Cis-to-Trans Isomerization of Azobenzene Derivatives Studied with Transition Path Sampling and Quantum Mechanical/Molecular Mechanical Molecular Dynamics

Muždalo, A.; Saalfrank, P.; Vreede, J.; Santer, M.

DOI
10.1021/acs.jctc.7b01120

Publication date
2018

Document Version
Final published version

Published in
Journal of Chemical Theory and Computation

JCTC Journal of Chemical Theory and Computation

Article

Cite This: J. Chem. Theory Comput. 2018, 14, 2042−2051

pubs.acs.org/JCTC

Cis-to-*Trans* Isomerization of Azobenzene Derivatives Studied with Transition Path Sampling and Quantum Mechanical/Molecular Mechanical Molecular Dynamics

Anja Muždalo,[†] Peter Saalfrank,[‡] Jocelyne Vreede,[§] and Mark Santer*[,†]

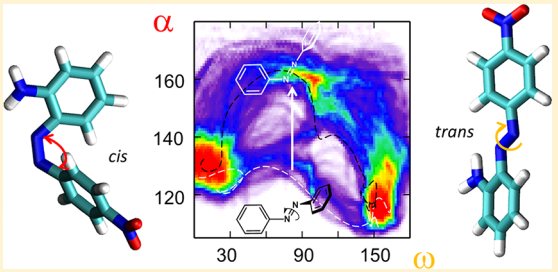
[†]Department of Theory and Biosystems, Max Planck Institute of Colloids and Interfaces, 14424 Potsdam, Germany

[‡]Department of Chemistry, University of Potsdam, 14476 Potsdam, Germany

[§]Computational Chemistry, Van't Hoff Institute for Molecular Sciences, University of Amsterdam, 1098 XH Amsterdam, The Netherlands

Supporting Information

ABSTRACT: Azobenzene-based molecular photoswitches are becoming increasingly important for the development of photoresponsive, functional soft-matter material systems. Upon illumination with light, fast interconversion between a more stable *trans* and a metastable *cis* configuration can be established resulting in pronounced changes in conformation, dipole moment or hydrophobicity. A rational design of functional photosensitive molecules with embedded azo moieties requires a thorough understanding of isomerization mechanisms and rates, especially the thermally activated relaxation. For small azo derivatives considered in the gas phase or simple solvents, Eyring's classical transition state theory (TST) approach yields useful predictions for trends in activation energies or corresponding half-life times of the *cis* isomer. However, TST or improved theories cannot easily be applied when the azo moiety is part of a larger molecular complex or embedded into a heterogeneous environment, where a multitude of possible reaction pathways may exist. In these cases, only the sampling of an ensemble of dynamic reactive trajectories (transition path sampling, TPS) with explicit models of the environment may reveal the nature of the processes involved. In the present work we show how a TPS approach can conveniently be implemented for the phenomenon of relaxation−isomerization of azobenzenes starting with the simple examples of pure azobenzene and a push−pull derivative immersed in a polar (DMSO) and apolar (toluene) solvent. The latter are represented explicitly at a molecular mechanical (MM) and the azo moiety at a quantum mechanical (QM) level. We demonstrate for the push−pull azobenzene that path sampling in combination with the chosen QM/MM scheme produces the expected change in isomerization pathway from inversion to rotation in going from a low to a high permittivity (explicit) solvent model. We discuss the potential of the simulation procedure presented for comparative calculation of reaction rates and an improved understanding of activated states.



1. INTRODUCTION

In recent years azobenzene-based derivatives have received considerable attention as photoswitchable moieties in functional biomolecules, such as photosensitive, shape-changing agents for microgels[1] and DNA[2−4] or as a switch to control the multivalent character of glycooligomers binding to lectins.[5] There is great interest in tuning the properties of azo derivatives such as to alter the optical wavelength they are responsive to or to increase their thermal stability. For instance, by applying various fluorinations of basic azobenzene, the back-isomerization may be triggered in the visible spectrum, while at the same time rather long lifetimes of the *cis* isomer can be established.[6]

Tailoring functional molecules containing such photosensitive moieties also requires a thorough understanding of how these properties are influenced by the environment in which they are embedded. There is a large body of experimental work devoted to a comprehensive characterization of reaction rates and their dependence on ring substituents or external solvent type or pressure.[7] Computational studies can provide essential insights into the nature of the isomerization mechanism, involving both, photoexcited pathways and thermal *cis*−*trans* relaxation in the ground state,[8,9] complementary to experiments. As long as the discussion can be restricted to the azo moiety itself, there usually are only a few distinct isomerization pathways that might compete, depending on substituents or solvent type. From the saddle point structure one may infer, e.g., a rotation mechanism, which involves mainly a torsion around the central C1′−N′=N−C1 dihedral and the inversion pathway through a transition state in which the atomic sequence in one of the bending angles becomes almost collinear;[10,11] see, for example, N′=N−C1 in the schematics of Figure 1a,c. In a seminal paper on isomerization

Received: November 13, 2017
Published: February 21, 2018

DOI: 10.1021/acs.jctc.7b01120
J. Chem. Theory Comput. 2018, 14, 2042−2051

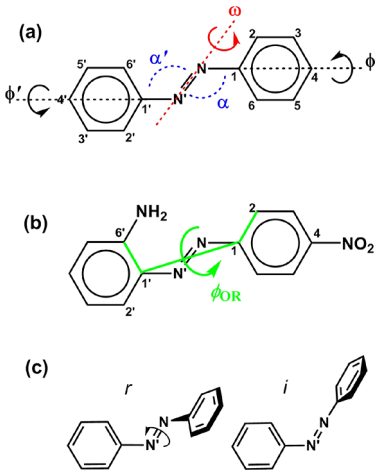


Figure 1. Systems studied and their relevant structural variables: (a) azobenzene (AB) and (b) the push−pull azobenzene (pp-AB) derivative $4NO_2$-6′NH_2-AB. The central torsion angle ω depicted in (a) around the double bond N′═N is defined by the atomic sequence C1′−N′═N−C1; the relevant bending angle α by N′═N−C1. In (b) the torsion angle ϕ_{OR} defined by the sequence C6′−C1′−C1−C2 is used to measure the relative orientation of the phenyl rings. (c) Major isomerization pathways on the ground-state BO surface: *r*, rotation around the central N′═N bond; *i*, inversion route, on which α becomes nearly 180°. We shall refer to the unprimed N atom as the inversion center.

on ground- and excited-state Born−Oppenheimer (BO) surfaces, the inversion mechanism was identified as the preferred one for thermal relaxation in the ground state with the phenyl rings perpendicular to each other, on excited state surfaces inversion as well as rotation can be expected to contribute;[12] in addition, a yet more complicated pathway was suggested in which the phenyl rings retain their relative orientation during the rotation around the central N′═N bond. This maneuver requires simultaneous twisting around an adjacent C−N bond, reminiscent of a "hula-twist".[13]

In general, starting from a saddle point analysis one may apply several techniques to characterize activation energies or reaction rates, starting from Eyring's classical transition state theory (TST),[14] which is usually applied to small-molecule reactions in the gas phase. Environmental effects such as solvent permittivity can be accounted for by implicit models;[8] the impact of noise by introducing models for frequency dependent friction.[15] If a reaction coordinate can be found that (approximately) defines a dividing surface at or in the vicinity of the barrier top, the so-called reactive flux passing through it can be measured by carrying out explicit molecular dynamics simulations with "off-the-barrier" trajectories.[16] This, however, is numerically feasible only if a convenient and simple set of order parameters can be found.[17] Indeed there are many occasions in which the influence of the environment becomes nontrivial, precluding a simple description of the barrier region. Ultimately we would like to consider azo moieties embedded in large molecular complexes such as polymer matrices[18] or attached to a surface forming a dense monolayer.[19] Reactive pathways then might largely be dictated by external restraints or reorganization dynamics. Even for a phenomenon as simple as the dissociation of ion pairs in water, the solvent coordination must be taken into account to understand the kinetics of separation.[20] The relevant information about reaction mechanisms and rates must then be inferred from the ensemble of MD trajectory segments connecting the stable states. When high energetic barriers make the corresponding transitions rare events, the transition path sampling (TPS) approach devised by Chandler, Bolhuis, Dellago, and others can be employed.[21,22]

