For this reason, the best performance of Method B was achieved





https://doi.org/10.1021/acs.jcim.6c02133
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

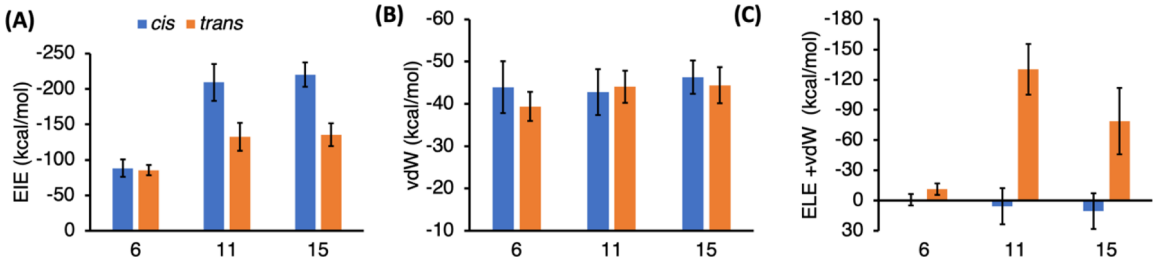



**Figure 6.** (A) Electrostatic interaction energy (EIE) and (B) van der Waals (vdW) interaction energy between eDHFR and TMP derivatives in their *cis* and *trans* isomers. (C) The sum of electrostatic and vdW ligand-water interaction energies. The ligands are in complex with eDHFR in all analyses.

with Protocol 2, i.e., setting the same protonation state for each ligand in both the protein and the aqueous solution. This approach removes the dominant source of error arising from p$K_a$ differences and focuses on the effects of chemical substitution on binding affinity in the same protonation state. While this treatment does not explicitly capture the coupling between protonation and binding, it provides a more physically interpretable estimate of substituent effects within the limitations of nonreactive force fields. It is worth noting that in our previous studies,[7,8] Method B was identified as the most accurate method for estimating substituent effects on light-responsive affinity differentials of photoswitchable inhibitors of tubulin and $\beta$-adrenergic receptors.

The comparative analysis highlights a critical limitation of Method B when applied to ionizable photoswitchable ligands that undergo changes in protonation state upon protein binding. The pH-REMD calculations showed that the ligands are predominantly N1-deprotonated in aqueous solution, whereas binding to eDHFR strongly favors the N1-protonated state. This shift in protonation state directly affects the thermodynamic cycle used in free-energy calculations. Method A explicitly accounts for the dominant protonation state in each environment, while Method B (Protocol 2) assumes a fixed protonation state throughout the calculation. As a result, Method A provides better agreement with the experimental measurements, confirming that binding-induced changes in protonation state make an important contribution to the calculated differences in light-responsive affinity. These results emphasize that accurate treatment of protonation equilibria is essential for reliable free-energy calculations of ionizable photoswitches and should be considered in their design for photopharmacological applications.

2. MM-PBSA and MM-GBSA

The *cis*-vs-*trans* affinity differentials ($\Delta\Delta G_{X,cis\rightarrow trans}$) obtained from MM-PB(GB)SA calculations are summarized in Table S2. The magnitude of the calculated values deviates significantly from the experimental values for all compounds. Additionally, both methods consistently predict a "*cis*-on" effect for all three compounds. However, this is inconsistent with experimental observations for **11**, which has *a "trans*-on" effect. This indicates that MM-GBSA is not able to consistently identify "*cis*-on" or "*trans*-on" effects qualitatively correctly.

The limitations of these methods become even more obvious for $\Delta\Delta\Delta G_{X\rightarrow Y,cis\text{-vs-}trans\ \text{binding}}$ (Table S3). MM-GBSA predicts large positive $\Delta\Delta\Delta G_{X\rightarrow Y,cis\text{-vs-}trans\ \text{binding}}$ values for all three transformations (**6** → **11**, **6** → **15**, and **11** → **15**), indicating a strong enhancement of "*cis*-on" effect upon these transformations. The magnitudes of these values are an order of magnitude higher than the experimental values, and the signs of them are both wrong (Table S3). These results indicate that, even after error cancellation by subtracting the $\Delta\Delta G_{X,cis\rightarrow trans}$ values, MM-GBSA still cannot qualitatively predict substituent effects on the differential light-responsive affinity. MM-PBSA, on the other hand, predicts negative $\Delta\Delta\Delta G_{X\rightarrow Y,cis\text{-vs-}trans\ \text{binding}}$ values for all three transformations, suggesting that they reduce the "*cis*-on" effect. While this qualitative trend partially agrees with the experiment for **6** → **11** and **6** → **15**, it is inconsistent for **11** → **15**, where the experiment indicates an increase in "*cis*-on" effect. Overall, these end-point free energy methods cannot qualitatively predict the *cis*-vs-*trans* affinity differentials or substituent effects.

## Protein−ligand Interactions in eDHFR

To understand the molecular basis of the *cis*-vs-*trans* affinity differentials of the TMP derivatives, we analyzed electrostatic and van der Waals (vdW) interaction energies between the ligands and protein, as well as between the ligands and surrounding water molecules, using unbiased production MD trajectories (Figure 6). The ligand-water interaction energies were calculated from protein-bound trajectories in which the ligands were in their N1-protonated state. Therefore, these values show the interactions between the bound ligand and nearby water molecules within or around the binding pocket. We also compared representative binding poses in the eDHFR active site (Figure 7).

