JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

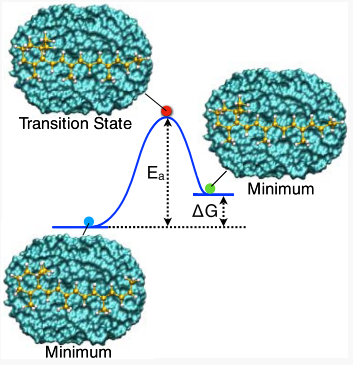
Free Energy Computation for an Isomerizing Chromophore in a Molecular Cavity via the Average Solvent Electrostatic Configuration Model: Applications in Rhodopsin and Rhodopsin-Mimicking Systems

Dmitrii M. Nikolaev, Madushanka Manathunga, Yoelvis Orozco-Gonzalez,* Andrey A. Shtyrov, Yansel Omar Guerrero Martínez, Samer Gozem, Mikhail N. Ryazantsev,* Kaline Coutinho,* Sylvio Canuto, and Massimo Olivucci*

Cite This: *J. Chem. Theory Comput.* 2021, 17, 5885−5895 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: We present a novel technique for computing the free energy differences between two chromophore "isomers" hosted in a molecular environment (a generalized solvent). Such an environment may range from a relatively rigid protein cavity to a flexible solvent environment. The technique is characterized by the application of the previously reported "average electrostatic solvent configuration" method, and it is based on the idea of using the free energy perturbation theory along with a chromophore annihilation procedure in thermodynamic cycle calculations. The method is benchmarked by computing the ground-state room-temperature relative stabilities between (i) the *cis* and *trans* isomers of prototypal animal and microbial rhodopsins and (ii) the analogue isomers of a rhodopsin-like light-driven molecular switch in methanol. Furthermore, we show that the same technology can be used to estimate the activation free energy for the thermal isomerization of systems i−ii by replacing one isomer with a transition state. The results show that the computed relative stability and isomerization barrier magnitudes for the selected systems are in line with the available experimental observation in spite of their widely diverse complexity.

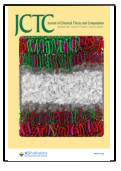


1. INTRODUCTION

Rhodopsins are a remarkable family of photoreceptor proteins found in all life domains.[1,2] Their functions in biological systems vary from vision in superior animals to ion transportation, phototaxis, and gene expression in archaea and eubacteria.[2] All members of the rhodopsin family feature a retinal protonated Schiff base (rPSB) chromophore, embedded inside a protein (opsin) cavity. The rhodopsin function is initiated through light-induced double-bond isomerization of rPSB. However, it has been recognized that the same isomerization process may be also initiated thermally and play an important role in the photoreceptor function. Indeed, understanding the thermal isomerization of such systems is of fundamental importance to investigate the photoreceptor background (dark) noise, in the engineering of photochromic (i.e., bistable) rhodopsins[3] as well as in the design of working biomimetic molecular switches and motors.[4−6]

The most accurate way to predict thermodynamic but also kinetic properties—such as the relative stabilities of reactants and products and the barrier for their reaction—in the type of systems mentioned above, requires the calculation of free energy changes rather than changes in potential energy. These calculations have to take into account that biological and biomimetic chromophores contain extended conjugated frameworks whose correct description requires electronic wave functions flexible enough to describe, in a balanced way, the closed-shell as well as open-shell regions of their potential energy surface. This is particularly important when investigating their double-bond isomerization. We recently reported a protocol to construct hybrid quantum mechanics/molecular mechanics (QM/MM) models of rhodopsin proteins and rhodopsin-like chromophores in solution that incorporates a level of QM theory suitable for dealing with large wave function character changes such as those that may occur during reactive and photoexcitation processes.[7−13] We also showed that such models can be employed to perform QM/MM geometry optimizations on an approximate free energy surface and can predict the absorption maximum wavelength (λ_{max}).[14] Such a protocol, called ASEC-FEG, was constructed based on the idea of combining the average solvent electrostatic

Received: March 2, 2021
Published: August 11, 2021



configuration (ASEC) model[15,16] and the free energy gradient (FEG) method.[17−19] However, no effort was made to exploit/adapt the protocol for the correct computation of free energy differences.

The ASEC-FEG protocol is based on the assumption that a protein provides an electrostatic and steric environment, much like a solvent. Therefore, a protein may be regarded as an "organized" solvent. Therefore, the generated ASEC model was adapted for rhodopsins, such that an ensemble of configurations represents a time-averaged interaction of the rPSB (QM subsystem) with the opsin environment (MM subsystem) at thermodynamic equilibrium. Then, the chromophore is optimized within the average protein environment. Methodologically, the treatment of solvents and proteins is different in a number of ways (see Sections 2 and 3). However, the basic theory that the solvent QM/MM and protein QM/MM models are built on is the same. Therefore, to avoid redundancy, in this work, we will refer to both an actual solvent environment and the protein environment as "solvent". The methodology will also be termed "ASEC" for both solvent and protein environments, despite some differences in the details of the methodology used in these two types of molecular systems.

As mentioned above, although ASEC-FEG was able to predict λ_{max} values of rhodopsins in good agreement with the experiment (within a limited absolute error), it was not able to correctly predict the relative stability between structurally distinct stationary structures (i.e., reactants, transition states, and products).[14] To find a solution to such limitations, in this work, we present an updated ASEC-FEG protocol for computing free energy differences between rhodopsin "isomers", namely, featuring isomers of the chromophore that can be interconverted via thermal double-bond isomerization. Thus, here, we are not dealing with the system photochemical reactivity. We are instead dealing with the chromophore thermal (equilibrium) isomerization exclusively.[20,21]

There are a few examples of theories that integrate QM/MM calculations into statistical mechanics tools to compute free energy differences. The details of the theory and implementation vary, however, among the different developed approaches. Hayashi et al.[22] developed the QM/MM reweighting free energy self-consistent field (QM/MM-RWFE-SCF) approach, which performed quantum mechanical calculations and optimizations on a free energy surface constructed by conformational sampling. They then applied this approach to optimize the reactant and transition state in protein systems and estimate the activation free energy.[23] More specifically, they use interpolation of coordinates between the reactant and transition state of the Ras-GAP complex and employ the Bennett acceptance ratio to evaluate free energy differences between two subsequent structures.[24]

A similar methodology, known as average solvent electrostatic potential/molecular dynamic (ASEP/MD), was previously proposed by Galvań et al.[25] and applied to the free energy optimization of a QM solute in an MM solvent environment. In both the Hayashi et al. and Galvań et al. methods, the QM subsystem is optimized in an average electrostatic potential created by running a molecular dynamics (MD) sampling of the solute MM environment. While this may appear to be similar to what happens in ASEC, in this last case, sampling focuses on the environment MM configurations rather than potential.

A number of hybrid QM/MM free energy methods have also been developed employing an interface between Q-Chem and CHARMM,[26] including a multienvironment single-system (MESS) QM/MM approach that samples the solvent environment.[27] These methods have largely been used to compute solvation free energies,[28,29] although they have also been used to compute reaction free energy profiles of a solute in a solvent and in a protein environment as well.[30−32]

The updated ASEC-FEG protocol described below leverages an adaptation of the thermodynamic cycle-based approach previously used to compute the free energy differences of a solute molecule in solution.[33,34] In such an approach, the free energy perturbation (FEP) theory is combined with the hypothetical annihilation of the solute molecule.[35] The same approach has been carefully benchmarked for tautomerization reactions in solution.[34] An alternative methodology has been also developed where the free energy profile between two isomers is computed by perturbing one structure to the nearest structure along a reaction path. The main deficiency of the latter approach is associated with the accuracy of the reaction path, which is normally computed in the gas phase[34] or obtained by linearly interpolating between equilibrium structures.[4] Although the above methodologies have been applied to study molecular systems in solution, to the best of our knowledge, there is no reported study on the application of the free energy perturbation method to compute isomer relative stability and thermal isomerization barriers of rhodopsins and rhodopsin-like systems. However, such an investigation has become important, especially, in the wake of experiments reporting an unusual kinetic behavior of the thermally induced isomerization of visual rhodopsin.[36]