In the present work, we shall present a convenient implementation of a TPS procedure for studying thermal *cis*−*trans* relaxation of azo derivatives, starting with two simple examples: pure azobenzene (AB) and a so-called push−pull derivative (pp-AB) with electron-donating (NH_2) and -withdrawing (NO_2) substituents; see Figure 1a,b. These species are studied in explicit solvents, DMSO (polar) and toluene (apolar), the latter two being treated at the molecular mechanical (MM) level. pp-AB is expected to respond to a change of solvent permittivity (apolar to polar solvent) with a change in reaction pathway from inversion to rotation, respectively,[7,8] and represents a simple albeit nontrivial example of environmental influence. Due to the complex conformation-dependent mesomeric characteristics of the azo moieties, the latter must be modeled at the quantum mechanical (QM) level. As a compromise between numerical efficiency and accuracy, we chose the recent variant of the semiempirical density functional tight binding method (DFTB3) to set up a QM/MM scheme to be outlined in the following section. Although the concept of combining TPS and QM/MM has successfully been employed by several groups before to study catalyzed chemical reactions,[23−25] it does not seem to be in widespread use; we shall therefore also give a somewhat extended account of how the TPS procedure is set up.

2. METHODS

2.1. QM/MM Scheme. All quantum-chemical (QM) calculations were performed using the Gaussian 03 program package.[26] With the QM/MM scheme in this work we employ electronic embedding as described by Fields,[27] using a third-order density functional tight binding Hamiltonian (DFTB3) introduced by Elstner and co-workers.[28−30] DFTB3 has been integrated into the GROMACS molecular dynamics package[31−34] for use with the particle-mesh Ewald (PME) method to treat long-range electrostatic interactions.[35] The Lennard-Jones (LJ) interaction parameters of the MM solvents were taken from the General Amber Force Field (GAFF).[36] A particular adaption (see, e.g., Freindorf et al.[37]) is not needed at this stage, as we are mainly interested in the coarse differences between vacuum and the presence of explicit solvent on the one hand, and strongly varying permittivities (conveyed by solvent polarity) on the other. Partial atomic charges were then assigned to DMSO and toluene molecules by restrained electrostatic potential (RESP)[38] fitting to the molecular electrostatic potential obtained from a calculation at the HF/6-31G* level. All MM parameters derived in this work are summarized in the Supporting Information, section 4. The van der Waals interactions were treated with a cutoff of 1.4 nm. PME was used with a grid spacing of 0.12 nm and an interpolation order 4. The systems in vacuum were set up starting from the optimized structure for the *cis* isomer followed by a 100 ps long equilibration run with stochastic dynamics. For the latter, a third order accurate stochastic integrator introduced by van Gunsteren and Berendsen[39] was employed with a time step of 1 fs, and a friction constant of τ = 0.1 ps, keeping an average temperature of 333 K.

The systems in solvent were set up by first solvating the optimized structures by random insertion into a cubic box with edge length of ∼4 nm, containing either toluene or DMSO.

DOI: 10.1021/acs.jctc.7b01120
J. Chem. Theory Comput. 2018, 14, 2042−2051

The solvated structures were then first relaxed in 500 steps with a steepest descent algorithm where the whole system was still described at the MM level (including the QM subsystem, with parameters from Duchstein et al.[40]), followed by a 1 ns long equilibration period in the NPT ensemble with the pressure kept at 1 bar. Finally, only the representation of the azobenzene derivatives was switched to the QM level and the whole system equilibrated for another 100 ps. The QM part now consists of 24 or 28 atoms for the derivatives AB and pp-AB, respectively, while the MM part of the system contained 308 and 458 solvent molecules for AB in toluene and DMSO, and 309 and 455 solvent molecules for pp-AB in toluene and DMSO, respectively. To facilitate the sampling of reactive pathways for *cis*−*trans* isomerization several different simulation techniques have been used, such as metadynamics and umbrella sampling for acquiring potential of mean force (PMF) maps, or steered molecular dynamics for generating initial pathways. These are all flavors of biased molecular dynamics, in our work facilitated by the PLUMED plugin v.2.1[41] to GROMACS v.5.1. In what follows we shall give an overview of how the TPS procedure is implemented for our case, details are given in the Results section.

2.2. Transition Path Sampling. In principle, transition pathways could be obtained from an asymptotically long equilibrium trajectory during which isomerization events are observed many times. Such a trajectory contains all information about reaction rates and mechanisms. If one is primarily interested in the latter, one may concentrate on those (short) segments of the trajectory that lead from a reactant state A (*cis* for azobenzene) to a product state B (*trans*) once without intermediate return to A, that is, the so-called ensemble of transition paths. Transition path sampling (TPS) is a technique by which these segments can be "harvested" very efficiently from a Metropolis Monte Carlo procedure:[21,22] trial trajectories are generated from already harvested transition pathways and added to the ensemble according to some acceptance criterion. This is a rather general idea that can be realized in many flavors. In the present work, we implement a stochastic "two-way shooting":[42] the system dynamics is restarted from a molecular configuration at an intermediate time t_i along an old trajectory, but with randomly perturbed momenta (e. g., by restarting the stochastic integrator). The dynamics is then integrated forward and backward in time. The new pathway is accepted if it starts in A at t_0 and ends in B at t_N, and rejected otherwise; here the number N of integration time steps is kept constant during a path sampling procedure.

In Figure 2 we illustrate the key aspects of the TPS implementation and introduce our notation. Ideally, reactant (A) and product state (B) are defined as well-separated low-energy regions in configuration (coordinate-) space (dark shaded compact areas in Figure 2). In our case these would be the set of molecular configurations visited during equilibrium in a metastable state *cis* or *trans*, that is, excluding the rare excursions to high-energy conformations. We first define a single order parameter X providing a foliation of hypersurfaces $X(\{\mathbf{r}_j\})$ = const between *cis* and *trans* useful for both types of transition pathways (see Figure 1c), $\{\mathbf{r}_j\}$ denoting the set of Cartesian coordinates of the system. By performing umbrella sampling in X, its potential of mean force $F(X)$ can be acquired, yielding the location X^* of the barrier top in this variable.

It is worth mentioning at this point that the definition of a useful order parameter indicating the progress/completion of a reaction depends on the method chosen, the problem type and also the nature of the dynamics (classic, stochastic, quantal, etc.). In the present case of TPS, strictly speaking, only a definition of the two stable states A and B is needed. In the reactive flux (RF)[16] method an extension up to the barrier top region is required in order to obtain a dividing hypersurface between the two basins of attraction; in classical TST or its extension variational TST[43] a dividing plane in configuration or phase space is sought running through the bottleneck of a reaction (provided such a unique rate-limiting bottleneck exists), and the corresponding order parameter need not necessarily be extended toward either stable state. In the conclusions section, we will further discuss the scope of the coordinate X to be defined shortly.

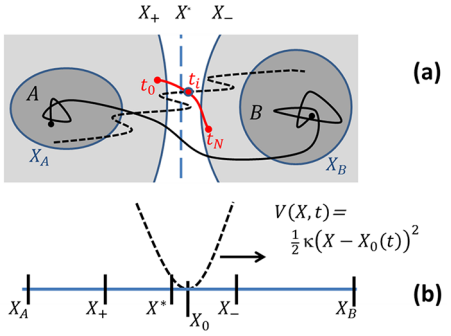


Figure 2. (a) Schematic view on the (free) energy landscape of the azobenzene isomerization dynamics. Dashed line: dynamic pathway connecting the basins A (*cis*) and B (*trans*) obtained from steered MD. Red: transition path segment obtained from the steered path by restarting the dynamics at an intermediate time t_i, representing a molecular configuration in the high energy interface region (blank concave area). This interface is bounded by two hypersurfaces in the reaction coordinate $X = X_-$ and $X = X_+$ (the dashed line $X = X^*$ represents the barrier top in $F(X)$), and the TPS procedure is carried out by restarting trajectories within the interface. Black solid line: transition path obtained from integrating the dynamics over a sufficiently long time such that the trajectory begins to settle in either state. (b) Schematic illustration of the steering process with a moving harmonic potential in X.