The energy decomposition reveals distinct interaction patterns among the three compounds. For compounds **11** and **15**, the *cis* isomers exhibit more favorable electrostatic interactions with eDHFR than the corresponding *trans* isomers, whereas compound **6** shows only a small *cis*-vs-*trans* electrostatic interaction energy difference (Figure 6A). In contrast, vdW interaction energies are relatively similar between *cis* and *trans* isomers for all three compounds (Figure 6B). These results suggest that electrostatic interactions, rather than vdW contacts, make the dominant contribution to the *cis*-vs-*trans* differential in the protein−ligand interaction energy.

For compounds **6** and **15**, the protein−ligand interaction analysis is consistent with the experimental and TI results showing "*cis*-on" effect. For compound **11**, however, the stronger electrostatic interaction of *cis*-**11** with eDHFR would appear to suggest "*cis*-on" effect, whereas both experiment and TI indicate a weak "*trans*-on" effect. This discrepancy indicates that protein−ligand interactions alone are insufficient to explain the *cis*-vs-*trans* affinity differential of compound **11**. We therefore further analyzed ligand-water interaction energies (Figure 6C). The *trans*-**11** shows the strongest stabilization from interacting with water among all systems, likely due to its




https://doi.org/10.1021/acs.jcim.6c02133
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

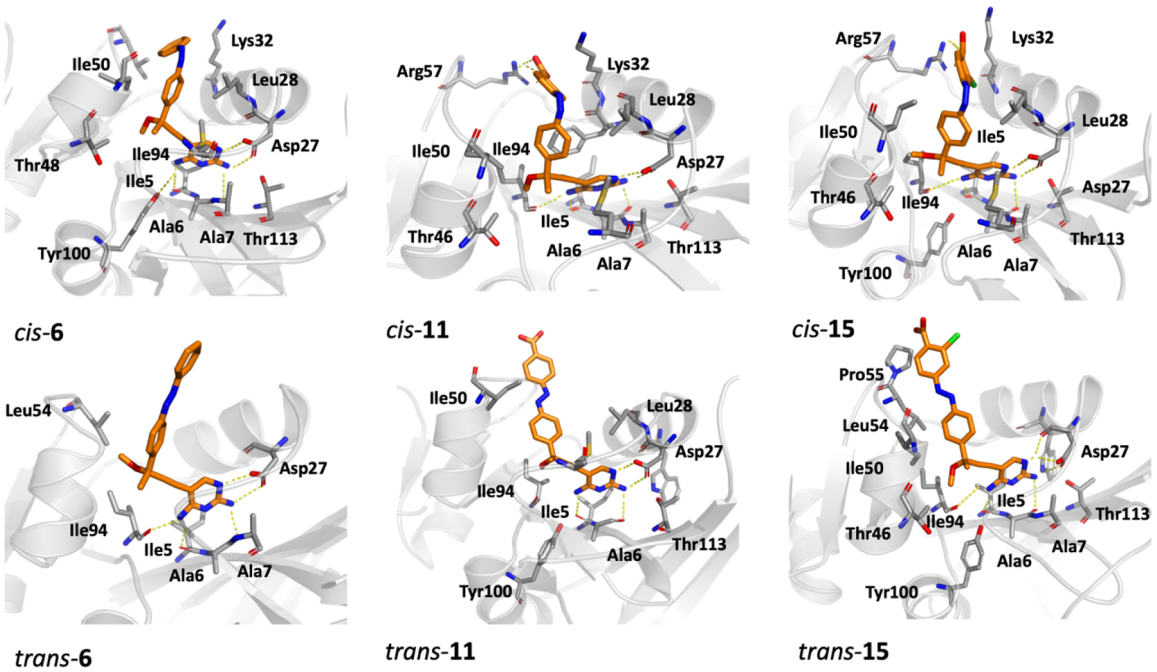



**Figure 7.** Binding poses of *cis*- and *trans*-TMP derivatives in complex with the eDHFR. The isomeric form of the ligands strongly influences their binding mode and the protein−ligand interactions in the binding site.

greater solvent exposure in the binding pocket (Figure 7). This enhanced ligand-water interaction of *trans*-**11** can compensate for its weaker ligand−protein interactions and shift the overall binding preference toward *trans*-**11** over *cis*-**11**. By contrast, for compound **15**, the "*cis*-on" effect due to ligand−protein interaction is stronger than **11**, while "*trans*-on" effect due to ligand-water interaction is weaker than **11**. These combined effects enhance the "*cis*-on" effect of compound **15**, consistent with experiment and TI simulations.