In conclusion, below, we show how a thermodynamic cycle method has been incorporated into the ASEC-FEG protocol[14] to compute free energy differences. Such a method is benchmarked using representative animal and microbial rhodopsins. For the former, we compute the relative stability between visual bovine rhodopsin (Rh) and its isolable primary photocycle isomer bathorhodopsin (bathoRh). We also compute the activation free energy of the charge-transfer (CT) transition state that has previously been reported to be associated with Rh → bathoRh thermal isomerization.[20] For the latter (i.e., for microbial rhodopsins), we focus on photochromic microbial sensory rhodopsin found in the freshwater cyanobacterium *Anabaena* (Nostoc) PCC7120 (*Anabaena* sensory rhodopsin, ASR).[37,38] We compute the relative stability between the 13-*cis* and all-*trans* isomers of ASR (hereafter ASR-13C and ASR-AT) and free energy barrier of the CT transition state controlling the experimentally observed ASR-13C → ASR-AT isomerization.[38] We also benchmark the updated ASEC-FEG protocol using the third system, the rhodopsin-like (i.e., biomimetic) *N*-alkyl-indanylidene-pyrrolinium (NAIP) switch that has been previously studied with the closely related but distinct ASEP method,[4] mentioned above. In this case, we compute the free energy difference between the *cis* (*Z*) and *trans* (*E*) isomers of the switch (from now on *Z*-NAIP and *E*-NAIP) and also predict the free energy barrier of the lower-energy CT transition state.

The next sections of this article are organized as follows. In Section 2, we briefly revise the theoretical details of ASEC-FEG protocol and introduce the methodology used to compute free energy differences. Section 3 is devoted to a description of the computational method and the benchmarking procedure. The results of the benchmarking procedure are

https://doi.org/10.1021/acs.jctc.1c00221
J. Chem. Theory Comput. 2021, 17, 5885−5895

presented and discussed in Section 4. Conclusions are drawn in Section 5.

2. THEORETICAL DETAILS

2.1. Free Energy Geometry Optimization. We benchmark the new ASEC-FEG capability by focusing on two types of systems: rhodopsin proteins and the rhodopsin chromophore-like NAIP in methanol solution. Since the first step of our methodology is the geometry optimization on the free energy surface, we now briefly discuss the structure of the models built for both types.

In rhodopsins, rPSB is bound to a lysine residue of the opsin cavity. In our QM/MM models, rPSB and few atoms of the lysine residue (see the shaded region in Figure 1) are treated at

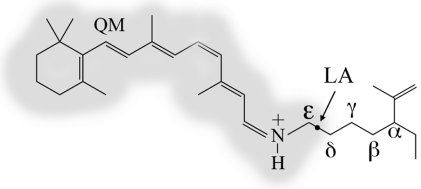


Figure 1. QM/MM model of rhodopsin-like systems used in the ASEC-FEG protocol. The shaded region represents the QM subsystem, whereas the rest belongs to the MM subsystem.

the QM level. The QM subsystem (hereafter QM-Lys) also comprises a hydrogen link atom (LA) bonded to Cε and directed toward Cδ. The position of the link atom is defined according to the Morokuma approach. The MM subsystem comprises Cδ, the rest of the lysine atoms, and the opsin protein. This QM/MM model is then optimized following the ASEC-FEG protocol, where the side chains of all atoms in the protein binding pocket, except for the rPSB chromophore and the connected lysine, are represented by the ASEC configuration (see below). The QM-Lys subsystem is then optimized in the ASEC environment. For a detailed theoretical description, the reader is directed to our previous paper.[14]

For the geometry optimization of NAIP switches in methanol, we have adapted the ASEC-FEG protocol to optimize a single molecule of the switch (i.e., the QM subsystem, Figure 2) in the solvent cavity and remaining bulk (i.e., the MM subsystem).

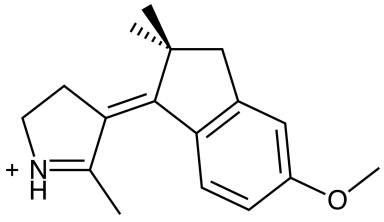


Figure 2. Structure of NAIP (QM subsystem) that was optimized in the methanol solution (MM subsystem) with the ASEC-FEG protocol.

In ASEC-FEG, the average force acting on each atom of the QM subsystem is calculated using the free energy gradient proposed by Nagaoka et al.,[17−19] which is obtained through the following relationship

$$F(q) = -\frac{\partial G(q)}{\partial q} = -\left\langle \frac{\partial V}{\partial q} \right\rangle \approx -\frac{\partial \langle V \rangle}{\partial q} \quad (1)$$

In this equation, q represents the nuclear coordinates of the QM subsystem, G is the free energy of the system, and V is the potential energy of the QM subsystem plus the interaction energy with the environment. These forces are equal to the time-averaged forces acting on each atom of the QM subsystem over the equilibrium distribution of the total system, as obtained from a molecular dynamics (MD) simulation.

The potential energy of the system in eq 1 is decomposed in the following way

$$V = V_{QM} + V_{QM/MM} \quad (2)$$

The term V_{QM} represents the potential energy of the QM part computed using quantum chemical methods, while $V_{QM/MM}$ comprises two components

$$V_{QM/MM} = V_{ele}(QM/MM) + V_{vdw}(QM/MM) \quad (3)$$

In this equation, $V_{ele}(QM/MM)$ and $V_{vdw}(QM/MM)$ are the electrostatic and van der Waals interaction energies, respectively, between the QM subsystem and the MM atoms of the solvent.

An issue in computing the average total energy from eq 1 is to find an efficient way to calculate the average interaction energy between the atoms of the QM subsystem and the MM atoms of the solvent. For this purpose, we use the ASEC model.[39−41] Accordingly, we sample 100 uncorrelated configurations of a selected part of the MM subsystem (i.e., the side chains of protein cavity residues, see Section 3) for rhodopsins and the 20 Å solvent shell (the solvent is treated with periodic boundary conditions) for the molecular switch. These configurations are taken as snapshots from an MD simulation. This means that in the MM subsystem we have a fixed part (far from the chromophore) and a flexible part (the one surrounding the chromophore). The "ASEC configuration" is then generated by the superposition of all of the sampled snapshots, in which case each atom of the solvent shell is replaced by 100 pseudoatoms in different positions with scaled charges and van der Waals parameters. This methodology is very similar to the one developed by Georg et al.[15,16] to optimize molecular systems in solution, albeit with two important differences: (i) we are incorporating the classical van der Waals interactions into the ASEC configuration, where van der Waals parameters are assigned both to the atoms of the QM subsystem and to the MM pseudoatoms of the MM subsystem. Details on the scaling of the van der Waals parameters can be found in the Supporting Information of ref 14. It can be observed that the interaction energy (electrostatic and van der Waals) between the atoms of the QM subsystem and the atoms of the solvent, calculated using the ASEC configuration, is the same as the time-averaged interaction energy taken over all of the individual sampled configurations. (ii) The electrostatic interaction energy between the solute and the MM pseudoatoms of the solvent is considered through the electrostatic potential fitting (ESPF) method, which computes the QM/MM electrostatic interaction in a uniquely defined way, see details in refs 42, 43.

Following this procedure, one could compute the gradient of the free energy from eq 1 and locate the stationary points (i.e., minima and transition states) on the free energy surface. This

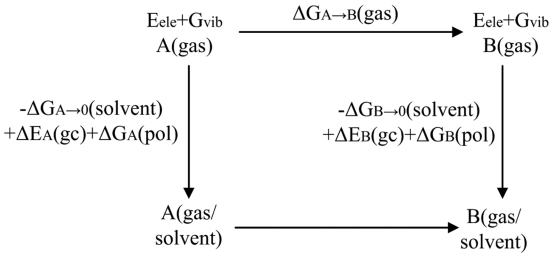
https://doi.org/10.1021/acs.jctc.1c00221
J. Chem. Theory Comput. 2021, 17, 5885−5895

procedure has been implemented iteratively. Extensive molecular dynamics is performed to sample the MM environment and generate the ASEC configuration, and then, a full geometry optimization of the QM subsystem is computed within the average external field generated by the ASEC configuration. At this point, the obtained QM/ASEC geometry is used to recompute the charges of the QM subsystem in the ASEC environment, and the new geometry and charges are used for a subsequent MD simulation to adapt the solvent to the new charges and geometry. A new ASEC environment is generated, a new geometry optimization is performed, and this iterative procedure is repeated until the total average energy of the system is converged to a defined threshold of about 0.5 kcal/mol. In addition to the average energy convergence, the convergence of the dipole moment of the QM subsystem is also required to properly polarize the QM subsystem due to the interaction with the solvent.[34,44,45]

The statistical contribution of the QM subsystem to the total free energy of the system is not considered during the geometry optimization due to computational resources required but can be taken into account by vibrational frequency calculations (see Section 2.2).