To obtain an initial pathway from which the TPS procedure can be started, we use steered molecular dynamics, dragging the molecule by a moving harmonic potential in X as indicated in Figure 2b. Path sampling is then conducted by generating trajectory segments that bridge a high-energy region bracketing X^* (blank, concave area in Figure 2a bounded by hypersurfaces $X = X_+$ and $X = X_-$). From within this interface, configurations corresponding to intermediate times t_i for trajectory restart are selected. To enhance the numerical efficiency of the procedure, we extend the definition *cis* and *trans* toward X_+ and X_-, respectively, adjacent to the interface region (light gray areas). Trajectories thus enter/leave the stable states when they leave/enter interface region $I_{\text{shoot}} = [X_-, X_+]$ according to $X(t)$. This reduces the computational cost related to the QM/MM setting significantly as the number of time steps N can be kept quite small; typically, a few thousand pathways must be sampled in order to achieve sufficient diversity. The above restriction is admissible, as we are mainly interested in the isomerization mechanism, which is mostly characterized by conformational changes close to the barrier top of the transition. We will choose I_{shoot} in such a way that sufficient distinction between the two stable states is provided, while obtaining a reasonable acceptance rate for new trajectories. It should be re-emphasized, however, that for now the actual estimation of

DOI: 10.1021/acs.jctc.7b01120
J. Chem. Theory Comput. 2018, 14, 2042−2051

Table 1. Energies and Transition Structure Geometry of AB in Various QM Descriptions[a]

method	ΔE_{ct}	$\Delta E^{\ddagger}$	$\Delta G^{\ddagger}$	ω	α	α'
B3LYP	63.58	105.97	96.97	179.10	179.43	117.00
PCM DMSO	54.93	111.38	104.00	171.90	179.25	117.62
PCM toluene	59.54	108.34	99.85	172.07	179.36	117.27
PBE	56.75	96.97	90.65	−143.63	178.76	116.58
HF	72.33	135.47	128.68	179.87	179.86	117.36
DFTBA	18.52	103.17	94.57	−107.64	179.95	114.96
PM3	9.22	113.45	111.50	−177.12	176.66	121.38
DFTB3	8.23 (7.41)	(95.95)	(89.63)	(−143.63)	(178.76)	(116.58)

[a]The basis set used for HF and DFT levels of theory was 6-31G*. The PCM method was used with B3LYP/6-31G*. Numbers in parentheses for DFTB3 refer to SP calculations on structures optimized with the PBE functional; in this case the value for $\Delta G^{\ddagger}$ was obtained by substituting in the expression $\Delta G^{\ddagger} = \Delta H^{\ddagger}(T) - T\Delta S^{\ddagger}$, $\Delta H^{\ddagger}(T) = \Delta E^{\ddagger} + \Delta E_{ZPE} + \Delta h(T)$, $\Delta E^{\ddagger}$ of the DFTB3 calculation for the PBE result. $\Delta h(T)$ is the temperature-dependent contribution to the enthalpy and ΔE_{ZPE} the difference in zero point energy (PBE). Energies are in kJ/mol, angles in degrees.

Table 2. Same as Table 1, but for pp-AB

method	ΔE_{ct}	$\Delta E^{\ddagger}$	$\Delta G^{\ddagger}$	ω	α	α'
B3LYP	59.29	80.22	74.51	−72.09	179.62	119.42
PCM DMSO	52.49	61.84	58.67	−89.24	134.43	121.10
PCM toluene	55.51	74.77	67.81	−52.27	179.13	119.99
PBE	51.75	70.91	63.85	102.56	179.42	119.34
HF	68.53	113.25	108.78	43.29	179.68	119.04
DFTBA	21.86	74.03	66.95	94.68	179.70	116.37
PM3	8.58	101.75	97.22	−91.57	177.56	119.37
DFTB3	11.91 (13.95)	(54.50)	(47.44)	(102.56)	(179.42)	(119.34)

reaction rates is set aside; in order to determine rates in the framework of TPS, additional, considerable computational labor would be required involving the sampling of a series of ensembles with increasing path lengths (durations).[21]

Note that the conception of a transition state as the lowest first-order saddle point of the BO surface in TST turns into one of a transition state *ensemble* (TSE) in TPS; for each transition path, a transition state (TS) may be defined as a configuration from which the reactant (product) state is reached with probability $p_{cis} = 0.5$ ($p_{trans} = 0.5$). This probability can be estimated as the fraction of trajectories that reach either state first, restarting from the same configuration and redrawing their initial velocities from a Maxwell−Boltzmann distribution each time. For estimating p_{trans} for a specific configuration, we used 100 restarted trajectories throughout.

3. RESULTS AND DISCUSSION

3.1. Structural Characterization. The choice of the QM method within the QM/MM framework is not only a natural compromise between accuracy and efficiency; the DFTB3 method employed in this work offers several advantages over other, semiempirical or first-principles methods. From a practical viewpoint, we can make use of a recent, very efficient new QM/MM implementation of DFTB3 into the Gromacs package.[35] Such a combination allows us to perform the necessary, extensive path sampling as outlined above. Furthermore, DFTB3 rests on a third-order expansion of the exchange-correlation functional around a reference density, and is known to perform well for systems involving H-bonding. It allows to treat (unlike other DFTB schemes) the attractive dispersion interactions within the QM part itself by using Grimme's D3 scheme.[44] These features will be useful for future studies, in particular when the investigations are extended to larger azo complexes, possibly involving strongly H-bonding solvents.

To characterize the DFTB3 method to some extent, we compare geometric parameters and single point (SP) energies of optimized structures of AB and pp-AB in their *cis* and *trans* conformations across various levels of theory. Several reference calculations were performed with non-semiempirical methods, namely HF and two flavours of DFT (B3LYP and PBE) all of them employing the 6-31G* basis set. For some of these, solvent effects were approximately included by a polarizable continuum model (PCM).[45] It has been shown elsewhere, that the combination B3LYP/6-31G* is indeed a surprisingly accurate reference for *cis*-to-*trans* isomerization of azobenzenes, also in comparison to highly correlated wave function methods with extended basis sets.[46] The results for stable state structures for B3LYP, PBE, DFTBA, and DFTB3 are rather close, and are summarized in the Supporting Information, section 1, Table 1. The energetics of the isomerization is reported here in Tables 1 and 2. For the energy difference of the optimized structures in *cis* and *trans* (ΔE_{ct}) one finds that all semiempirical methods (PM3, DFTBA and DFTB3) clearly underestimate this quantity in comparison to B3LYP. However, as has been pointed out in the previous section, for the present purpose the beyond-barrier dynamics is of minor importance.