Representative binding poses further reveal the molecular origins of these trends (Figure 7). For compound **6**, both isomers maintain the conserved hydrogen bond between the diaminopyrimidine N1 atom and Asp27, while *cis*-**6** forms closer contacts with Ala6, Ala7, Ile5, and Ile94. For compound **11**, Asp27 remains the primary stabilizing residue for both isomers, but the added carboxylate group strengthens electrostatic interactions in *cis*-**11** through hydrogen bond with Lys32 and Arg57. However, this stronger protein−ligand stabilization of *cis*-**11** is offset by stronger ligand-water stabilization of *trans*-**11**. For compound **15**, *cis*-**15** forms a cooperative interaction network: Asp27 stabilizes the diaminopyrimidine ring, Lys32 and Arg57 interact favorably with the carboxylate group, and Ile5, Ile94, Leu28, and Ile50 create a compact hydrophobic environment around the ligand's aromatic rings. In *trans*-**15**, these contacts are weakened because the ligand extends away from the binding cavity. Overall, the active isomer is determined by a balance between protein−ligand electrostatics, hydrophobic packing, and ligand solvation rather than by interactions with a single residue.

Admittedly, without quantifying ligand-water interactions in bulk solution via rigorous solvation free-energy calculations, this standard MD-based analysis is only qualitative rather than quantitative. However, such analyses provide useful insight into the balance between protein−ligand and ligand-water interactions when the ligand is bound. This qualitative trend explains the light-responsive binding behaviors observed from experiments and our more rigorous TI free-energy calculations, which explicitly account for solvation free energy in both aqueous solution and protein-bound systems.

## DISCUSSION

Photoswitchable ligands with titratable groups whose protonation-state equilibrium shifts in response to changes in their molecular environments represent an important class of compounds with various applications in photopharmacology. To uncover how their light-responsive binding affinities are coupled to changes in protonation, we performed a comprehensive computational investigation of TMP-derived photoswitchable inhibitors of eDHFR, developed recently through a systematic, hypothesis-driven, and structure-based experimental design campaign.[6] The pH-REMD and *ab initio* QM/MM MD US simulations established that the ligand's diaminopyrimidine ring is protonated upon protein binding. The protonation state change of the ligand is due to the significant increase in the p$K_a$ of this functional group, which is attributed to the strong salt bridge interaction with the Asp27 residue in the binding pocket.

Additionally, the performance of TI with different choices of alchemical transformation pathways and ligand protonation states was benchmarked against experimental data.[6] The benchmark focused on the quantitative accuracy for predicting two critical properties for photopharmacology: (1) the *cis*-vs-*trans* affinity differential of an individual ligand, and (2) the substituent effects on the affinity differentials among a group of ligands. TI simulations employing a *cis*-to-*trans* alchemical transformation pathway while accounting for the protonation-state change upon protein binding (Method A) emerged as the most robust approach. Importantly, without explicitly accounting for binding-induced changes in protonation state, this approach fails to faithfully reproduce the substituent effects on the experimentally observed affinity differentials.

TI simulations employing compound-to-compound alchemical transformations (Method B) were less robust. When




https://doi.org/10.1021/acs.jcim.6c02133
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

explicitly treating protonation state changes of the ligand, Method B's performance relies on cancellations of large errors in the p$K_a$ prediction of distinct compounds if the ligand's protonation state change is explicitly treated. The p$K_a$ errors arise from inaccuracies in the force field parameters, which typically do not cancel out between a pair of compounds connected by the alchemical transformation. When not explicitly treating the ligand's protonation state changes, Method B cannot consistently provide qualitatively correct *cis*-vs-*trans* affinity differentials or the substituent effects on them.

One could, in principle, perform rigorous QM-based calculations to correct the p$K_a$ for each compound in each isomer form to remedy this issue of Method B. However, such an approach would not only incur significantly higher computational cost but may also be infeasible for large, flexible molecular photoswitches whose conformational distributions depend on the ligand's protonation state. In our aqueous solution MD simulations, the ligand can sample many conformations, which differ in different protonation states. A simple QM-based approach employing geometry minimization and single-point energy calculation is not compatible with the large conformational ensemble of protonated and deprotonated states combined. More complex methods based on QM/MM free-energy calculations could be applied, but this would incur higher computational costs, since they need to be applied to each isomer of each ligand.

Our results also indicate that the MM-PB(GB)SA methods cannot consistently predict either the *cis*-vs-*trans* affinity differential or the substituent effects. Also, standard MD simulations suggest that the light-responsive affinity differential arises from a balance of protein−ligand and ligand-water interactions, whereas simply considering protein−ligand interaction can lead to qualitatively incorrect interpretations and predictions, as highlighted by the case of compound **11**. This compound exhibits a cancellation of strong "*cis*-on" and "*trans*-on" effects arising from protein−ligand and ligand-water interactions, respectively. Additionally, the light-responsive affinity differential is a result of a network of ligand−protein interactions. Persistent H-bond to key residues such as Arg57 and Lys32 stabilizes the *cis*-**15** over *trans*-**15**, contributing to its strong "*cis*-on" effect.

Our earlier studies[1,2] addressed the free energy calculation of light-responsive affinity differentials for photoswitches without a binding-induced change in ligand protonation. Among the tested approaches for these systems, the TI method employing the compound-to-compound transformation pathway (Method B in this study) had the most consistent and quantitative agreement with experiment, because it largely avoids the sampling convergence issues associated with large structural reorganization in the protein compared to other free energy methods directly sampling the binding/unbinding pathways.

