

## Introduction to QM/MM Simulations

Gerrit Groenhof

### Abstract

Hybrid quantum mechanics/molecular mechanics (QM/MM) simulations have become a popular tool for investigating chemical reactions in condensed phases. In QM/MM methods, the region of the system in which the chemical process takes place is treated at an appropriate level of quantum chemistry theory, while the remainder is described by a molecular mechanics force field. Within this approach, chemical reactivity can be studied in large systems, such as enzymes. In the first part of this contribution, the basic methodology is briefly reviewed. The two most common approaches for partitioning the two subsystems are presented, followed by a discussion on the different ways of treating interactions between the subsystems. Special attention is given on how to deal with situations in which the boundary between the QM and MM subsystems runs through one or more chemical bonds. The second part of this contribution discusses what properties of larger system can be obtained within the QM/MM framework and how. Finally, as an example of a QM/MM application in practice, the third part presents an overview of recent QM/MM molecular dynamics simulations on photobiological systems. In addition to providing quantities that are experimentally accessible, such as structural intermediates, fluorescence lifetimes, quantum yields and spectra, the QM/MM simulations also provide information that is much more difficult to measure experimentally, such as reaction mechanisms and the influence of individual amino acid residues.

**Key word:** Quantum mechanics, Molecular mechanics, QM/MM, Molecular dynamics

---

### 1. Introduction

In this chapter we present a short introduction into the development and application of computational techniques for modelling chemical reactions in the condensed phase. We start by reviewing the basic concepts of these methods. We then discuss how these methods can be used in practical computations and what kind of information can be obtained. We conclude this chapter with a short review of an application on a photobiological system, for which the simulations not only revealed the detailed sequence of events that follow photon absorption but also demonstrate how the biological environment controls the photochemical reaction.

## 2. QM/MM: Theory and Implementation

The size and complexity of a typical biomolecular system, together with the timescales that must be reached, necessitate the use of classical molecular dynamics for the nuclear degrees of freedom. In molecular dynamics (MD) simulations, Newton's equations of motion are solved numerically to obtain a trajectory of the dynamics of a molecule over a period of time (1). To model electronic rearrangements during a chemical reaction, a quantum mechanical description (QM) is required for those parts of the system that are involved in the reaction. For the remainder, a simple molecular mechanics force field model suffices (MM). The interactions in the system are thus computed within a hybrid QM/MM framework.

### 2.1. Molecular Mechanics

Molecular dynamics simulations of biological systems have come of age (2). Since the first application of MD on a small protein in vacuum more than three decades ago (3), advances in computer power, algorithmic developments and improvements in the accuracy of the used interaction functions have established MD as an important and predictive technique to study dynamic processes at atomic resolution (4). In the interaction functions, the so-called molecular mechanics force field, simple chemical concepts are used to describe the potential energy of the system (1):

$$V_{\text{MM}} = \sum_i^{N_{\text{bonds}}} V_i^{\text{bond}} + \sum_j^{N_{\text{angles}}} V_j + \sum_l^{N_{\text{torsions}}} V_l^{\text{torsion}} + \sum_i^{N_{\text{MM}}} \sum_{j>i}^{N_{\text{MM}}} V_{ij}^{\text{Coul}} + \sum_i^{N_{\text{MM}}} \sum_{j>i}^{N_{\text{MM}}} V_{ij}^{\text{LJ}}, \quad (1)$$

where  $N_{\text{MM}}$  is the number of atoms in the system. Bonds and angles ( $V^{\text{bond}}$ ,  $V^{\text{angle}}$ ) are normally modelled by harmonic functions, and torsions by periodic functions ( $V^{\text{torsion}}$ ). The pairwise electrostatic interaction between atoms with a partial charge ( $Q_i$ ) is given by Coulomb's law:

$$V_{ij}^{\text{Coul}} = \frac{e^2 Q_i Q_j}{4\pi\epsilon_0 R_{ij}}, \quad (2)$$

in which  $R_{ij}$  denotes the interatomic distance,  $e$  the unit charge and  $\epsilon_0$  the dielectric constant. Van der Waals interactions, for example the combination of short-range Pauli repulsion and long-range dispersion attraction, are most often modelled by the Lennard-Jones potential:

$$V_{ij}^{\text{LJ}} = \left( \frac{C_{12}^{ij}}{R_{ij}} \right)^{12} - \left( \frac{C_6^{ij}}{R_{ij}} \right)^6, \quad (3)$$

with  $C_{12}^{ij}$  and  $C_6^{ij}$  a repulsion and attraction parameter, respectively, which depend on the atomtypes of the atoms  $i$  and  $j$ .