Except for DFTB3,[47] the low-energy saddle points of the BO surface were determined for each method, defining activation energies $\Delta E^{\ddagger}$ and activation free energies $\Delta G^{\ddagger}$. The latter are exploited in TST approaches to estimate reaction rates. The $\Delta G^{\ddagger}$ are determined from a quasi-harmonic approximation of the BO surface at either state (stable or saddle point) followed by a corresponding evaluation of the partition sums of the approximated vibrational Hamiltonian, duly accounting for zero-point energies.[48] A $\Delta G^{\ddagger}$ is thus determined from just two points on the BO surface, and it will be interesting to compare them with the barrier heights inferred from the potential of mean force profiles $F(X)$ to be acquired for DFTB3 with umbrella sampling. HF and PM3 as its semiempirical approximation lead to the highest activation (free) energies

DOI: 10.1021/acs.jctc.7b01120
J. Chem. Theory Comput. 2018, 14, 2042−2051

$\Delta E^{\ddagger}$, $\Delta G^{\ddagger}$. For DFTB3 we tentatively report in Tables 1 and 2 SP calculations on PBE optimized structures (the DFTB3 Hamiltonian derives from the PBE exchange correlation functional). Whereas for AB there appears to be only a small difference between PBE and PBE/DFTB3, the results for pp-AB indicates that the activation (free) energies are expected to be somewhat smaller for DFTB3. Further, both DFTB3 and PBE underestimate the B3LYP activation (free) energies somewhat. In general we observe a reduced barrier height of push−pull azobenzene (pp-AB) compared to pure azobenzene (AB) for all levels of theory; the introduction of different PCM solvents in the B3LYP case reduces the barrier for pp-AB even further, with $\Delta G^{\ddagger}$(vacuum) > $\Delta G^{\ddagger}$(toluene) > $\Delta G^{\ddagger}$(DMSO), compare Table 2. Note also that in DMSO the character of the transition state (saddle point) structure has changed: from an inversion-like geometry (inversion center N closer to the NO_2-carrying ring) to a rotation-like one in agreement with earlier studies.[8] In all other cases, each QM method (including the semiempirical) predicts the inversion structure with N′═N−C1 close to 180°. With emphasis on the solvent effects just outlined, the *cis*−*trans* isomerization relaxation will now be explicitly modeled with our QM/MM sampling scheme.

3.2. Order Parameters for *Cis*-to-*Trans* Isomerization.

In general, the isomerization of azobenzene derivatives involves changes in the torsion angle ω and the bending angle α occurring simultaneously (recall Figure 1 for the definition of variables). To obtain a single order parameter, we change coordinates by setting

$$X = \sin(\alpha)\cos(\omega), \qquad Y = \sin(\alpha)\sin(\omega) \tag{1}$$

To illustrate the effect of this transformation, we consider the example of pp-AB in vacuum and acquire a PMF map in the variable pair (ω,α) using well-tempered metadynamics during a 10 ns long *ab initio* simulation at 333 K, see Figure 3a. By inspecting the PMF landscape, it is evident that the isomerization should first proceed with an increase in α in order to pass the barrier in the low energy region of the central ridge at $\omega \approx 90°$ (the second, equivalent pathway goes to *trans* from 0° through −90°); the transition state for pp-AB is thus of the (rotation assisted) inversion type. In Figure 3b the same data set is shown but now as a function of (X,Y). In these coordinates, the inversion pathway now runs parallel to the X axis at small positive Y, crossing the lines (hypersurfaces) X = const, which provide the desired foliation between *cis* and *trans* including a well-defined barrier top at $X \simeq 0$. We now determine the potential of mean force $F(X)$ in order to obtain some relation to trends in the (free) energy of activation reported in Tables 1 and 2, but also to better motivate the choice of the shooting interval I_{shoot} for the TPS procedure.

To acquire $F(X)$ we employ umbrella sampling with overlapping distributions in X from a set of 30, 5 ns long restrained simulations. Applying weighted histogram analysis (WHAM)[49,50] we obtain the reconstructed curves displayed in Figure 3c. To rationalize their appearance we have superimposed in Figure 3b ensembles of points corresponding to snapshots of restrained simulations in selected umbrella windows. When viewed in the original (α,ω) chart (Figure 3a), the ensemble at X = −0.06 spreads across a broad region close to α = 180° where the transition state for pp-AB is expected. The ensembles at X = 0.84 and X = −0.9 are well situated in the low-energy basins of *cis* and *trans*, respectively. With respect to $X \to \pm 1$, these are the outermost reference points for the umbrella potential to produce reasonable

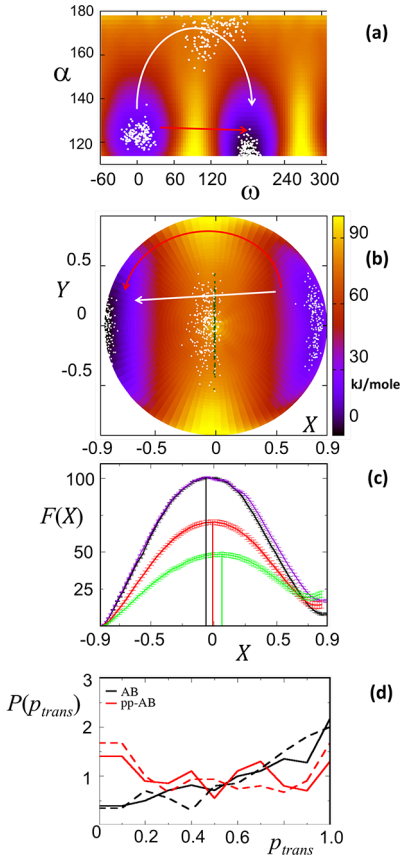


Figure 3. PMF map of pp-AB (a) from a metadynamics simulation in (ω,α) and (b) transformed into (X,Y) coordinates according to eq 1. The white arrows indicate the approximate course of an inversion-type isomerization pathway, red arrows a rotation-like trajectory. The three sets of scattered dots represent snapshots of MD runs with the molecule restrained in selected umbrella windows for X = −0.9 (*trans*), X = +0.84(*cis*), and the transition state region (X = −0.06). The green set of dots in (b) corresponds to configurations taken when crossing the barrier top at $X \approx -0.01$. (c) Potential of mean force $F(X)$ (in kJ/mol) inferred from umbrella sampling for AB in the gas phase (black), AB in DMSO (purple), pp-AB in the gas phase (red), and pp-AB in DMSO (green). The solid vertical lines indicate the respective location X^* of the barrier top in this coordinate. (d) Committor distributions $P(p_{trans})$ for pp-AB (red) and AB (black) in gas phase (solid line) and in DMSO (dashed line) at T = 333 K.

distributions, because the modulus of X (or $X^2 + Y^2$) cannot exceed unity. We find that the extent to which $F(X)$ can reliably be reconstructed is restricted to the interval [−0.9,0.9]; see also Supporting Information, Figure 1. This also limits the extent to which the *cis* and *trans* minima can be resolved in X. Consider the circular boundary $X^2 + Y^2 = 0.9$ of the PMF map in Figure 3b representing the lower edge of the (α, ω) window in Figure 3a at $\alpha \simeq 115°$. If we compare the locations of the two ensembles at *cis* and *trans*, we find the one at X = 0.84 well centered within the *cis* basin roughly half of it is recovered with the profile of $F(X)$ in Figure 3c, with the minimum at X_{cis} = 0.84. In the *trans* minimum α is somewhat smaller and the ensemble at X = −0.9 can only partially cover the low energy region. Consequently, the corresponding minimum at negative X in Figure 3c is not fully resolved and we get only a rough measure for the PMF difference ΔF_{ct} between *cis* and *trans*. The latter then is represented by $F(X_{cis})$, because X = −0.9 is chosen as the reference point for the WHAM procedure ($F(-0.9) = 0$

DOI: 10.1021/acs.jctc.7b01120
J. Chem. Theory Comput. 2018, 14, 2042−2051

Table 3. Locations X^* and X_{cis} As Well As the Corresponding Barrier Heights ΔF^* and Estimated Differences ΔF_{ct} between *Cis* and *Trans* for AB and pp-AB in Vacuum and DMSO Corresponding to the PMF Curves in Figure 3[a]

compound	X^*	X_{cis}	ΔF^*	ΔF_{ct}	$\langle\omega\rangle$	$\langle\alpha\rangle$
AB (vac)	−0.06	0.84	92.5(1.6)	7.9(0.9)	126.1(28.6)	171.3(4.6)
AB (DMSO)	−0.06	0.84	83.8(1.4)	17.0(0.9)	126.1(28.6)	170.7(5.2)
pp-AB (vac)	−0.01	0.84	56.3(4.0)	14.1(2.5)	85.9(34.4)	169.6(6.9)
pp-AB (DMSO)	0.07	0.72	29.2(3.6)	19.3(2.2)	68.8(22.9)	166.7(9.2)

[a]The average values $\langle\omega\rangle$ for the torsion around N′═N and $\langle\alpha\rangle$ for the bending angle at the inversion center refer to the restraint simulations at the respective barrier tops X^*. Free energies are in kJ/mol, and angles are given in degrees.

for all PMF curves). By contrast, the barrier height $\Delta F^* = F(X^*) - F(X_{cis})$ is somewhat better defined; Table 3 summarizes the values for the observables related to $F(X)$. The results appear to reflect the behavior inferred from Tables 1 and 2. In vacuum, the magnitude ΔF_{ct} compares well to ΔE_{ct}(DFTB3), which is considerably smaller than that of the full DFT methods. Also the behavior of the free energies of activation pointed out above is reproduced: we observe a lowering of ΔF^* in going from AB to pp-AB in vacuum and the further reduction in the case of pp-AB in going from apolar to polar solvent.