The present study addresses a new question: does the same free energy method (Method B) remains reliable when ligand binding changes the dominant protonation state? The pH-REMD and QM/MM free energy calculations show that the TMP-derived ligands are predominantly N1-deprotonated in aqueous solution but become N1-protonated in the eDHFR binding site. Once this protonation change is included in the TI employing the compound-to-compound transformation pathway, the thermodynamic cycle becomes sensitive to errors in the relative p$K_a$ values of distinct ligands. As a result, the compound-to-compound TI pathway that performed the best in the earlier systems[1,2] becomes unreliable here, whereas the *cis*-to-*trans* TI pathway performs better when the dominant protonation state is assigned differently in bulk solution and in the protein, since the errors in the relative p$K_a$ values of distinct isomers of the same ligand largely cancel out each other. A key methodological insight provided by the present work is therefore that the reliability of an alchemical pathway cannot be separated from the ligand's acid-base reactivity: a pathway that works well for ligands with fixed protonation states may become unreliable when protonation and binding are coupled.

Beyond this specific system, the methodological contribution from this work has broader implications for the design of photoswitchable ligands. The ability to computationally predict small *cis*-vs-*trans* affinity differentials and the substituent effects is essential for the rational design of photoswitchable ligands, where energetic errors of less than 1 kcal/mol can completely reverse the functional outcome of the biomolecular system under photocontrol. The present results provide useful guidelines for the design of ionizable photoswitches in photopharmacology. First, changes in protonation state upon protein binding should be explicitly considered during ligand design, since binding-induced shifts in p$K_a$ can influence the light-responsive affinity differential and the extent to which substituents affect this quantity. Second, substituent modifications can affect the balance between ligand-water and protein−ligand interactions, thereby shifting the light-responsive affinity differential and even changing the identity of the active isomer. Together, these findings indicate that successful optimization of ionizable photoswitches requires balancing protonation equilibrium, protein−ligand interactions, and solvent effects rather than focusing solely on binding affinity.

One limitation of this study is the relatively small number of experimentally characterized compounds available for benchmarking. The analysis was performed using three TMP-derived photoswitchable eDHFR inhibitors for which quantitative measurements of both *cis* and *trans* binding affinities and substituent effects were reported. However, it is worth noting that the three ligands were *not* randomly chosen. They represent three key stages of a previous experimental hypothesis-driven design campaign for effective photoswitchable ligands targeting DHFR,[6] a well-established system for traditional drug design. The design principles employed in this systematic experimental study were representative of and applicable to a wide range of biomolecular systems in photopharmacology. By selecting these three compounds in our computational study, our aim was *not* to perform a large-scale benchmark across diverse chemical series and biomolecules. Instead, we aim to elucidate the molecular origins of these experimental design principles in a carefully characterized model system, such as the effects of introducing salt−bridge interactions, changes in ligand solvation, and changes in protonation state on the light-responsive binding behavior of these ligands. The availability of high-quality experimental data made these three compounds particularly suitable for our goals.

Additionally, in the present work, we do *not* aim to establish a universal ranking of free energy methods for all ionizable photoswitches. Instead, we aim to demonstrate that explicitly accounting for protein-binding-induced changes in protonation state can improve predictions of the light-responsive binding behavior of photoswitchable ligands in experimentally well-characterized systems. To the best of our knowledge, this is the




https://doi.org/10.1021/acs.jcim.6c02133
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

first study to investigate how changes in protonation state influence the binding affinity difference between the *cis* and *trans* isomers of photoswitchable ligands. Our findings highlight protonation as an important factor in accurately predicting light-induced affinity differential and provide valuable insight for the rational design of ionizable photoswitchable compounds. Although this data set is sufficient to evaluate the impact of binding-induced changes in protonation state on free-energy calculations within this series of ligands, the improved performance of Method A observed in this work should be viewed as a case study for this particular system. A comprehensive assessment across the broader chemical space of ionizable photoswitches is needed in future studies. Additional validation with larger and more diverse classes of photoswitchable ligands will be important for establishing the approach's general applicability.

## ■ ASSOCIATED CONTENT

### Data Availability Statement

We refer the readers to the Supporting Information file for additional data related to the properties of the molecular systems investigated in this manuscript. The structures and inputs for the initial simulation setup of the eDHFR systems are included in the "setup_files.zip" file.

### Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.6c02133.

Figures S1−S2 and Tables S1−S3, as referenced throughout the main text. The figures include the MM and QM PES of compound **15** in vacuum, and the QM/MM partitioning scheme. The table includes the alchemical free energy changes from TI and MM/PBSA(GBSA) simulations. It also contains structures for the initial simulation setup of the systems (PDF)

## ■ AUTHOR INFORMATION


### Corresponding Author

**Ruibin Liang** − *Department of Chemistry and Biochemistry, Texas Tech University, Lubbock 79409 Texas, United States;* orcid.org/0000-0001-8741-1520; Email: rliang@ttu.edu

### Authors

**Mohammad Khavani** − *Department of Chemistry and Biochemistry, Texas Tech University, Lubbock 79409 Texas, United States*

**Kambham Devendra Reddy** − *Department of Chemistry and Biochemistry, Texas Tech University, Lubbock 79409 Texas, United States*