2.2. Computing Free Energy Differences. Once the stationary points have been located with ASEC-FEG (see Figure 3 in ref 14 and the related text for a schematic representation of the ASEC-FEG workflow), the free energy difference between two stationary points (e.g., corresponding to two isomers) and indicated as A(gas/solvent) → B(gas/solvent) can be computed using the thermodynamic cycle,[33,34] as shown in Scheme 1. In this diagram, A and B represent the

Scheme 1. Thermodynamic Cycle for Computing the Free Energy Difference between Two Chromophore Configurations A and B



QM subsystem of the two stationary points (e.g., two rhodopsin isomers or two isomers of a molecular switch in solution). A(gas) and A(gas/solvent) correspond to the different equilibrium geometries in isolated conditions and in the solvent environment, respectively. $\Delta G_{A->B}(gas)$ can be computed from the potential energy (electronic energy plus nuclear repulsion energy) plus thermal corrections from vibrational frequency calculations at the CASSCF/6-31G* level of theory. The vertical arrows in the cycle include three terms. $\Delta E_A(gc)$ and $\Delta E_B(gc)$ stands for the variation of the gas-phase QM subsystem potential energy going from the gas-phase geometry to the solvated equilibrium geometry. $\Delta G_A(pol)$ and $\Delta G_B(pol)$ correspond to the variation of the free energy caused by the electronic polarization of the QM subsystem when interacting with the environment.

The remaining term $\Delta G_{A\to 0}(solvent)$ and $\Delta G_{B\to 0}(solvent)$ can be thought of as the variation of the free energy required by the QM subsystem to disappear inside the solvent. Therefore, the negative of this term would be the energy required for the chromophore to appear inside the environment. This term is computed through a hypothetical procedure that involves annihilation of the QM subsystem. More specifically, in the annihilation (indicated by A → 0 and B → 0), the parameters defining the nonbonded interactions between the chromophore and the solvent are gradually zeroed. In this case, the variation of the free energy for the solute to vanish in the corresponding environment (i.e., the terms $\Delta G_{A\to 0}(solvent)$ and $\Delta G_{B\to 0}(solvent)$) is computed through the fundamental equation of the free energy perturbation theory,[46] expressed in terms of the variations of the enthalpy of the system, H, as shown in the following equation[47,48]

$$\Delta G_{i\to j}(solvent) = -k_B T \ln\langle \exp[-(H_j - H_i)/k_B T]\rangle_i \quad (4)$$

In this expression, H represents the enthalpy of the system, k_B and T are the Boltzmann constant and temperature, respectively, and i and j represent two different, but close, configurations of the chromophore embedded in the environment. In this case, the term "configurations" stands for the intermediate steps of the chromophore annihilation procedure going from A to 0 and from B to 0. Note that large values of ΔH, due to very different configurations, would lead to unreliable free energy difference calculation. The angle brackets $\langle\ \rangle_i$ in eq 4 indicate that the average of the ΔH in the exponential is computed in the ensemble of configurations generated for i. Accordingly, only the interactions between the chromophore and environment vary.

The desired free energy difference A(gas/solvent) → B(gas/solvent) can then be computed following the thermodynamic cycle represented in Scheme 1. According to this cycle, the free energy difference can be calculated using eq 5.

$$\begin{aligned}\Delta G_{A\to B}(gas/solvent) &= \Delta G_{A\to B}(gas) - \Delta G_{B\to 0}(solvent) + \Delta E_B(gc) \\ &\quad + \Delta G_B(pol) + \Delta G_{A\to 0}(solvent) - \Delta E_A(gc) \\ &\quad - \Delta G_A(pol)\end{aligned} \quad (5)$$

Equation 4 is used to compute the values of $\Delta G_{A\to 0}(solvent)$ and $\Delta G_{B\to 0}(solvent)$.

For rhodopsins, a series of MD simulations are performed to gradually annihilate the QM part in each structure (A and B). This process was divided into three stages. In the first stage, the electrostatic interaction is turned off using eight simulations, where the atomic charges are multiplied by the scaling factor μ varying from 1 to 0 with intervals of 0.125. In the second stage, the attractive part of the Lennard−Jones (LJ) interaction is almost turned off using eight simulations with the ε parameter multiplied by μ varying from 1 to 0.01 with the same intervals. Finally, in the third stage, the repulsive part of the LJ interaction is turned off using four simulations and choosing the σ parameter multiplied by μ varying from 1 to 0 with intervals of 0.25. Therefore, a total of 20 MD simulations were performed to complete the annihilation. The energy variations were found to be always smaller than 5 kcal/mol, thus allowing a smooth convergence of the free energy in each simulation interval, $\Delta G(\mu_i \to \mu_{i+1})$. For the NAIP switches, we performed an equivalent annihilation procedure but using MC simulations that allowed the double-wide sampling for the atomic charge and vanishing ε parameter. In this case, because

https://doi.org/10.1021/acs.jctc.1c00221
J. Chem. Theory Comput. 2021, 17, 5885−5895

in a single simulation two free energy variations are calculated (i.e., $\Delta G(\mu_i \rightarrow \mu_{i-1})$ and $\Delta G(\mu_i \rightarrow \mu_{i+1})$), only four simulations were needed for the first and second stages. In this way, a total of 12 MC simulations were performed to completely annihilate the NAIP switches when considering the same scaling factor μ_i.

For the rhodopsin-type proteins, the above annihilation procedure has been implemented through a collection of Bash and Python scripts and incorporated into the ASEC-FEG protocol. Such an update makes the protocol capable of computing the free energy difference between rhodopsin isomers automatically. The input for the procedure would be stationary structures (in .pdb format, see also below), and the corresponding charges of the retinal chromophore would be computed, again, by the ASEC-FEG protocol. As for the solvated NAIPs, the well-established methodology implemented in the DICE code[49] is used to compute the free energy difference of the *E* and *Z* molecular switches in solution.[49] We note that although DICE is used to compute the free energy difference, the switch stationary structures used in the calculations are the ones optimized by the ASEC-FEG methodology (see Section 3.2).

3. COMPUTATIONAL DETAILS

3.1. Rhodopsin Proteins. The free energy geometry optimizations of Rh and bathoRh isomers and of the ASR-13C and ASR-AT isomers were performed using the ASEC-FEG protocol, consistent with the benchmarking calculations reported in ref 14. More specifically, the X-ray crystallographic structures of Rh, bathoRh, and ASR isomers (PDB ID: 1U19, 2G87, and 1XIO, respectively) were obtained from the Protein Data Bank and the cavities were selected using the CASTp online server.[50] The crystal structures and cavity residue files were then used as the input to the ASEC-FEG protocol. As for the geometry optimizations, the QM-Lys subsystem was treated at 3-root state average (with equal state weights) CASSCF/6-31G* level of theory[51] with an active space comprising of the full π system (12 electrons in 12 orbitals) of the rPSB chromophore. The MM subsystem was described using the AMBER force field.[52] The resulting optimized Rh and ASR-13C structures were used to construct the input for locating the transition states (from now on Rh-TS and ASR-TS, respectively). Transition-state structures were optimized starting from optimized guess structures corresponding to regular (i.e., featuring a single selected MM subsystem representing the chromophore environment) QM/MM models already reported in the literature (see, for instance, the structures from ref 19 for Rh). When such guess structures are not available, we run a constrained geometry optimization constraining the isomerizing double bond to a 90° dihedral angle and starting from geometrical parameters ensuring an initial CT character (i.e., selecting a suitable initial bond-length alternation). Starting from such a guess, we employ the findTS algorithm implemented in MOLCAS to optimize the transition state. The obtained guess structure is used to carry out the ASEC-FEG computation. In all cases, the cavities and the level of theory for the transition state optimization were the same used for energy minimizations. The necessary guess Hessian is from the guess structure. The cavities and the level of theory for the transition state optimizations were the same as used for the minimizations.