Yet we should emphasize that ΔF^* as a difference in potential of mean force is conceptually different from $\Delta G^{\ddagger}$, as the latter is determined by evaluating the partition sums with respect to the local expansion of the BO surface at the *cis* state and the (lowest energy) saddle point. Moreover, from restraint simulations at X^* it can be concluded that in the vicinity of the barrier top many configurations are sampled that do not belong to the transition state ensemble as defined above (i.e., the distribution of commitment probabilities $P(p_{trans})$ shows a narrow peak around $p_{trans} = 0.5$). In particular we can show that the surface $X = X^*$ is not an *iso-committor surface*.[51] We consider system configurations located precisely at X^* forming the seam indicated in Figure 3b. It represents a total of 200 molecular geometries visited in simulations restrained at the barrier top, $X^* = -0.01$, and recorded at the times t_c when the trajectory $X(t_c) = X^*$ crosses the surface. Each of these configurations is obtained from an independent run to reduce correlation effects. The corresponding distribution of committor probabilities $P(p_{trans})$ is shown in Figure 3d and appears to be rather uniform; if the barrier top was representative of the transition state ensemble or mainly captured molecular configurations close to a saddle point, we would also expect a sharp peak in $P(p_{trans})$ around $p_{trans} = 0.5$. Indeed, we remark that the transition from inversion to rotation for pp-AB in DMSO that was predicted from a saddle point analysis cannot be reproduced by the restraint simulations at the barrier top, compare the averages $\langle\alpha\rangle$ in Table 3. We will now show that the unbiased TPS procedure does exhibit this change in isomerization pathway.

3.3. Results from TPS. The TPS, as explained in Methods, was implemented by us with a shell script, based on a generic algorithm (Scheme 3 in Bolhuis, Dellago, and Geissler[52]). We choose the interface as $I_{shoot} = [-0.5, 0.5]$; in this way the probability for new pathways to be accepted can be kept around 30%. *Cis* and *trans* are then still reasonably well defined as $X > 0.5$ and $X < -0.5$, respectively, compare the corresponding limits in (ω,α) variables summarized in Table 4. Note that with this choice of I_{shoot} both, inversion and also pure rotation type routes can be tracked equally well, the schematic course of both pathways is depicted in Figure 3a,b. We find that for $I_{shoot} = [-0.5, 0.5]$, 500 integration time steps (time step set to 0.5 fs) in the gas phase and 1000 steps in solvent are sufficient for most trajectories to leave the interface region and to commit to either *cis* or *trans*. Initial trajectories were produced by steered molecular dynamics starting from a randomly chosen snapshot of an equilibrium trajectory in *cis*, until arriving in *trans* (according to the definition in Table 4). The path statistics for all TPS runs, such as the number of shooting points, the number of accepted paths, and the acceptance ratio (the ratio of the number of accepted paths and the number of TPS shooting points), are summarized in Table 5.

Table 4. Definition of *Cis* and *Trans* Stable States for the TPS Simulations in Terms of X and the Angles (ω,α)[a]

stable state	X	α	ω	α_{max}	ω_{max}
cis	$X \geq 0.5$	$\alpha < 150$	$\omega < 30$	$\alpha_{max} < 150$	$\omega_{max} < 45$
trans	$X \leq -0.5$	$\alpha < 150$	$\omega > 140$	$\alpha_{max} < 140$	$\omega_{max} > 150$

[a]ω_{max} and α_{max} designate some upper/lower limit to the range in these variables that is typically visited in equilibrium simulations of a few 100 ps duration (at 333 K).

Table 5. Path Statistics for All TPS Ensembles Acquired in This Work: (a) Fraction of Accepted Pathways, (b) Total Number of Accepted Pathways, and (c) Total Number of Shooting Points from Which Attempts Were Made

	AB			pp-AB		
	vacuum	toluene	DMSO	vacuum	toluene	DMSO
(a)	0.40	0.51	0.52	0.42	0.46	0.29
(b)	1196	1074	1540	2077	1100	865
(c)	3000	2093	2982	5000	2415	3000

We should reemphasize that the choice of the order parameter does not bias the reactive trajectory, its purpose is solely to distinguish between arrival/departure in/from the reactant/product state. Nevertheless, convergence of the TPS procedure can be an issue if there are multiple reaction channels separated by high free energy barriers. Indeed, we find empirically that during steering usually an inversion type pathway is produced in both, vacuum and solvent as just discussed along with the analysis of restraint simulations at the barrier top. We have therefore tested explicitly the impact of the initial choice on the convergence of TPS in the case of pp-AB in vacuum. Figure 4 shows a path density plot in (ω,α) space, compiled from a TPS run in which the initial pathway was produced by steered MD in ω (white dashed line) that exhibits only minor variations in α. For comparison, we show an analogous initial trajectory for steering in X (black dashed line). A path density plot emphasizes the occurrence of pathways over pure densities (the frequency at which a certain point is visited) by increasing the counts by +1 in each

DOI: 10.1021/acs.jctc.7b01120
J. Chem. Theory Comput. 2018, 14, 2042−2051

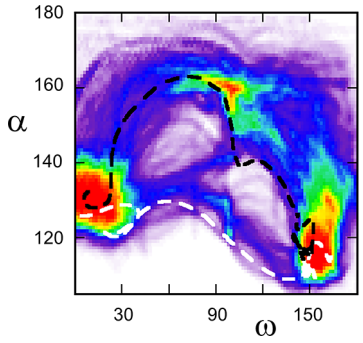


Figure 4. Path density plot of the convergence test of the TPS sampling for pp-AB in vacuum. The data comprise 1473 accepted pathways harvested from 5000 shooting points. The initial steered trajectory in ω is indicated by the white dashed line. The black dashed line corresponds to a trajectory steered in X, i.e., the initial steered trajectory for the TPS run in Figure 5a. The density shown in the heatmap represents the scaled number of counts within bins of size 1.8° × 1.8° (red; 1.0, maximum count; white, 0.0, 0 counts).

consecutive pair of bins visited along a trajectory $(\omega(t), \alpha(t))$. In the course of the path sampling, three successive channels develop along which reactive trajectories can be found with high probability. The rotation-like channel lies close to the steered trajectory in ω, and quickly branches off to an intermediate channel at a very early stage in the TPS run. The path ensemble then spreads out into the broad region enclosing the steered path in X, thus indicating convergence to the inversion route. Figure 5a shows the path density plot of a

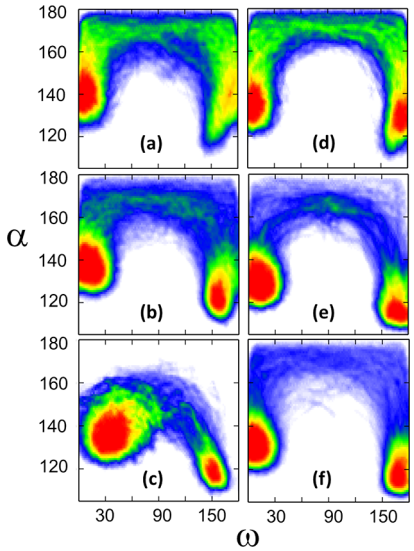


Figure 5. Path density plots projected onto (ω,α) for pp-AB (left column) and AB (right), in varying solvent conditions: vacuum (a,d), toluene (b,e), and DMSO (c,f). Bin size of histogram is 1.8° × 1.8°.

corresponding production run for pp-AB in vacuum, started with the initial trajectory steered in X as shown in Figure 4. The inversion route is now clearly established: in (ω,α) space the pathway appears almost rectangular with initial and final variation of the most probable paths involving only α. In toluene the character of the conversion of both systems remains inversion-like, in accord with the results of the PCM model for toluene applied to B3LYP, $\alpha \simeq 179°$, see Tables 1 and 2. However, in both cases the addition of small rotation-like components upon explicit solvation can be observed broadening the overall distribution of the pathways as compared to the vacuum case (Figure 5a,d).