**Pauf Neupane** − *Department of Chemistry and Biochemistry, Texas Tech University, Lubbock 79409 Texas, United States*

**Gustavo J. Costa** − *Department of Chemistry and Biochemistry, Texas Tech University, Lubbock 79409 Texas, United States*

**Laleh Khalvati** − *Department of Chemistry and Biochemistry, Texas Tech University, Lubbock 79409 Texas, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.6c02133


### Author Contributions

†M.K. and K.D.R. Equal contribution. Ruibin Liang designed the research project. Mohammad Khavani, Kambham Devendra Reddy, Pauf Neupane, Gustavo J. Costa and Laleh Khalvati performed the simulations and analyzed the data. Ruibin Liang, Mohammad Khavani, Kambham Devendra Reddy, Pauf Neupane, and Laleh Khalvati wrote and revised the manuscript.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health (grant number: R35GM150780) and the Texas University Fund (TUF) for the Human Molecular Aging Center at Texas Tech University under internal award #TUF6656. We also acknowledge the computing facilities provided by the High-Performance Computing Center at Texas Tech University.

## ■ REFERENCES


(1) (a) 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* (8), 6114−6178. (b) Kobauri, P.; Dekker, F. J.; Szymanski, W.; Feringa, B. L. Rational Design in Photopharmacology with Molecular Photoswitches. *Angew. Chem. Int.* **2023**, *62* (30), No. e202300681. (c) Niraula, S.; Seruga, B.; Ocana, A.; Shao, T.; Goldstein, R.; Tannock, I. F.; Amir, E. The price we pay for progress: a meta-analysis of harms of newly approved anticancer drugs. *J. Clin. Oncol.* **2012**, *30* (24), 3012−3019. (d) Wise, L.; Parkinson, J.; Raine, J.; Breckenridge, A. New approaches to drug safety: a pharmacovigilance tool kit. *Nat. Rev. Drug Discovery* **2009**, *8* (10), 779−782.

(2) (a) Beharry, A. A.; Woolley, G. A. Azobenzene photoswitches for biomolecules. *Chem. Soc. Rev.* **2011**, *40* (8), 4422−4437. (b) Lerch, M. M.; Hansen, M. J.; van Dam, G. M.; Szymanski, W.; Feringa, B. L. Emerging targets in photopharmacology. *Angew. Chem. Int.* **2016**, *55* (37), 10978−10999. (c) Hüll, K.; Morstein, J.; Trauner, D. In Vivo Photopharmacology. *Chem. Rev.* **2018**, *118* (21), 10710−10747.

(3) (a) Broichhagen, J.; Frank, J. A.; Trauner, D. A Roadmap to Success in Photopharmacology. *Acc. Chem. Res.* **2015**, *48* (7), 1947−1960. (b) Tochitsky, I.; Kienzler, M. A.; Isacoff, E.; Kramer, R. H. Restoring Vision to the Blind with Chemical Photoswitches. *Chem. Rev.* **2018**, *118* (21), 10748−10773. (c) 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. (d) Klippenstein, V.; Mony, L.; Paoletti, P. Probing Ion Channel Structure and Function Using Light-Sensitive Amino Acids. *Trends Biochem. Sci.* **2018**, *43* (6), 436−451. (e) Mart, R. J.; Allemann, R. K. Azobenzene photocontrol of peptides and proteins. *Chem. Commun.* **2016**, *52* (83), 12262−12277. (f) Gautier, A.; Gauron, C.; Volovitch, M.; Bensimon, D.; Jullien, L.; Vriz, S. How to control proteins with light in living systems. *Nat. Chem. Biol.* **2014**, *10*, 533. (g) Borowiak, M.; Nahaboo, W.; Reynders, M.; Nekolla, K.; Jalinot, P.; Hasserodt, J.; Rehberg, M.; Delattre, M.; Zahler, S.; Vollmar, A.; Trauner, D.; Thorn-Seshold, O. Photoswitchable Inhibitors of Microtubule Dynamics Optically Control Mitosis and Cell Death. *Cell* **2015**, *162* (2), 403−411.

(4) 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* (46), 15764−15773.

(5) Bosma, R.; Dijon, N. C.; Zheng, Y.; Schihada, H.; Hauwert, N. J.; Shi, S.; Arimont, M.; Riemens, R.; Custers, H.; van de Stolpe, A.; Vischer, H. F.; Wijtmans, M.; Holliday, N. D.; Kuster, D. W. D.; Leurs, R. Optical control of the $\beta$2-adrenergic receptor with opto-