The optimized Rh, bathoRh, ASR-13C, ASR-AT, Rh-TS, and ASR-TS structures were provided as inputs for the annihilation procedure described in Section 2.2. The values of the scaling factor (μ) were selected as follows: for chromophore charges, μ = 1.0, 0.875, 0.75, 0.625, 0.5, 0.375, 0.25, 0.125, 0.0; for well depth (ε) of the Lennard−Jones potential (which approximates van der Waals interactions), μ = 1.0, 0.875, 0.75, 0.625, 0.5, 0.375, 0.25, 0.125, 0.01; and for the van der Waals radii, μ = 1.0, 0.75, 0.5, 0.25, 0.0. The rPSBs from each structure were isolated in the gas phase and optimized at the CASSCF/6-31G* level of theory with the same active space selected for QM/MM calculations. Frequency and vibrational free energy calculations were performed for all optimized structures at the same level of theory. To account for dynamical electron correlation, the energies of the initial (the geometry of the chromophore as optimized inside the protein) and optimized structures were recomputed at the CASPT2//CASSCF/6-31G* level.[53] In line with previously reported benchmark studies on a minimal rhodopsin chromophore model, no IPEA shift was employed in the calculations. In fact, the IPEA shift is expected to have a limited effect on the activation barrier of CT transition states. A level shift of 0.2 is used in CASPT2 calculations to exclude intruder states. Most likely, the ca. +3 kcal/mol blue-shifting effects seen at the CASPT2//CASSCF QM/MM calculations originate from the specific error cancellation associated with the selected/benchmarked protocol.

All QM calculations (i.e., geometry optimizations, vibrational free energy computations) were carried out using the MOLCAS computer package.[54] The corrected annihilation energies were computed (see Scheme 1). The weight of the CASSCF reference function in the CASPT2 reference was found to be similar to those in the different states and for all systems (e.g., for the Rh equilibrium structures were found to be 0.519, 0.511, and 0.513 for S_0, S_1, and S_2 respectively).

3.2. NAIP Molecular Switch. The construction of the solute−solvent model and the QM/MM free energy geometry optimization of NAIP switch configurations were carried out following the protocol mentioned in Section 2.1. To be more specific, the following steps are followed:

(1) The solute is optimized in its ground state at the MP2/6-31G*/PCM level of theory.

(2) The optimized structure is embedded at the center of a cubic methanol solvent box, and a 5 ns long (1 ns heating and equilibration, 4 ns production) classical MD simulation is performed at room temperature with the NPT ensemble. Both the solute and the methanol solvent are described by the OPLS-AA force field parameters,[55] and periodic boundary conditions are employed to simulate the solvent. This MD simulation is performed using the GROMACS molecular dynamics package.[56]

(3) An ASEC configuration is constructed by extracting 100 uncorrelated snapshots from the production stage of the MD simulation. Only solvent molecules within 20 Å of the solute are included in this ASEC configuration. The solute structure is optimized inside the average environment (i.e., inside the ASEC configuration) at the CASSCF/6-31G*/OPLS-AA level using MOLCAS/TINKER.[43,54,57] In this case, the solute is treated at the QM level and the solvent environment is treated at the MM level.

(4) The optimized solute structure is re-embedded in a solvent box and steps 2 and 3 are iteratively carried out

https://doi.org/10.1021/acs.jctc.1c00221
J. Chem. Theory Comput. 2021, 17, 5885−5895

until the energies of the optimized structure from two consecutive iterations become less than 0.5 kcal/mol.

The free energy geometry optimizations of *E*-NAIP, *Z*-NAIP, and the transition state connecting them (NAIP-TS) were all performed using the above protocol.

To compute the free energy difference between the isomers of the NAIP switches, the procedure described in Section 2.2 was used, as implemented in the DICE code[49] for solute molecules. In this case, the solvent environment was sampled using Monte Carlo (MC) simulations in the NPT ensemble and standard room conditions of temperature and pressure (298 K temperature and 1 atm, respectively) using the standard Metropolis Sampling technique[35] and periodic boundary conditions. Each MC simulation comprises an equilibration stage of 3.0×10^5 MC steps followed by a production stage of 6.0×10^5 MC steps. The OPLS-AA force field parameters[55] were used to describe both solute and the solvent during the MC simulations.

3.3. Transition State Calculation. For each system, a transition state was located using the following procedure: (i) compute, for each chromophore, a gas-phase transition-state structure using the default method implemented in Molcas 7.8. To prove that the optimized structure is a saddle point, we calculated the vibrational frequencies and intrinsic coordinate reaction (IRC). (ii) The gas-phase structure was then inserted into the protein/solvent environment and reoptimized at the QM/MM level. (iii) Because we could not perform the frequency/IRC analysis at such level, we compared the optimized transition-state energy with the energies of the structures generated by displacing the transition-state geometry backward and forward along the Hessian eigenvector corresponding to the negative eigenvalue of the Hessian (this Hessian is, again, calculated in the gas phase). If the energies of the two structures are lower than of the optimized transition state, we consider the optimization successful.

4. RESULTS AND DISCUSSION

4.1. Visual Bovine Rhodopsin. Rh and bathoRh were optimized using the ASEC-FEG protocol, and their vertical excitation energies (corresponding to λ_{max}) were computed. For Rh, the CASPT2//CASSCF/6-31G* value of 494 nm (57.8 kcal/mol) was obtained, which is consistent with our previously reported value as well as with experimental observations (see ref 14 and references cited therein). For bathoRh, we obtained a λ_{max} value of 503 nm (56.8 kcal/mol), and it is therefore correctly red-shifted, consistent with the experiment. The optimized Rh structure provided the starting point for the transition state (Rh-TS) optimization. Accordingly, we optimized the charge-transfer (CT) transition state of Rh both using ASEC-FEG and in the gas phase. To approximately check the nature of the optimized stationary point, we carried out vibrational frequency calculations in the gas phase, which yielded a single imaginary frequency (see the Supporting Information (SI) for the molecular structures of rPSB in Rh, bathoRh, and Rh-TS). The visualization of this vibrational mode shows that it corresponds to C11═C12 torsion. This provides evidence that the computed Rh transition state is the one related to the isomerization around the C11═C12 double bond. Furthermore, the total charge calculated for the fractions of rPSB divided by this double bond indicated that our transition state has a CT character as expected. For the optimized Rh, bathoRh, and Rh-TS structures, we calculated the free energy of solvation using the free energy perturbation approach. The resulting annihilation free energies are summarized in Table 1. To

Table 1. Calculated Values of Annihilation Free Energy (ΔG_X(solvent)), Free Energy in the Gas Phase (ΔG_X(gas)), Energy for Geometry Change (ΔE_X(gc)), and Free Energy for Polarization (ΔG_X(pol)) for Rh, bathoRh, and Rh-TS[a]

	Rh	bathoRh	Rh-TS
ΔG_X(solvent)[b]	−43.7	−42.3	−42.7
ΔG_X(gas)[b]	−546 213.4	−54 6224.1	−54 6191.5
ΔE_X(gc)[b]	−1.3	16.5	−1.53
ΔG_X(pol)[b]	6.16	4.74	6.70

[a]X has the Rh, bathoRh, and Rh-TS values. [b]The energies are in kcal/mol.

demonstrate that the FEP simulations have a sufficient number of intermediate states (scaling factors μ_i), we analyzed the potential energy differences, $\Delta U_{i\rightarrow j} = U(\mu_j) - U(\mu_i)$, where $j = i + 1$, for all successive perturbations to ensure the convergence of the average shown in eq 4. The distributions of $\Delta U_{i\rightarrow j}$ and the values of $\Delta G_{i\rightarrow j}$ are shown in the SI. Furthermore, the rPSB structure from each model was isolated in the gas phase and reoptimized and the vibrational free energy was calculated. The difference in energy arising from the geometric change was accounted for by performing CASPT2//CASSCF/6-31G* energy computations for the rPSB structures before and after the gas-phase geometry optimizations. The results of these calculations are summarized in Table 1 and Figure 3.