For AB, however, there is no further shift with respect to DMSO whereas for pp-AB, a clear transition from inversion to rotation can now be observed. This agrees with the TST calculations employing implicit polarizable continuum model (PCM).[8]

Note the slight squeezing in ω of the path density distribution for pp-AB in going from vacuum/toluene to DMSO, Figure 5a,b and Figure 5c, respectively. We recall that in toluene and DMSO the number of integration time steps is doubled (from 500 to 1000) in anticipation of the isomerization to proceed under displacement and rearrangement of solvent molecules, and Figure 5c appears to indicate that the full transition to stable state values for ω and α is not yet complete (for pp-AB, the equilibrium values corresponding to Figure 5a−c for ω/α are all within 11 ± 6/123 ± 4 for *cis*, and within 173 ± 5/115 ± 3 for *trans*, respectively; compare also Table 1 in the Supporting Information). This does not, however, imply a slow-down of the dynamics related to barrier crossing: measuring the time spend to traverse the interface region (according to the definition in Table 4) we find a value of ∼100 fs quite independent of whether explicit solvent is present or not; the remaining time is spend outside I_{shoot} in a stationary fashion (represented primarily by the red regions of high path density in Figure 5). Obvious candidates for impacting molecular motion beyond the interface region are the rather bulky phenyl rings, and we shall conclude this work with a first look at their dynamics to be characterized by their relative orientation depicted in Figure 1b.

In addition to ω and α, ϕ_{OR} is also notably governed by the phenyl ring torsions ϕ and ϕ'. The saddle point analysis of AB at the CASSCF level,[12] for example, predicts $\phi_{OR} \approx 90°$ as the orientation for the low energy saddle point. The structure resembles the invertomer shown in Figure 1c-*i*; but also the "planar invertomer" with $\phi_{OR} = 180°$ exists at higher energies. For our cases we find the perpendicular invertomer in vacuum ($\phi_{OR} \simeq 106°(-85°)$) for full DFT at the B3LYP(PBE) level in the case of AB; a similar result is obtained for pp-AB ($\phi_{OR} \simeq -97°(-97°)$). The path density plots for AB and pp-AB in vacuum are shown in Figure 6a,b as a function of X and ϕ_{OR}. The vast majority of pathways assumes a fixed characteristic relative orientation of the phenyl rings, which varies only little along the trajectory (see the sample trajectories highlighted with dashed lines and also Supporting Information, Figure 2). Surprisingly, pp-AB appears to represent the perpendicular invertomer whereas AB settles on the planar geometry. Figure 6c shows the evolution of ϕ_{OR} during the progress of the TPS procedure, for a selected subset of the transition state ensemble (points with committor probability $p_{trans} \approx 0.5$). The subset has been chosen as follows: to determine the transition state for a given reactive pathway, one usually has to trace out the trajectory in short time intervals and calculate p_{trans} after each step to successively locate the snapshot with $p_{trans} = 0.5$. To bypass this computationally very expensive process, we randomly picked 332 accepted trajectories for pp-AB and 217 for AB, and for each path a random snapshot within the interval I_{shoot}, for which p_{trans} was determined and counted as a transition state if $0.4 < p_{trans} < 0.6$. The values ϕ_{OR} for these points are plotted in Figure 6a,b as scattered dots, and in Figure 6c as a function of the number variable i_{sel} representing the order by which they appear in the TPS iterations (not the

DOI: 10.1021/acs.jctc.7b01120
J. Chem. Theory Comput. 2018, 14, 2042−2051

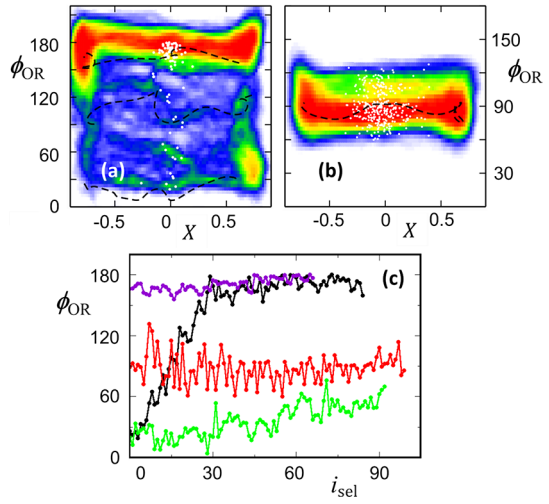


Figure 6. (a) Path density plot of ϕ_{OR} (AB) for the TPS run in vacuum projected onto (ϕ_{OR},X) space; (b) same for pp-AB. Dashed black lines: characteristic transition pathways; scattered dots: subset of the TSE with configurations for which $p_{trans} \approx 0.5$ (see text for explanation). (c) Subsets of TSEs plotted according to the order of iterations; black circles, TSE subset in (a), red circles correspond to (b). In addition, the corresponding selections for AB in DMSO (purple) and pp-AB in DMSO (green) are shown.

absolute iteration number). The TPS runs for AB in vacuum and for pp-AB in DMSO both started out with $\phi_{OR} \approx 10°$ from the initially steered path, yet they are seen to gradually approach the levels of ϕ_{OR} for AB in DMSO and pp-AB in vacuum, respectively. So, at the later stages of the TPS runs of Figure 6 there is only little variation along the complete reactive pathways. This can be thought to be conveyed by a small hula-twist component (see Figure 3 in the Supporting Information). Figure 6a,b indicates the extent of the TSE in the variable X. This is complementary to Figure 3d, which shows that an ensemble constrained at the barrier top X^* does not only contain transition states; now we observe that the TSE forms bands that are smaller than the interface region defined by I_{shoot}, which is in turn contained well within the range where ϕ_{OR} is approximately constant. This could imply that a characteristic value of ϕ_{OR} is not restricted to the transition state ensemble but rather typical of high energy conformations, a conjecture which is further supported by simulations restrained at X^* (comprising a majority of nontransition states; see Figure 4 in the Supporting Information). In other words, for the rather short transition path segments considered in this work the variable ϕ_{OR} approximately assumes a constant value; for pp-AB in the (equilibrium) *trans* state, the perpendicular arrangement appears only as a second, elevated minimum (see Figure 5 in the Supporting Information) indicating explicitly that the system could not completely settle in its stable state.

4. SUMMARY, CONCLUSIONS, AND OUTLOOK

In this paper we investigated the thermal *cis* → *trans* isomerization for azobenzene and a push−pull derivative in the gas phase as well as in two types of explicit solvents, employing a QM/MM description in combination with the transition path sampling approach. The semiempirical self-consistent density functional tight binding scheme DFTB3 as the QM method (coded into GROMACS) makes the TPS procedure quite efficient if the sampling of reactive trajectories is restricted to high energy regions. We have demonstrated that this was sufficient to obtain detailed insight into the nature of the transition, the change in isomerization mechanism of pp-AB in going from apolar to polar solvent, and the broadening of the transition path ensemble upon explicit solvation for both derivatives.