Electrons are thus ignored in molecular mechanics force fields. Their influence is expressed by empirical parameters that are valid for the ground state of a given covalent structure. Therefore, processes that involve electronic rearrangements, such as chemical reactions, cannot be described at the MM level. Instead, these processes require a quantum mechanics description of the electronic degrees of freedom. However, the computational demand for evaluating the electronic structure places severe constraints on the size of the system that can be studied.

## 2.2. Hybrid Quantum Mechanics/Molecular Mechanics Models

Most biochemical systems, such as enzymes, are too large to be described at any level of *ab initio* theory. At the same time, the available molecular mechanics force fields are not sufficiently flexible to model processes in which chemical bonds are broken or formed. To overcome the limitations of a full quantum mechanical description on the one hand, and a full molecular mechanics treatment on the other hand, methods have been developed that treat a small part of the system at the level of quantum chemistry (QM), while retaining the computationally cheaper force field (MM) for the larger part. This hybrid QM/MM strategy was originally introduced by Warshel and Levitt (5) and is illustrated in Fig. 1. The justification for dividing a system into regions that are described at different levels of theory is the local character of most chemical reactions in condensed phases. A distinction can usually be made between a “reaction centre” with atoms that are directly involved in the reaction and a “spectator” region, in which the atoms do not directly participate in the reaction. For example, most reactions in solution involve the reactants and the first few solvation shells. The bulk solvent is hardly affected by the reaction, but can influence the reaction via long-range interactions. The same is true for most enzymes, in which the catalytic process is restricted to an active site located somewhere inside the protein. The rest of the protein provides an electrostatic background that may or may not facilitate the reaction.

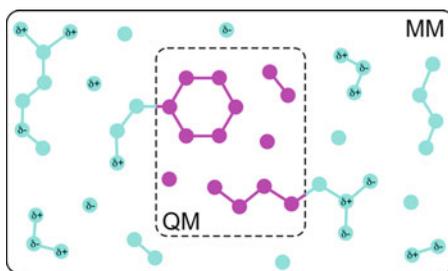


Fig. 1. Illustration of the QM/MM concept. A small region, in which a chemical reaction occurs and therefore cannot be described with a force field, is treated at a sufficiently high level of QM theory. The remainder of the system is modelled at the MM level.

The hybrid QM/MM potential energy contains three classes of interactions: interactions between atoms in the QM region, between atoms in the MM region and interactions between QM and MM atoms. The interactions within the QM and MM regions are relatively straightforward to describe, that is at the QM and MM level, respectively. The interactions between the two subsystems are more difficult to describe, and several approaches have been proposed. These approaches can be roughly divided into two categories: subtractive and additive coupling schemes.

### 2.3. Subtractive QM/MM Coupling

In the subtractive scheme, the QM/MM energy of the system is obtained in three steps. First, the energy of the total system, consisting of both QM and MM regions, is evaluated at the MM level. The QM energy of the isolated QM subsystem is added in the second step. Third, the MM energy of the QM subsystem is computed and subtracted. The last step corrects for including the interactions within the QM subsystem twice:

$$V_{\text{QM/MM}} = V_{\text{MM}}(\text{MM} + \text{QM}) + V_{\text{QM}}(\text{QM}) - V_{\text{MM}}(\text{QM}). \quad (4)$$

The terms QM and MM stand for the atoms in the QM and MM subsystems, respectively. The subscripts indicate the level of theory at which the potential energies ( $V$ ) are computed. The most widely used subtractive QM/MM scheme is the ONIOM method, developed by the Morokuma group (6, 7), and is illustrated in Fig. 2.

The main advantage of the subtractive QM/MM coupling scheme is that no communication is required between the quantum chemistry and molecular mechanics routines. This makes the implementation relatively straightforward. However, compared to the more advanced schemes that are discussed below, there are also disadvantages.