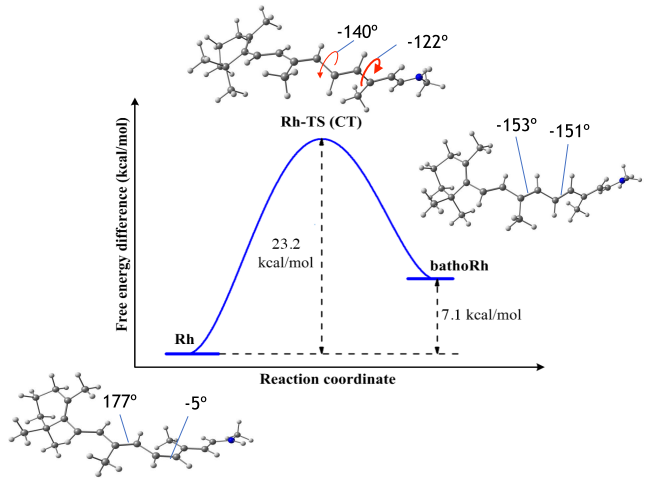


Figure 3. Reaction coordinate diagram for the thermal isomerization of Rh. The energies are in kcal/mol. The structure of the reactant (Rh), transition state (Rh-TS), and product (bathoRh) are also given together with the main torsional deformation (in terms of the corresponding C−C−C−C dihedral angle) characterizing the reaction coordinate.

Using eq 5 and data reported in Table 1, the calculated free energy barrier for Rh is 23.2 kcal/mol. This value is very close to the experimentally observed barrier (22.3 kcal/mol) for the 310.5−317.65 K temperature range.[36] This value is 10.8 kcal/mol lower than the previously reported potential energy barrier for the charge-transfer transition state (34 kcal/mol)[20] and 1.7 kcal/mol higher than the potential energy barrier calculated in the current study (21.5 kcal/mol). The resulting calculated free

https://doi.org/10.1021/acs.jctc.1c00221
J. Chem. Theory Comput. 2021, 17, 5885−5895

energy difference between Rh and bathoRh is 7.1 kcal/mol. This value is much lower than the experimentally measured enthalpy difference between bathoRh and Rh (32−35 kcal/mol)[58,59] and the enthalpy difference calculated in previous theoretical studies (16−17[60] and 21 kcal/mol[61]).

4.2. Anabaena Sensory Rhodopsin (ASR). The construction of models for ASR and the corresponding CT transition state was performed in the same manner as we have described for the Rh and Rh-TS above. The calculated λ_{max} values for the constructed ASR-13C and ASR-AT models are 535 and 543 nm, respectively. The computed transition state was verified by visual inspection of the vibrational motions and frequency analysis (see the SI for the molecular structures of all-*trans*, 13-*cis* isomers, and the transition state). Annihilation free energy, vibrational free energy computed in the gas phase, and the energy difference between gas-phase-isolated and gas-phase-optimized rPSB structures are summarized in Table 2. A summary of energies related to each annihilation step is reported in the SI (Table S2). The calculated free energy differences between two ASR isomers and the free energy barrier are summarized in Figure 4.

Table 2. Calculated Values of Annihilation Free Energy (ΔG_X(solvent)), Free Energy in the Gas Phase (ΔG_X(gas)), Energy for Geometry Change (ΔE_X(gc),) and Free Energy for Polarization (ΔG_X(pol)) for ASR Isomers and the Transition State

	ASR-13C	ASR-AT	ASR-TS
ΔG_X(solvent)[a]	−47.1	−49.0	−46.6
ΔG_X(gas)[a]	−546 216.7	−546 217.8	−546 195.4
ΔE_X(gc)[a]	2.9	5.3	−0.2
ΔG_X(pol)[a]	3.02	3.21	3.30

[a]The energies are in kcal/mol.

As reported in Figure 4, the computed free energy difference between ASR-AT and ASR-13C is 0.41 kcal/mol. This is an important improvement over previous theoretical investigations, where calculations predicted an incon-sistent stability trend relative to the ex-periment and with a 4.2 kcal/mol energy difference.[62,63] Here, the free energy corrections result in a trend of stability that is consistent with the experimental observations reported in Kandori et al.[62] The authors reported that the half-life of ASR-13C -> ASR-AT conversion is 90 min at 277 K temperature. We calculated the rate and the half-life for the same conversion at 300 K by the computed free energy barrier (19.4 kcal/mol) in the Eyring equation. This resulted in a rate constant of 0.0561 s^{-1} and 12.4 s half-life.

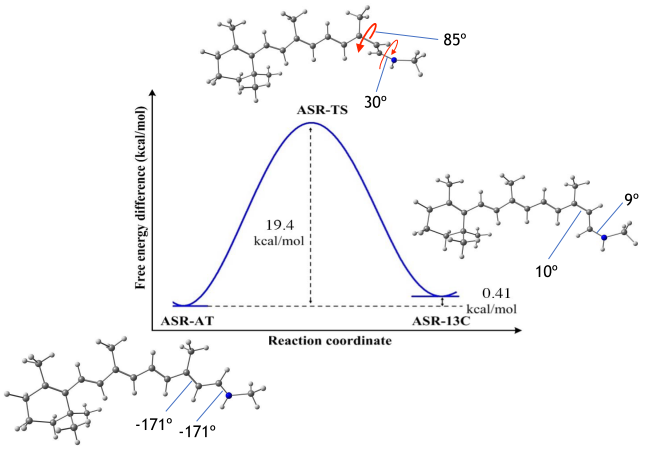


Figure 4. Reaction coordinate diagram for the thermal isomerization of ASR. The energies are in kcal/mol. The structures of the reactant (ASR-AT), transition state (ASR-TS), and product (ASR-13C) are also given together with the main torsional deformation (in terms of the corresponding C−C−C−C and C−C−N−C dihedral angles) characterizing the reaction coordinate.

4.3. NAIP Molecular Switch. The QM/MM models for solute−solvent systems of *E*-NAIP, *Z*-NAIP, and NAIP-TS were constructed using the procedure described in Section 3.2. The λ_{max} values computed for *E* and *Z* isomers are reported in Table 3. These values are compared with previously documented values based on the mentioned ASEP method and experiments.

Table 3. Comparison of λ_{max} Values Computed for NAIP Isomers Using ASEC-FEG against ASEP Results and Experimental Observations Reported in ref 4[a]

	E-NAIP	*Z*-NAIP
ASEC-FEG	364.3	367.3
ASEP	372.3	376.7
experiments	376.7	384.8

[a]The reported λ_{max} values are in nanometer.

As evident from Table 3, the predicted λ_{max} values using ASEC-FEG are blue-shifted with respect to ASEP and experimental values. The magnitude of this blue shift is 12−18 nm in comparison to experiments. The transition state optimization of the NAIP-TS model resulted in a solute structure ∼90° twisted around the exocyclic double bond. Frequency calculations carried out for this structure resulted in a single imaginary frequency, providing evidence for a transition state. The visual observation of the vibrational motions confirmed that the obtained transition state corresponds to the isomerization of the exocyclic double bond connecting the *Z*-NAIP to the *E*-NAIP structures (see the SI for structures of *E*-, *Z*-NAIP, and the transition state). The free energies of annihilation, geometry change, and polarization and the vibrational free energy computed in the gas phase are reported in Table 4. The distributions of potential energy differences $\Delta U_{i\rightarrow j}$ and the summary of energies related to each annihilation step are reported in the SI (Figures S4−S6, Table S3). The reaction coordinate diagram is presented in Figure 5.

Table 4. Calculated Values of Annihilation Free Energy (ΔG_X(solvent)), Free Energy in the Gas Phase (ΔG_X(gas)), Energy for Geometry Change (ΔE_X(gc)), and Free Energy for Polarization (ΔG_X(pol)) for NAIP Switch Isomers and the Transition State

	E-NAIP	*Z*-NAIP	NAIP-TS
ΔG_X(sol/prot)[a]	−67.25	−65.7	−60.14
ΔG_X(gas)[a]	−494 487.7	−494 490.6	−494 481.2
ΔG_X(gc)[a]	3.26	3.21	12.75
ΔG_X(pol)[a]	2.90	2.73	0.73

[a]The energies are in kcal/mol.