Although the strength of the approach and its future developments should clearly be devoted to studying complex environments, the TPS/QM/MM approach might also prove useful as a supporting technique to TST in order to provide additional insight when the latter scheme runs into conceptual problems. For instance, one recurrent artifact of TST in the case of azobenzenes[6] is the large discrepancy between theory and experiment with respect to the magnitude of the computed prefactors in the Arrhenius expression $k = A \exp(\Delta E^{\ddagger}/k_B T)$ for the rate constant k; experimentalists use this expression to infer A from the temperature dependence of k. In TST A is related to the entropy of activation $\Delta S^{\ddagger}$, and thus to the density of states contributing to the reaction, but only in the vicinity of the saddle point. The inspection of the transition state ensemble on an extended portion of the iso-committor surface could provide a useful means to probe this approximation. In using X as an order parameter for TPS sampling, we have seen that it does not exhibit an iso-committor property but may be considered a good first approximation, which can be improved by including more variables. With our numerical setup in its present form a systematic search process could straightforwardly be carried out using the aimless-shooting approach of Peters and Trout[53] with a subsequent likelihood maximization for the best combination of selected variables.[25,54]

Furthermore, in a preliminary investigation, we found that using only X (or ω, in the case of a purely rotational pathway) as order parameter, we obtain with a reactive-flux calculation transmission factors that range between 0.2 and 0.4 for the systems of this study, such that the actual estimation of relaxation rates becomes feasible.[16,55] Here, however, an improved global definition of the order parameter/reaction coordinate is desirable to provide more accurate accounts of free energy differences between the barrier top and the stable states. Yet also path sampling strategies could be employed for rate calculations directly. In 2003, van Erp, Moroni, and Bolhuis suggested a novel TPS scheme termed transition interface sampling (TIS),[56] by which the computational effort of obtaining rate constants could be considerably reduced compared to the original prescriptions.[21] For carrying out TIS, the order parameter should conveniently provide a foliation between the two stable states as the coordinate X does. In addition, TIS does not require fixed-length path ensembles, and this would be advantageous in the present case as the trajectories spend a lot of time in the complement of I_{shoot}. The optimization of local (in the transition state region) and global (for computing free energy differences) order parameters as well as alternative path sampling schemes are the subjects of our current investigations and will be published elsewhere.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.7b01120.

Details on all numerical procedures used in this work; force field parameters for MM models of toluene, DMSO, and azo derivatives (PDF)

DOI: 10.1021/acs.jctc.7b01120
J. Chem. Theory Comput. 2018, 14, 2042−2051

AUTHOR INFORMATION

Corresponding Author

*E-mail: mark.santer@mpikg.mpg.de. Phone: +49(0) 3315679610. Fax: +49(0)3315679602.

ORCID

Mark Santer: 0000-0002-5545-5155

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors thank Gerritt Groenhof and Tomas Kubar for technical support about the QM/MM implementation used in this work. We thank Ana Vila Verde for many stimulating and insightful discussions in the early stages of this project. Marko Wehle supervised the charge parametrization of all MM models used herein. This work was funded by the *International Max Planck Research School (IMPRS) for Multiscale Biosystems*.

REFERENCES

(1) Zakrevskyy, Y.; Richter, M.; Zakrevska, S.; Lomadze, N.; von Klitzing, R.; Santer, S. Stimuli-Responsive Materials: Light-Controlled Reversibel Manipulation of Microgel Particle Size Using Azobenzene-Containing Surfactant. *Adv. Funct. Mater.* **2012**, *22*, 5000−5009.

(2) Zakrevskyy, Y.; Roxlau, J.; Brezesinski, G.; Lomadze, N.; Santer, S. Photosensitive surfactants: Micellization and interaction with DNA. *J. Chem. Phys.* **2014**, *140*, 044906.

(3) Geoffroy, M.; Faure, D.; Oda, R.; Bassani, D. M.; Baigl, D. Photocontrol of Genomic DNA Conformation by Using a Photosensitive Gemini Surfactant: Binding Affinity versus Reversibility. *ChemBioChem* **2008**, *9*, 2382−2385.

(4) Le Ny, A.-L. M.; Lee, C. T. Photoreversible DNA Condensation Using Light-Responsive Surfactants. *J. Am. Chem. Soc.* **2006**, *128*, 6400−6408.

(5) Ponader, D.; Igde, S.; Wehle, M.; Märker, K.; Santer, M.; Bléger, D.; Hartmann, L. Photoswitchable precision glycooligomers and their lectin binding. *Beilstein J. Org. Chem.* **2014**, *10*, 1603−1612.

(6) Knie, C.; Utecht, M.; Zhao, F.; Kulla, H.; Kovalenko, S.; Brouwer, A. M.; Saalfrank, P.; Hecht, S.; Bléger, D. *ortho*-Fluoroazobenzenes: Visible Light Switches with Very Long-Lived *Z* Isomers. *Chem. - Eur. J.* **2014**, *20*, 16492−16501.

(7) Nishimura, N.; Kosako, S.; Sueishi, Y. Yoshimi, The Thermal Isomerization of Azobenzenes. III. Substituent, Solvent and Pressure Effects on the Thermal Isomerization of Push-pull Azobenzenes. *Bull. Chem. Soc. Jpn.* **1984**, *57*, 1617−1625.

(8) Dokić, J.; Gothe, M.; Wirth, J.; Peters, M. V.; Schwarz, J.; Hecht, S.; Saalfrank, P. Quantum Chemical Investigation of Thermal Cis-to-Trans Isomerization of Azobenzene Derivatives: Substituent Effects, Solvent Effects, and Comparison to Experimental Data. *J. Phys. Chem. A* **2009**, *113*, 6763−6773.

(9) Böckmann, M.; Doltsinis, N. L.; Marx, D. Nonadiabatic Hybrid Quantum and Molecular Mechanic Simulations of Azobenzene Photoswitching in Bulk Liquid Environment. *J. Phys. Chem. A* **2010**, *114*, 745−754.

(10) Hofmann, H. J.; Cimiraglia, R.; Tomasi, J. A Conformational Basis for the Description of the Thermal E/Z Isomerization of Aromatic Azo and Azomethine compounds. *J. Mol. Struct.: THEOCHEM* **1987**, *152*, 19−33.

(11) Cimiraglia, R.; Hofmann, H.-J. Rotation and inversion states in thermal E/Z isomerization of aromatic azo compounds. *Chem. Phys. Lett.* **1994**, *217*, 430−435.

(12) Cattaneo, P.; Persico, M. An ab-initio study of the photochemistry of azobenzene. *Phys. Chem. Chem. Phys.* **1999**, *1*, 4739−4743.

(13) Quick, M.; Dobryakov, A. L.; Gerecke, M.; Richter, C.; Berndt, F.; Ioffe, I. N.; Granovsky, A.; Mahrwald, R.; Ernsting, N.; Kovalenko, S. A. Photoisomerization Dynamics and Pathways of trans- and cis-Azobenzene in Solution from Broadband Femtosecond Spectroscopies and Calculations. *J. Phys. Chem. B* **2014**, *118*, 8756−8771.

(14) Eyring, H. The activated complex and the absolute rate of chemical reactions. *Chem. Rev.* **1935**, *17*, 65−77.

(15) Hänggi, P.; Talkner, P.; Borkovec, M. Reaction-rate theory: fifty years after Kramers. *Rev. Mod. Phys.* **1990**, *62*, 251−341.

(16) Chandler, D. Statistical mechanics of isomerization dynamics in liquids and the transition state approximation. *J. Chem. Phys.* **1978**, *68*, 2959−2970.

(17) Frenkel, D.; Smit, B. *Understanding Molecular Simulation*, 2nd ed.; Academic Press: New York, 2002; Chapter 16.4, p 450.