A major disadvantage is that a force field is required for the QM subsystem, which may not always be available. In addition, the force field needs to be sufficiently flexible to describe the effect of chemical changes when a reaction occurs.

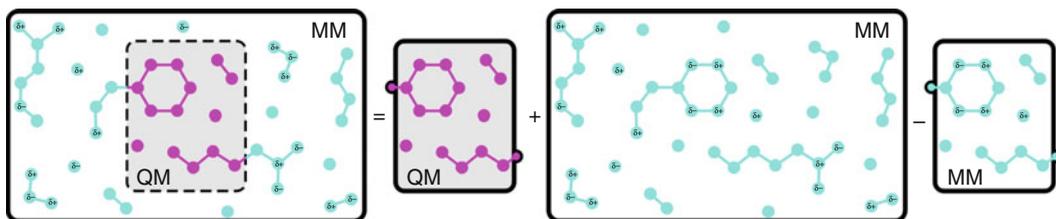


Fig. 2. Subtractive QM/MM coupling: The QM/MM energy of the total system (*left hand side of the equation*) is assumed to be equal to the energy of the isolated QM subsystem, evaluated at the QM level, plus the energy of the complete system evaluated at the MM level, minus the energy of the isolated QM subsystem, evaluated at the MM level. The last term is subtracted to correct for double counting of the contribution of the QM subsystem to the total energy. A prerequisite for the calculation is that a force field for the QM subsystem is available.

A further drawback of this method is absence of polarization of the QM electron density by the MM environment. This shortcoming can be particularly problematic for modelling biological charge transfer processes, since these are usually mediated by the protein environment. For a realistic description of such reactions a more consistent treatment of the interactions between the electrons and their surrounding environment is needed.

#### **2.4. Additive QM/MM Coupling**

In additive schemes, the QM system is embedded within the larger MM system, and the potential energy for the whole system is a sum of MM energy terms, QM energy terms and QM/MM coupling terms:

$$V_{\text{QM/MM}} = V_{\text{QM}}(\text{QM}) + V_{\text{MM}}(\text{MM}) + V_{\text{QM-MM}}(\text{QM} + \text{MM}). \quad (5)$$

Here, only the interactions within the MM region are described at the force field level,  $V_{\text{MM}}(\text{MM})$ . In contrast to the subtractive scheme, the interactions between the two subsystems are treated explicitly:  $V_{\text{QM-MM}}(\text{QM}+\text{MM})$ . These interactions can be described at various degrees of sophistication.

##### **2.4.1. Mechanical Embedding**

In the most basic approach, all interactions between the two subsystems are handled at the force field level. The QM subsystem is thus kept in place by MM interactions. This is illustrated in Fig. 3. Chemical bonds between QM and MM atoms are modelled by harmonic potentials ( $V^{\text{bond}}$ ), angles defined by one QM atom, and two MM atoms are described by the harmonic potential as well ( $V^{\text{angles}}$ ), while torsions involving at most two QM atoms are commonly modelled by a periodic potential function ( $V^{\text{torsion}}$ ). Non-bonded interactions, that is those between atoms separated by three or more atoms, are also modelled by force field terms: Van der Waals by the Lennard-Jones potential ( $V^{\text{LJ}}$ ) and electrostatics by the Coulomb potential ( $V^{\text{Coul}}$ ). In the most simple implementation of mechanical embedding, the electronic wave function is evaluated for an isolated QM subsystem. Therefore, the MM environment cannot induce polarization of the electron density in the QM region. For calculating the electrostatic interactions between the subsystems, one can either use a fixed set of charges for the QM region, for example, those given by the force field, or re-compute the partial charges on the QM atoms at every integration step of the simulation. In the second strategy, a least-squares fitting procedure is used to derive atomic charges that optimally reproduce the electrostatic potential at the surface of the QM subsystem (8, 9).