As evident from Figure 5, the computed free energy difference between *E*- and *Z*-NAIP is very close to the experimental value (1.4 kcal/mol). In fact, the difference we

https://doi.org/10.1021/acs.jctc.1c00221
J. Chem. Theory Comput. 2021, 17, 5885−5895

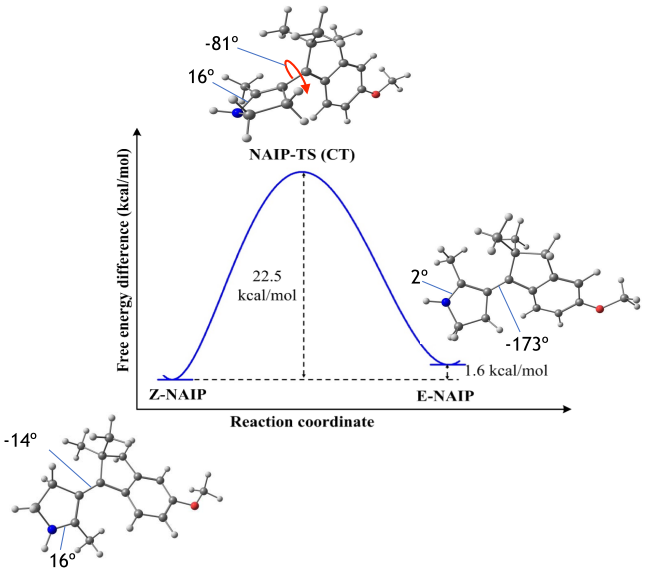


Figure 5. Reaction coordinate diagram for thermal isomerization of the NAIP molecular switch. The energies are in kcal/mol. The structures of the reactant (*Z*-NAIP), transition state (NAIP-TS), and product (*E*-NAIP) are also given together with the main torsional deformation (in terms of the corresponding C−C−C−C and C−C−N−C dihedral angles) characterizing the reaction coordinate.

have obtained is improved with respect to that obtained using the ASEP method reported in ref 4. Furthermore, the 22.5 kcal/mol free energy barrier, per the Eyring equation, corresponds to a half-life of 45 min at room temperature, which is in excellent agreement with the observation that the sample isomerizes in several hours when it is left in the dark (Personal communication: Marco Paolino).

5. CONCLUSIONS

We have reported and applied a novel computational method for computing the free energy difference between isomers and between an isomer and a transition state for a few selected rhodopsin and rhodopsin-like systems. Such a technique has been incorporated into the ASEC-FEG protocol and automated. More specifically, we computed the structure and relative stability of the reactant, product, and charge-transfer transition states of selected animal rhodopsin (Rh), microbial rhodopsin (ASR), and a biomimetic molecular switch on the free energy surface. The calculated free energy values at the Rh and ASR transition states indicate that the methodology correctly predicts the magnitude of the rPSB isomerization barriers even when this is embedded in a complex, organized, and relatively stiff molecular environment, even if relatively simple QM/MM models (but employing a multiconfigurational QM level of theory) are used. The same appears to be true for the NAIP chromophore isomers in methanol solution, which feature a rhodopsin-like chromophore embedded in the flexible solvent molecular environment. In the case of Rh, however, the bathoRh stability appears to be quantitatively overestimated. It is not clear if this is due to the approximations adopted in the bathoRh model construction, which assumes conserved ionization states of the residues of Rh and bathoRh. This also could be the cause of a too blue-shifted calculated bathoRh λ_{max} value. We note, however, an inconsistency between the reported experimentally measured barrier for isomerization in Rh and the measured bathoRh energy storage. This incon-sistency puts into question the accu-racy of the experi-mentally determined bathoRh stability.[58,59]

The calculation of thermal reaction barriers also gave barriers that are in good agreement with experiments for all three systems and, therefore, diverse molecular environments. Further testing and benchmarking of this methodology is certainly needed, but the results reported here indicate that the approach used is promising, even for transition states featuring a charge-transfer character with respect to the reactant and that are therefore strongly influenced by the solvent environment. Overall, we believe that the developed methods give access to an efficient way to investigate both the thermal equilibrium and reaction rates in systems that require a multiconfigurational QM treatment at the QM/MM level. Most importantly, the tested multiconfigurational QM treatment indicates that the same strategy/technology could be tested/applied to the description of other inherently difficult processes such as energy transfer and excited state luminescence and reactivity when these occur at the equilibrium. Since these processes are common in photobiological systems as well as synthetic molecular devices, the presented updated ASEC-FEG protocol should soon find other applications in these fields.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.1c00221.

Cartesian coordinates of the optimized structures and the values of all terms contributing the free energy differences (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Yoelvis Orozco-Gonzalez − *Department of Chemistry, Bowling Green State University, Bowling Green, Ohio 43403, United States; Department of Chemistry, Georgia State University, Atlanta, Georgia 30303, United States;* orcid.org/0000-0001-7225-2424; Email: yoelvis.orozco@gmail.com

Mikhail N. Ryazantsev − *Institute of Biomedical Systems and Biotechnologies, Peter the Great St. Petersburg Polytechnic University, St. Petersburg 195251, Russia; Saint Petersburg State University, St. Petersburg 199034, Russia;* orcid.org/0000-0003-3413-1706; Email: mikhail.n.ryazantsev@gmail.com

Massimo Olivucci − *Department of Chemistry, Bowling Green State University, Bowling Green, Ohio 43403, United States; Dipartimento di Biotecnologie, Chimica e Farmacia, Università di Siena, I-53100 Siena, Italy; Institut de Physique et Chimie des Matériaux de Strasbourg, Université de Strasbourg—CNRS, F-67034 Strasbourg, France;* orcid.org/0000-0002-8247-209X; Email: massimo.olivucci@unisi.it

Kaline Coutinho − *Instituto de Física, Universidade de São Paulo, São Paulo 05508-090, Brazil;* orcid.org/0000-0002-7586-3324; Email: kaline@if.usp.br

Authors

Dmitrii M. Nikolaev − *Nanotechnology Research and Education Centre RAS, Saint Petersburg Academic University, St. Petersburg 194021, Russia*

https://doi.org/10.1021/acs.jctc.1c00221
J. Chem. Theory Comput. 2021, 17, 5885−5895

Madushanka Manathunga − *Department of Chemistry, Bowling Green State University, Bowling Green, Ohio 43403, United States;* orcid.org/0000-0002-3594-8112
Andrey A. Shtyrov − *Nanotechnology Research and Education Centre RAS, Saint Petersburg Academic University, St. Petersburg 194021, Russia*
Yansel Omar Guerrero Martínez − *Instituto de Física, Universidade de São Paulo, São Paulo 05508-090, Brazil*
Samer Gozem − *Department of Chemistry, Georgia State University, Atlanta, Georgia 30303, United States;* orcid.org/0000-0002-6429-2853
Sylvio Canuto − *Instituto de Física, Universidade de São Paulo, São Paulo 05508-090, Brazil;* orcid.org/0000-0002-9942-8714

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.1c00221

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

M.O. is grateful to the Center for Photochemical Sciences of the Bowling Green State University, the Human Frontier Science Program Organization under Grant RGP0049/385, the National Science Foundation under Grant CHE-1152070, and the Institute for Advanced Studies of the University of Strasbourg for a USIAS fellowship. The research was supported by the Ohio Super-computing Center (OSC). The results of the work were obtained using computational resources of Peter the Great Saint-Petersburg Polytechnic University Supercomputing Center (www.spbstu.ru). The work has been performed under the Project HPC-EUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the D.M.N. and M.N.R. gratefully acknowledge the support of Prof. Massimo Olivucci and the computing resources and technical support provided by CINECA. M.N.R. acknowledges funding by St. Petersburg State University, travel grant COLLAB2019_2 (No. 41131617). S.G. acknowledges the NSF for its support through Grant CHE-2047667 and for computational resources through XSEDE research allocation CHE180027. K.C and S.C. thank the Brazilian funding agencies for Grants 2014/50983-3 of FAPESP and 465259/2014-6 of CNPq from the National Institute of Science and Technology Complex Fluids (INCT-FCx) and Grant 23038.004630/2014-35 of CAPES from the BioMol project. Code development for the application of FEP methodology for rPSB in the protein environment was supported by RSF under Grant 20-13-00303.