(18) Yu, X.; Wang, Z.; Buchholz, M.; Füllgrabe, N.; Grosjean, S.; Bebensee, F.; Bräse, S.; Wöll, C.; Heinke, L. *cis*-to-*trans* isomerization of azobenzene investigated by using thin films of metal−organic frameworks. *Phys. Chem. Chem. Phys.* **2015**, *17*, 22721−22725.

(19) Füchsel, G.; Klamroth, T.; Dokić, J.; Saalfrank, P. On the Electronic Structure of Neutral and Ionic Azobenzenes and Their Possible Roles as Surface Mounted Molecular Switches. *J. Phys. Chem. B* **2006**, *110*, 16337−16345.

(20) Geissler, P. L.; Dellago, C.; Chandler, D. Kinetic Pathways of Ion Pair Dissociation in Water. *J. Phys. Chem. B* **1999**, *103*, 3706−3710.

(21) Dellago, C.; Bolhuis, P. G.; Csajka, F. S.; Chandler, D. Transition path sampling and the calculation of rate constants. *J. Chem. Phys.* **1998**, *108*, 1964−1977.

(22) Dellago, C.; Bolhuis, P. G.; Chandler, D. On the calculation of reaction rate constants in the transition path ensemble. *J. Chem. Phys.* **1999**, *110*, 6617−6625.

(23) Basner, J. E.; Schwartz, S. D. How Enzyme Dynamics Helps Catalyze a Reaction in Atomic Detail: A Transition Path Sampling Study. *J. Am. Chem. Soc.* **2005**, *127*, 13822−13831.

(24) Quaytman, S. L.; Schwartz, S. D. Reaction coordinate of an enzymatic reaction revealed by transition path sampling. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 12253−12258.

(25) Knott, B. C.; Momeni, M. H.; Crowley, M. F.; Mackenzie, L. F.; Götz, A. W.; Sandgren, M.; Withers, S. G.; Stahlberg, J.; Beckham, G. T. The Mechanism of Cellulose Hydrolysis by a Two-Step, Retaining Cellobiohydrolase Elucidated by Structural and Transition Path Sampling Studies. *J. Am. Chem. Soc.* **2014**, *136*, 321−329.

(26) Frisch, M.; Trucks, G.; Schlegel, H.; Scuseria, G.; Robb, M.; Cheeseman, J.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K.; Burant, J.; Millam, J.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H.; Cross, J.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R.; Yazyev, O.; Austin, A.; Cammi, R.; Pomelli, C.; Ochterski, J.; Ayala, P.; Morokuma, K.; Voth, G.; Salvador, P.; Dannenberg, J.; Zakrzewski, V.; Dapprich, S.; Daniels, A.; Strain, M.; Farkas, O.; Malick, D. K.; Rabuck, A.; Raghavachari, K.; Foresman, J.; Ortiz, J.; Cui, Q.; Baboul, A.; Clifford, S.; Cioslowski, J.; Stefanov, B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R.; Fox, D.; Keith, T.; Al-Laham, M.; Peng, C.; Nanayakkara, A.; Challacombe, M.; Gill, P.; Johnson, B.; Chen, W.; Wong, M.; Gonzalez, C.; Pople, J. *Gaussian 03*, Revision C.02; Gaussian Inc.: Wallingford, CT, 2004.

(27) Field, M. J.; Bash, P. A.; Karplus, M. A Combined Quantum Mechanical and Molecular Mechanical Potential for Molecular Dynamics Simulations. *J. Comput. Chem.* **1990**, *11*, 700−733.

(28) Gaus, M.; Cui, Q.; Elstner, M. DFTB3: Extension of the Self-Consistent-Charge Density-Functional Tight-Binding Method (SCC-DFTB). *J. Chem. Theory Comput.* **2011**, *7*, 931−948.

(29) Gaus, M.; Goez, A.; Elstner, M. Parametrization and Benchmark of DFTB3 for Organic Molecules. *J. Chem. Theory Comput.* **2013**, *9*, 338−354.

(30) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, Th.; Suhai, S.; Seifert, G. Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1998**, *58*, 7260−7268.

DOI: 10.1021/acs.jctc.7b01120
J. Chem. Theory Comput. 2018, 14, 2042−2051

(31) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(32) van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, Flexible and Free. *J. Comput. Chem.* **2005**, *26*, 1701−1719.
(33) Lindahl, E.; Hess, B.; van der Spoel, D. GROMACS 3.0: A package for molecular simulation and trajectory analysis. *J. Mol. Model.* **2001**, *7*, 306−317.
(34) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **1995**, *91*, 43−56.
(35) Kubar, T.; Welke, K.; Groenhof, G. New QM/MM Implementation of the DFTB3Method in Gromacs Package. *J. Comput. Chem.* **2015**, *36*, 1978−1989.
(36) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(37) Freindorf, M.; Shao, Y.; Furlani, T. R.; Kong, J. Lennard-Jones Parameters for the Combined QM/MM Method Using the B3LYP/6-31+G*/AMBER Potential. *J. Comput. Chem.* **2005**, *26*, 1270−1278.
(38) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993**, *97*, 10269−10280.
(39) Van Gunsteren, W. F.; Berendsen, H. J. C. A Leap-frog Algorithm for Stochastic Dynamics. *Mol. Simul.* **1988**, *1*, 173−185.
(40) Duchstein, P.; Neiss, C.; Görling, A.; Zahn, D. Molecular mechanics modeling of azobenzene-based photoswitches. *J. Mol. Model.* **2012**, *18*, 2479−2482.
(41) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED2: new feathers for an old bird. *Comput. Phys. Commun.* **2014**, *185*, 604.
(42) Dellago, C.; Bolhuis, P. G.; Chandler, D. Efficient transition path sampling: Application to Lennard-Jones cluster rearrangements. *J. Chem. Phys.* **1998**, *108*, 9236−9245.
(43) Truhlar, D. G.; Garrett, B. C. Variational Transition State Theory. *Acc. Chem. Res.* **1980**, *13*, 440−448.
(44) 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*, 154104.
(45) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum Mechanical Continuum Solvation Models. *Chem. Rev.* **2005**, *105*, 2999−3094.
(46) Rietze, C.; Titov, E.; Lindner, S.; Saalfrank, P. *J. Phys.: Condens. Matter* **2017**, *29*, 314002.
(47) At the time of writing, a saddle point search algorithm to be used with DFTB3 was not available to us; we are currently working on an implementation of a quasi-Newton scheme, the results of which could be compared to that of a corresponding calculation with Gaussian. A more complete exhibition of transition-state structures using DFTB3 will be published elsewhere.
(48) Cramer, C. J. *Essentials of Computational Chemistry*, 2nd ed.; Wiley: Weinheim, 2004; Chapter 15.
(49) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem.* **1992**, *13*, 1011−1021.
(50) Grossfield, A. *WHAM*: an implementation of the weighted histogram analysis method, Version 2.0.9; http://membrane.urmc.rochester.edu/content/wham/.
(51) Metzner, P.; Schütte, C.; Vanden-Eijnden, E. Illustration of transition path theory on collection of simple examples. *J. Chem. Phys.* **2006**, *125*, 084110.
(52) Dellago, C.; Bolhuis, P. G.; Geissler, P. L. Transition Path Sampling. *Adv. Chem. Phys.* **2003**, *123*, 1−78.
(53) Peters, B.; Trout, B. L. Obtaining reaction coordinates by likelihood maximization. *J. Chem. Phys.* **2006**, *125*, 054108.
(54) Vreede, J.; Juraszek, J.; Bolhuis, P. G. Prediting the reaction coordinates of millisecond light-induced conformational changes in photoactive yellow protein. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 2397−2402.
(55) Hummer, G. From transition paths to transions states and rate coefficients. *J. Chem. Phys.* **2004**, *120*, 516−523.
(56) van Erp, T. S.; Moroni, D.; Bolhuis, P. G. A novel path sampling method for the calculation of rate constants. *J. Chem. Phys.* **2003**, *118*, 7762−7774.

DOI: 10.1021/acs.jctc.7b01120
J. Chem. Theory Comput. 2018, 14, 2042−2051