prop-2: A cis-active azobenzene analog of propranolol. *iScience* **2022**, *25* (9), 104882.
(6) Kobauri, P.; Galenkamp, N. S.; Schulte, A. M.; de Vries, J.; Simeth, N. A.; Maglia, G.; Thallmair, S.; Kolarski, D.; Szymanski, W.; Feringa, B. L. Hypothesis-Driven, Structure-Based Design in Photopharmacology: The Case of eDHFR Inhibitors. *J. Med. Chem.* **2022**, *65* (6), 4798−4817.
(7) Khavani, M.; Bakhtiiari, A.; Khalvati, L.; Leurs, R.; Liang, R. A Systematic Benchmark Study of Free Energy Methods for Quantifying Light-Responsive Binding Affinities of Photoswitchable Antagonists of Beta-Adrenergic Receptors. *J. Med. Chem.* **2026**, *69*, 9265.
(8) Bakhtiiari, A.; Khavani, M.; Costa, G. J.; Liang, R. A Multiscale Simulation Framework for Elucidating Photochemical Structure−Activity Relationships of Photoswitchable Ligands in Complex Biomolecular Environments. *J. Chem. Inf. Model.* **2025**, *65* (23), 12861−12878.
(9) Zhang, Z.; Le, G. N. T.; Ge, Y.; Tang, X.; Chen, X.; Ejim, L.; Bordeleau, E.; Wright, G. D.; Burns, D. C.; Tran, S.; Axerio-Cilies, P.; Wang, Y. T.; Dong, M.; Woolley, G. A. Isomerization of bioactive acylhydrazones triggered by light or thiols. *Nat. Chem.* **2023**, *15* (9), 1285−1295.
(10) (a) Lin, Y.-L.; Aleksandrov, A.; Simonson, T.; Roux, B. An overview of electrostatic free energy computations for solutions and proteins. *J. Chem. Theory Comput.* **2014**, *10* (7), 2690−2709. (b) Papaneophytou, C. P.; Grigoroudis, A. I.; McInnes, C.; Kontopidis, G. Quantification of the effects of ionic strength, viscosity, and hydrophobicity on protein−ligand binding affinity. *ACS Med. Chem. Lett.* **2014**, *5* (8), 931−936.
(11) Koźmiński, P.; Halik, P. K.; Chesori, R.; Gniazdowska, E. Overview of dual-acting drug methotrexate in different neurological diseases, autoimmune pathologies and cancers. *Int. J. Mol. Sci.* **2020**, *21* (10), 3483.
(12) Schnell, J. R.; Dyson, H. J.; Wright, P. E. Structure, dynamics, and catalytic function of dihydrofolate reductase. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33* (1), 119−140.
(13) (a) Crellin, E.; Mansfield, K. E.; Leyrat, C.; Nitsch, D.; Douglas, I. J.; Root, A.; Williamson, E.; Smeeth, L.; Tomlinson, L. A. Trimethoprim use for urinary tract infection and risk of adverse outcomes in older patients: cohort study. *BMJ.* **2018**, *360*, 1. (b) Schneider, P.; Hawser, S.; Islam, K. Iclaprim, a novel diaminopyrimidine with potent activity on trimethoprim sensitive and resistant bacteria. *Bioorg. Med. Chem. Lett.* **2003**, *13* (23), 4217−4221.
(14) (a) Wegener, M.; Hansen, M. J.; Driessen, A. J.; Szymanski, W.; Feringa, B. L. Photocontrol of antibacterial activity: shifting from UV to red light activation. *J. Am. Chem. Soc.* **2017**, *139* (49), 17979−17986. (b) Mashita, T.; Kowada, T.; Takahashi, H.; Matsui, T.; Mizukami, S. Light-Wavelength-Based Quantitative Control of Dihydrofolate Reductase Activity by Using a Photochromic Isostere of an Inhibitor. *ChemBioChem* **2019**, *20* (11), 1382−1386.
(15) Cocco, L.; Roth, B.; Temple, C.; Montgomery, J. A.; London, R. E.; Blakley, R. L. Protonated state of methotrexate, trimethoprim, and pyrimethamine bound to dihydrofolate reductase. *Arch. Biochem. Biophys.* **1983**, *226* (2), 567−577.
(16) Matthews, D. A.; Bolin, J. T.; Burridge, J. M.; Filman, D. J.; Volz, K. W.; Kraut, J. Dihydrofolate reductase. The stereochemistry of inhibitor selectivity. *J. Biol. Chem.* **1985**, *260* (1), 392−399.
(17) (a) Christ, C. D.; Fox, T. Accuracy assessment and automation of free energy calculations for drug design. *J. Chem. Inf. Model.* **2014**, *54* (1), 108−120. (b) Homeyer, N.; Stoll, F.; Hillisch, A.; Gohlke, H. Binding free energy calculations for lead optimization: assessment of their accuracy in an industrial drug design context. *J. Chem. Theory Comput.* **2014**, *10* (8), 3331−3344.