■ REFERENCES

(1) Spudich, J. L.; Yang, C.-S.; Jung, K.-H.; Spudich, E. N. Retinylidene proteins: structures and functions from archaea to humans. *Annu. Rev. Cell Dev. Biol.* **2000**, *16*, 365−392.
(2) Ernst, O. P.; Lodowski, D. T.; Elstner, M.; Hegemann, P.; Brown, L. S.; Kandori, H. Microbial and animal rhodopsins: structures, functions, and molecular mechanisms. *Chem. Rev.* **2014**, *114*, 126−163.
(3) Fenno, L.; Yizhar, O.; Deisseroth, K. The development and application of optogenetics. *Annu. Rev. Neurosci.* **2011**, *34*, No. 389.
(4) Sinicropi, A.; Martin, E.; Ryazantsev, M.; Helbing, J.; Briand, J.; Sharma, D.; Léonard, J.; Haacke, S.; Cannizzo, A.; Chergui, M.; et al. An artificial molecular switch that mimics the visual pigment and completes its photocycle in picoseconds. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 17642−17647.
(5) Melloni, A.; Rossi Paccani, R.; Donati, D.; Zanirato, V.; Sinicropi, A.; Parisi, M. L.; Martin, E.; Ryazantsev, M.; Ding, W. J.; Frutos, L. M.; et al. Modeling, preparation, and characterization of a dipole moment switch driven by Z/E photoisomerization. *J. Am. Chem. Soc.* **2010**, *132*, 9310−9319.
(6) Gozem, S.; Melaccio, F.; Luk, H.; Rinaldi, S.; Olivucci, M. Learning from photobiology how to design molecular devices using a computer. *Chem. Soc. Rev.* **2014**, *43*, 4019−4036.
(7) Pedraza-González, L.; Marin, M. D. C.; Jorge, A. N.; Ruck, T. D.; Yang, X.; Valentini, A.; Olivucci, M.; De Vico, L. Web-ARM: A web-based interface for the automatic construction of QM/MM models of rhodopsins. *J. Chem. Inf. Model.* **2020**, *60*, 1481−1493.
(8) Pedraza-González, L.; De Vico, L.; Marín, M. d. C.; Fanelli, F.; Olivucci, M. a-ARM: Automatic Rhodopsin Modeling with Chromophore Cavity Generation, Ionization State Selection, and External Counterion Placement. *J. Chem. Theory Comput.* **2019**, *15*, 3134−3152.
(9) Sumita, M.; Ryazantsev, M. N.; Saito, K. Acceleration of the Z to E photoisomerization of penta-2, 4-dieniminium by hydrogen out-of-plane motion: theoretical study on a model system of retinal protonated Schiff base. *Phys. Chem. Chem. Phys.* **2009**, *11*, 6406−6414.
(10) Nikolaev, D. M.; Shtyrov, A. A.; Mereshchenko, A. S.; Panov, M. S.; Tveryanovich, Y. S.; Ryazantsev, M. N. An assessment of the influence of applied water placement algorithms on Quantum Mechanics/Molecular Mechanics model quality: The case of rhodopsins first spectral absorption band maxima. *Phys. Chem. Chem. Phys.* **2020**, *22*, 18114−18123.
(11) Nikolaev, D. M.; Shtyrov, A. A.; Panov, M. S.; Jamal, A.; Chakchir, O. B.; Kochemirovsky, V. A.; Olivucci, M.; Ryazantsev, M. N. A Comparative Study of Modern Homology Modeling Algorithms for Rhodopsin Structure Prediction. *ACS Omega* **2018**, *3*, 7555−7566.
(12) Melaccio, F.; del Carmen Marín, M.; Valentini, A.; Montisci, F.; Rinaldi, S.; Cherubini, M.; Yang, X.; Kato, Y.; Stenrup, M.; Orozco-Gonzalez, Y.; et al. Toward automatic rhodopsin modeling as a tool for high-throughput computational photobiology. *J. Chem. Theory Comput.* **2016**, *12*, 6020−6034.
(13) Shtyrov, A. A.; Nikolaev, D. M.; Mironov, V. N.; Vasin, A. V.; Panov, M. S.; Tveryanovich, Y. S.; Ryazantsev, M. N. Simple Models to Study Spectral Properties of Microbial and Animal Rhodopsins: Evaluation of the Electrostatic Effect of Charged and Polar Residues on the First Absorption Band Maxima. *Int. J. Mol. Sci.* **2021**, *22*, No. 3029.
(14) Orozco-Gonzalez, Y.; Manathunga, M.; Marin, M. d. C.; Agathangelou, D.; Jung, K.-H.; Melaccio, F.; Ferré, N.; Haacke, S.; Coutinho, K.; Canuto, S.; et al. An average solvent electrostatic configuration protocol for Qm/mm free energy optimization: Implementation and application to rhodopsin systems. *J. Chem. Theory Comput.* **2017**, *13*, 6391−6640.
(15) Georg, H. C.; Canuto, S. Electronic properties of water in liquid environment. A sequential QM/MM study using the free energy gradient method. *J. Phys. Chem. B* **2012**, *116*, 11247−11254.
(16) Bistafa, C.; Georg, H. C.; Canuto, S. Combining ab initio multiconfigurational and Free Energy Gradient methods to study the $\pi-\pi^*$ excited state structure and properties of uracil in water. *Comput. Theor. Chem.* **2014**, *1040−1041*, 312−320.
(17) Okuyama-Yoshida, N.; Nagaoka, M.; Yamabe, T. Transition-state optimization on free energy surface: Toward solution chemical reaction ergodography. *Int. J. Quantum Chem.* **1998**, *70*, 95−103.
(18) Okuyama-Yoshida, N.; Kataoka, K.; Nagaoka, M.; Yamabe, T. Structure optimization via free energy gradient method: Application to glycine zwitterion in aqueous solution. *J. Chem. Phys.* **2000**, *113*, 3519−3524.
(19) Hirao, H.; Nagae, Y.; Nagaoka, M. Transition-state optimization by the free energy gradient method: Application to aqueous-phase Menshutkin reaction between ammonia and methyl chloride. *Chem. Phys. Lett.* **2001**, *348*, 350−356.