(18) Darve, E. Thermodynamic Integration Using Constrained and Unconstrained Dynamics. In *Free Energy Calculations: Theory and Applications in Chemistry and Biology*; Chipot, C., Pohorille, A., Eds.; Springer Berlin Heidelberg, 2007; pp 119−170.
(19) (a) Miller III, B. R.; McGee Jr, T. D.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E. MMPBSA. py: an efficient program for end-state free energy calculations. *J. Chem. Theory Comput.* **2012**, *8* (9), 3314−3321. (b) Lu, Q.; Luo, R. A Poisson−Boltzmann dynamics method with nonperiodic boundary condition. *J. Chem. Phys.* **2003**, *119* (21), 11035−11047.
(20) (a) Wang, E.; Sun, H.; Wang, J.; Wang, Z.; Liu, H.; Zhang, J. Z.; Hou, T. End-point binding free energy calculation with MM/PBSA and MM/GBSA: strategies and applications in drug design. *Chem. Rev.* **2019**, *119* (16), 9478−9508. (b) Srivastava, H. K.; Sastry, G. N. Molecular dynamics investigation on a series of HIV protease inhibitors: assessing the performance of MM-PBSA and MM-GBSA approaches. *J. Chem. Inf. Model.* **2012**, *52* (11), 3088−3098.
(21) Khavani, M.; Bakhtiiari, A.; Khalvati, L.; Leurs, R.; Liang, R. A Systematic Benchmark Study of Free Energy Methods for Quantifying Light-Responsive Binding Affinities of Photoswitchable Antagonists of Beta-Adrenergic Receptors. *J. Med. Chem.* **2026**, *69* (8), 9265−9283.
(22) Bennett, B. C.; Wan, Q.; Ahmad, M. F.; Langan, P.; Dealwis, C. G. X-ray structure of the ternary MTX· NADPH complex of the anthrax dihydrofolate reductase: A pharmacophore for dual-site inhibitor design. *J. Struct. Biol.* **2009**, *166* (2), 162−171.
(23) 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* (2), 455−461.
(24) Anandakrishnan, R.; Aguilar, B.; Onufriev, A. V. H++ 3.0: automating p K prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Res.* **2012**, *40* (W1), W537−W541.
(25) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79* (2), 926−935.
(26) 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. Theory Comput.* **2015**, *11* (8), 3696−3713.
(27) Cummins, P. L.; Ramnarayan, K.; Singh, U.; Gready, J. E. Molecular dynamics/free energy perturbation study on the relative affinities of the binding of reduced and oxidized NADP to dihydrofolate reductase. *J. Am. Chem. Soc.* **1991**, *113* (22), 8247−8256.
(28) (a) 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* (9), 1157−1174. (b) 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* (2), 247−260.
(29) (a) Liang, R.; Bakhtiiari, A. Multiscale simulation unravels the light-regulated reversible inhibition of dihydrofolate reductase by phototrexate. *J. Chem. Phys.* **2022**, *156* (24), 245102. (b) Bakhtiiari, A.; Liang, R. Unraveling Solvent and Substituent Effects in the Photodynamics of Light-Dependent Microtubule Inhibitors for Cancer Phototherapy. *J. Comput. Chem.* **2025**, *46* (7), No. e70076.
(30) Bakhtiiari, A.; Costa, G. J.; Liang, R. On the Simulation of Thermal Isomerization of Molecular Photoswitches in Biological Systems. *J. Chem. Theory Comput.* **2023**, *19* (18), 6484−6499.
(31) 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* (2), 024110.
(32) 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* (9), 5499−5511.
(33) (a) Liang, R. First-Principles Nonadiabatic Dynamics Simulation of Azobenzene Photodynamics in Solutions. *J. Chem. Theory Comput.* **2021**, *17* (5), 3019−3030. (b) Liang, R.; Das, D.; Bakhtiiari, A. Protein confinement fine-tunes aggregation-induced emission in human serum albumin. *Phys. Chem. Chem. Phys.* **2021**, *23* (46), 26263−26272. (c) Liang, R.; Bakhtiiari, A. Effects of Enzyme−Ligand Interactions on the Photoisomerization of a Light-Regulated

Chemotherapeutic Drug. *J. Phys. Chem. B* **2022**, *126* (12), 2382−2393. (d) 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 nπ* and ππ* Excited States. *J. Am. Chem. Soc.* **2020**, *142* (49), 20680−20690.
(34) Seritan, S.; Bannwarth, C.; Fales, B. S.; Hohenstein, E. G.; Kokkila-Schumacher, S. I.; 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* (22), 224110.
(35) Seritan, S.; Bannwarth, C.; Fales, B. S.; Hohenstein, E. G.; Isborn, C. M.; Kokkila-Schumacher, S. I.; Li, X.; Liu, F.; Luehr, N.; Snyder, J. W., Jr; Song, C.; et al. A graphical processing unit-accelerated electronic structure package for large-scale ab initio molecular dynamics. *WIREs Comput. Mol. Sci.* **2021**, *11* (2), No. e1494.
(36) Sindhikara, D. J.; Kim, S.; Voter, A. F.; Roitberg, A. E. Bad seeds sprout perilous dynamics: Stochastic thermostat induced trajectory synchronization in biomolecules. *J. Chem. Theory Comput.* **2009**, *5* (6), 1624−1631.
(37) Åqvist, J.; Wennerström, P.; Nervall, M.; Bjelic, S.; Brandsdal, B. O. Molecular dynamics simulations of water and biomolecules with a Monte Carlo constant pressure algorithm. *Chem. Phys. Lett.* **2004**, *384* (4−6), 288−294.
(38) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N log (N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089.
(39) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23* (3), 327−341.
(40) Case, D. A.; Aktulga, H. M.; Belfon, K.; Ben-Shalom, I. Y.; Berryman, J. T.; Brozell, S. R.; Cerutti, D. S.; Cheatham III, T. E.; Cisneros, G. A.; Cruzeiro, V. W. D. *Amber 2023*; University of California: San Francisco, 2023.
(41) Swails, J. M.; York, D. M.; Roitberg, A. E. Constant pH Replica Exchange Molecular Dynamics in Explicit Solvent Using Discrete Protonation States: Implementation, Testing, and Validation. *J. Chem. Theory Comput.* **2014**, *10* (3), 1341−1352.
(42) Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **2008**, *120* (1), 215−241.
(43) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. *J. Comput. Chem.* **2003**, *24* (6), 669−681.
(44) (a) Ho, J.; Coote, M. L. A universal approach for continuum solvent pKa calculations: are we there yet? *Theor. Chem. Acc.* **2010**, *125* (1), 3−21. (b) Mlotkowski, A. J.; Schlegel, H. B.; Chow, C. S. Calculated pKa Values for a Series of Aza- and Deaza-Modified Nucleobases. *J. Phys. Chem. A* **2023**, *127* (15), 3526−3534.
(45) Torrie, G. M.; Valleau, J. P. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.* **1977**, *23* (2), 187−199.
(46) (a) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A:At., Mol., Opt. Phys.* **1988**, *38* (6), 3098−3100. (b) 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* (2), 785−789. (c) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98* (7), 5648−5652. (d) Becke, A. D. A new mixing of Hartree−Fock and local density-functional theories. *J. Chem. Phys.* **1993**, *98* (2), 1372−1377.
(47) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132* (15), 154104.
(48) (a) Costa, G. J.; Liang, R. Understanding the multifaceted mechanism of Compound I formation in Unspecific Peroxygenases through multiscale simulations. *J. Phys. Chem. B* **2023**, *127* (41), 8809−8824. (b) Costa, G. J.; Egbemhenghe, A.; Liang, R. Computational characterization of the reactivity of Compound I in unspecific peroxygenases. *J. Phys. Chem. B* **2023**, *127* (51), 10987−10999. (c) Shaik, S.; Cohen, S.; Wang, Y.; Chen, H.; Kumar, D.; Thiel, W. P450 Enzymes: Their Structure, Reactivity, and Selectivity—Modeled by QM/MM Calculations. *Chem. Rev.* **2010**, *110* (2), 949−1017. (d) Shaik, S.; Kumar, D.; de Visser, S. P.; Altun, A.; Thiel, W. Theoretical perspective on the structure and mechanism of cytochrome P450 enzymes. *Chem. Rev.* **2005**, *105* (6), 2279−2328.
(49) Altun, A.; Breidung, J. r.; Neese, F.; Thiel, W. Correlated ab initio and density functional studies on H2 activation by FeO+. *J. Chem. Theory Comput.* **2014**, *10* (9), 3807−3820.
(50) Jankowska, J.; Sobolewski, A. L. Modern theoretical approaches to modeling the excited-state intramolecular proton transfer: an overview. *Molecules* **2021**, *26* (17), 5140.
(51) Radoń, M.; Broclawik, E. Peculiarities of the Electronic Structure of Cytochrome P450 Compound I: CASPT2 and DFT Modeling. *J. Chem. Theory Comput.* **2007**, *3* (3), 728−734.
(52) Liang, R.; Liu, F.; Martínez, T. J. Nonadiabatic Photodynamics of Retinal Protonated Schiff Base in Channelrhodopsin 2. *J. Phys. Chem. Lett.* **2019**, *10* (11), 2862−2868.
(53) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.* **1992**, *13* (8), 1011−1021.
(54) 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* (1), 461−469.
(55) Bonomi, M.; Branduardi, D.; Bussi, G.; Camilloni, C.; Provasi, D.; Raiteri, P.; Donadio, D.; Marinelli, F.; Pietrucci, F.; Broglia, R. A.; et al. PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Comput. Phys. Commun.* **2009**, *180* (10), 1961−1972.
(56) Gao, Y.; Zhu, T.; Chen, J. Exploring drug-resistant mechanisms of I84V mutation in HIV-1 protease toward different inhibitors by thermodynamics integration and solvated interaction energy method. *Chem. Phys. Lett.* **2018**, *706*, 400−408.
(57) Chen, J.; Wang, X.; Zhu, T.; Zhang, Q.; Zhang, J. Z. H. A Comparative Insight into Amprenavir Resistance of Mutations V32I, G48V, I50V, I54V, and I84V in HIV-1 Protease Based on Thermodynamic Integration and MM-PBSA Methods. *J. Chem. Inf. Model.* **2015**, *55* (9), 1903−1913.
(58) Han, Y.; Dawson, J. R.; DeMarco, K. R.; Rouen, K. C.; Bekker, S.; Yarov-Yarovoy, V.; Clancy, C. E.; Xiang, Y. K.; Vorobyov, I. Elucidation of a dynamic interplay between a beta-2 adrenergic receptor, its agonist, and stimulatory G protein. *Proc. Natl. Acad. Sci. U. S. A.* **2023**, *120* (10), No. e2215916120.
(59) Roth, B.; Strelitz, J. Z. Protonation of 2, 4-diaminopyrimidines. I. Dissociation constants and substituent effects. *J. Org. Chem.* **1969**, *34* (4), 821−836.