(20) Gozem, S.; Schapiro, I.; Ferré, N.; Olivucci, M. The molecular mechanism of thermal noise in rod photoreceptors. *Science* **2012**, *337*, 1225−1228.
(21) Govardovskii, V. I.; Astakhova, L. A.; Rotov, A. Y.; Firsov, M. L. Rejection of the biophoton hypothesis on the origin of photoreceptor dark noise. *J. Gen. Physiol.* **2019**, *151*, 887−897.
(22) Hayashi, S.; Uchida, Y.; Hasegawa, T.; Higashi, M.; Kosugi, T.; Kamiya, M. QM/MM geometry optimization on extensive free-Energy surfaces for examination of enzymatic reactions and design of novel functional properties of proteins. *Annu. Rev. Phys. Chem.* **2017**, *68*, 135−154.
(23) Kumon, K.; Higashi, M.; Saito, S.; Hayashi, S. Protein flexibility catalyzes a cell signaling reaction of the Ras-GAP complex *ChemRxiv* 2019, DOI: 10.26434/chemrxiv.9679028.
(24) Bennett, C. H. Efficient estimation of free energy differences from Monte Carlo data. *J. Comput. Phys.* **1976**, *22*, 245−268.
(25) Galván, I. F.; Sánchez, M.; Martín, M.; Olivares del Valle, F.; Aguilar, M. Geometry optimization of molecules in solution: Joint use of the mean field approximation and the free-energy gradient method. *J. Chem. Phys.* **2003**, *118*, 255−263.
(26) Woodcock, H. L., III; Hodošček, M.; Gilbert, A. T.; Gill, P. M.; Schaefer, H. F., III; Brooks, B. R. Interfacing Q-Chem and CHARMM to perform QM/MM reaction path calculations. *J. Comput. Chem.* **2007**, *28*, 1485−1502.
(27) König, G.; Mei, Y.; Pickard, F. C., IV; Simmonett, A. C.; Miller, B. T.; Herbert, J. M.; Woodcock, H. L.; Brooks, B. R.; Shao, Y. Computation of hydration free energies using the multiple environment single system quantum mechanical/molecular mechanical method. *J. Chem. Theory Comput.* **2016**, *12*, 332−344.
(28) Pickard, F. C., IV; König, G.; Simmonett, A. C.; Shao, Y.; Brooks, B. R. An efficient protocol for obtaining accurate hydration free energies using quantum chemistry and reweighting from molecular dynamics simulations. *Bioorg. Med. Chem.* **2016**, *24*, 4988−4997.
(29) Jia, X.; Wang, M.; Shao, Y.; König, G.; Brooks, B. R.; Zhang, J. Z.; Mei, Y. Calculations of solvation free energy through energy reweighting from molecular mechanics to quantum mechanics. *J. Chem. Theory Comput.* **2016**, *12*, 499−511.
(30) Tao, P.; Sodt, A. J.; Shao, Y.; König, G.; Brooks, B. R. Computing the free energy along a reaction coordinate using rigid body dynamics. *J. Chem. Theory Comput.* **2014**, *10*, 4198−4207.
(31) Li, P.; Jia, X.; Pan, X.; Shao, Y.; Mei, Y. Accelerated computation of free energy profile at ab initio quantum mechanical/molecular mechanics accuracy via a semi-empirical reference potential. I. Weighted thermodynamics perturbation. *J. Chem. Theory Comput.* **2018**, *14*, 5583−5596.
(32) Pan, X.; Li, P.; Ho, J.; Pu, J.; Mei, Y.; Shao, Y. Accelerated computation of free energy profile at ab initio quantum mechanical/molecular mechanical accuracy via a semi-empirical reference potential. II. Recalibrating semi-empirical parameters with force matching. *Phys. Chem. Chem. Phys.* **2019**, *21*, 20595−20605.
(33) da Cunha, A. R.; Duarte, E. L.; Lamy, M. T.; Coutinho, K. Protonation/deprotonation process of Emodin in aqueous solution and pKa determination: UV/Visible spectrophotometric titration and quantum/molecular mechanics calculations. *Chem. Phys.* **2014**, *440*, 69−79.
(34) Lima, M. C. P.; Coutinho, K.; Canuto, S.; Rocha, W. R. Reaction mechanism and tautomeric equilibrium of 2-mercaptopyrimidine in the gas phase and in aqueous solution: a combined Monte Carlo and quantum mechanics study. *J. Phys. Chem. A* **2006**, *110*, 7253−7261.
(35) Jorgensen, W. L.; Buckner, J. K.; Boudon, S.; Tirado-Rives, J. Efficient computation of absolute free energies of binding by computer simulations. Application to the methane dimer in water. *J. Chem. Phys.* **1988**, *89*, 3742−3746.
(36) Guo, Y.; Sekharan, S.; Liu, J.; Batista, V. S.; Tully, J. C.; Yan, E. C. Unusual kinetics of thermal decay of dim-light photoreceptors in vertebrate vision. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 10438−10443.
(37) Jung, K. H.; Trivedi, V. D.; Spudich, J. L. Demonstration of a sensory rhodopsin in eubacteria. *Mol. Microbiol.* **2003**, *47*, 1513−1522.
(38) Schapiro, I.; Ruhman, S. Ultrafast photochemistry of anabaena sensory rhodopsin: Experiment and theory. *Biochim. Biophys. Acta, Bioenerg.* **2014**, *1837*, 589−597.
(39) Coutinho, K.; Georg, H.; Fonseca, T.; Ludwig, V.; Canuto, S. An efficient statistically converged average configuration for solvent effects. *Chem. Phys. Lett.* **2007**, *437*, 148−152.
(40) Coutinho, K.; Rivelino, R.; Georg, H. C.; Canuto, S. The Sequential QM/MM Method and Its Applications to Solvent Effects in electronic and Structural Properties of solutes. *Solvation Effects on Molecules and Biomolecules*; Springer, 2008; pp 159−189.
(41) Cardenuto, M. H.; Cezar, H. M.; Mikkelsen, K. V.; Sauer, S. P.A.; Coutinho, K.; Canuto, S. A QM/MM study ofthe conformation stability and elec-tronic structure of the photochromic switches derivatives of DHA/VHF in acetonitrile solution. *Spectrochim. Acta, Part A* **2021**, *251*, 119434.
(42) Ferré, N.; Ángyán, J. G. Approximate electrostatic interaction operator for QM/MM calculations. *Chem. Phys. Lett.* **2002**, *356*, 331−339.
(43) Melaccio, F.; Olivucci, M.; Lindh, R.; Ferré, N. Unique QM/MM potential energy surface exploration using microiterations. *Int. J. Quantum Chem.* **2011**, *111*, 3339−3346.
(44) Kongsted, J.; Osted, A.; Mikkelsen, K. V.; Christiansen, O. The QM/MM approach for wavefunctions, energies and response functions within self-consistent field and coupled cluster theories. *Mol. Phys.* **2002**, *100*, 1813−1828.
(45) Martın, M.; Sánchez, M.; Olivares del Valle, F.; Aguilar, M. A multiconfiguration self-consistent field/molecular dynamics study of the (n→ π^*) 1 transition of carbonyl compounds in liquid water. *J. Chem. Phys.* **2000**, *113*, 6308−6315.
(46) Zwanzig, R. W. High-temperature equation of state by a perturbation method. I. Nonpolar gases. *J. Chem. Phys.* **1954**, *22*, 1420−1426.
(47) Kollman, P. Free energy calculations: applications to chemical and biochemical phenomena. *Chem. Rev.* **1993**, *93*, 2395−2417.
(48) von Ragué Schleyer, P.; Schreiner, P. R.; Schaefer, H. F., III; Jorgensen, W. L.; Thiel, W.; Glen, R. C.; Allinger, L. N.; Clark, T.; Gasteiger, J.; Chandler, D. *Encyclopedia of Computational Chemistry*; John Wiley & Sons, Ltd., 1998.
(49) Cezar, H. M.; Canuto, S.; Coutinho, K. DICE: A Monte Carlo Code for Mo-lecular Simulation Including the Configurational Bias Monte Carlo Method. *J. Chem. Inf. Model.* **2020**, *60*, 3472−3488.
(50) Dundas, J.; Ouyang, Z.; Tseng, J.; Binkowski, A.; Turpaz, Y.; Liang, J. CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res.* **2006**, *34*, W116−W118.
(51) Roos, B. O. The complete active space self-consistent field method and its applications in electronic structure calculations. *Adv. Chem. Phys.* **1987**, *69*, 399−445.
(52) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(53) Andersson, K.; Malmqvist, P. Å.; Roos, B. O. Second-order perturbation theory with a complete active space self-consistent field reference function. *J. Chem. Phys.* **1992**, *96*, 1218−1226.
(54) Aquilante, F.; De Vico, L.; Ferré, N.; Ghigo, G.; Malmqvist, P. å.; Neogrády, P.; Pedersen, T. B.; Pitoňák, M.; Reiher, M.; Roos, B. O.; et al. MOLCAS 7: the next generation. *J. Comput. Chem.* **2010**, *31*, 224−247.
(55) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.
(56) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der

Spoel, D.; et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **2013**, *29*, 845−854.
(57) Ponder, J. W.; Richards, F. M. An efficient newton-like method for molecular mechanics energy minimization of large molecules. *J. Comput. Chem.* **1987**, *8*, 1016−1024.
(58) Cooper, A. Energy uptake in the first step of visual excitation. *Nature* **1979**, *282*, 531−533.
(59) Schick, G. A.; Cooper, T. M.; Holloway, R. A.; Murray, L. P.; Birge, R. R. Energy storage in the primary photochemical events of rhodopsin and isorhodopsin. *Biochemistry* **1987**, *26*, 2556−2562.
(60) Khrenova, M.; Bochenkova, A.; Nemukhin, A. Modeling reaction routes from rhodopsin to bathorhodopsin. *Proteins* **2010**, *78*, 614−622.
(61) Strambi, A.; Coto, P. B.; Frutos, L. M.; Ferré, N.; Olivucci, M. Relationship between the excited state relaxation paths of rhodopsin and isorhodopsin. *J. Am. Chem. Soc.* **2008**, *130*, 3382−3388.
(62) Kawanabe, A.; Furutani, Y.; Jung, K.-H.; Kandori, H. Photochromism of Anabaena sensory rhodopsin. *J. Am. Chem. Soc.* **2007**, *129*, 8644−8649.
(63) Strambi, A.; Durbeej, B.; Ferré, N.; Olivucci, M. Anabaena sensory rhodopsin is a light-driven unidirectional rotor. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 21322−21326.

https://doi.org/10.1021/acs.jctc.1c00221
J. Chem. Theory Comput. 2021, 17, 5885−5895