Lennard-Jones parameters are normally not updated. Therefore, problems may arise if during the simulation, changes occur in the chemical character of the atoms in the QM region, for example, in reactions that involve changes in the hybridization state of the

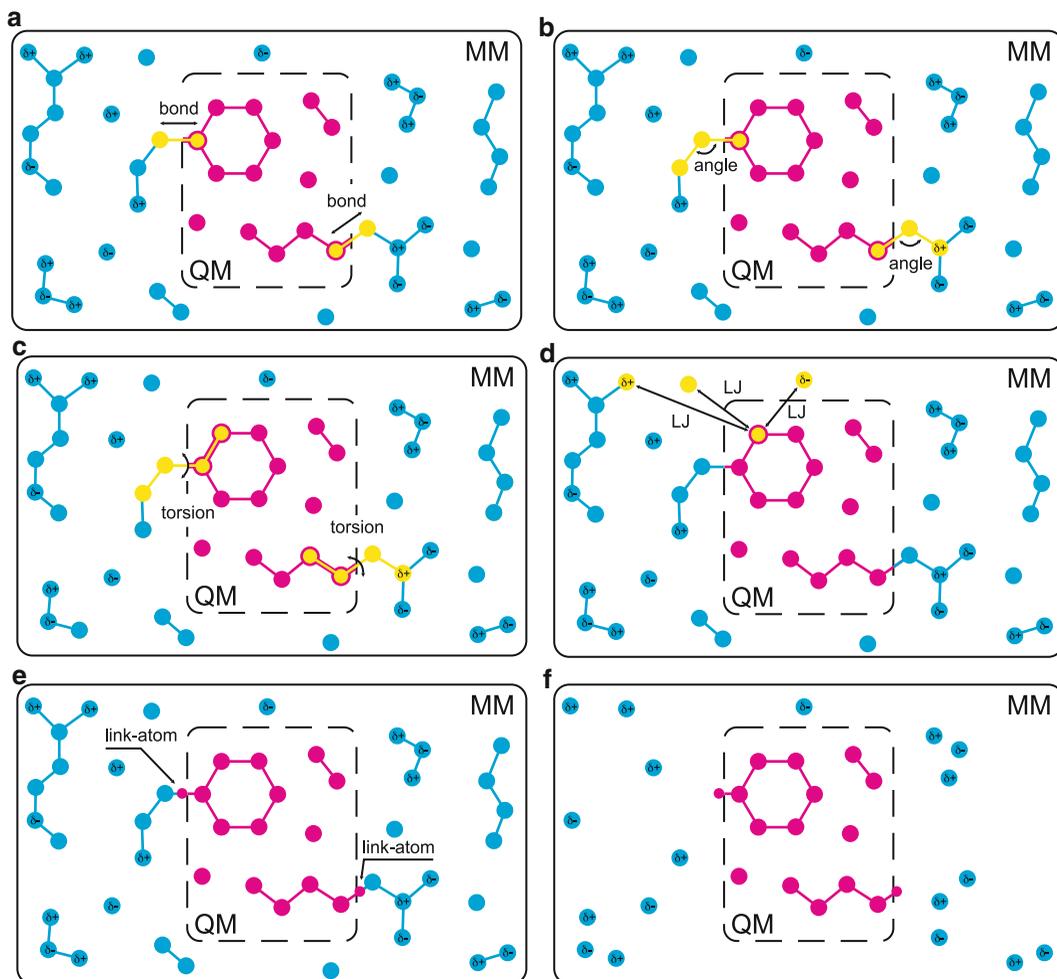


Fig. 3. Coupling between the QM and MM subsystems in the additive QM/MM schemes. The top panels (a)–(c) show bonded interactions between QM and MM atoms. These interactions are handled at the force field level (MM). Panel d shows the Van der Waals interactions between an atom in the QM region and three MM atoms. These interactions are modelled by the Lennard-Jones potential. Panel e illustrates the link atom concept. This atom caps the QM subsystem and is present only in the QM calculation. Panel f demonstrates how the electrostatic QM/MM interactions are handled. In the electrostatic embedding approach, the charged MM atoms enter the electronic Hamiltonian of the QM subsystem. In the mechanical embedding, partial MM charges are assigned to the QM atoms and the electrostatic interactions are computed by the pairwise Coulomb potential.

atoms. However, since the Lennard-Jones potential is a relatively short-ranged function, the error introduced by keeping the same parameters throughout the simulations is most likely not very large.

#### 2.4.2. Electrostatic Embedding

An improvement of mechanical embedding is to include polarization effects. In the electrostatic embedding scheme, the electrostatic interactions between the two subsystems are handled during the

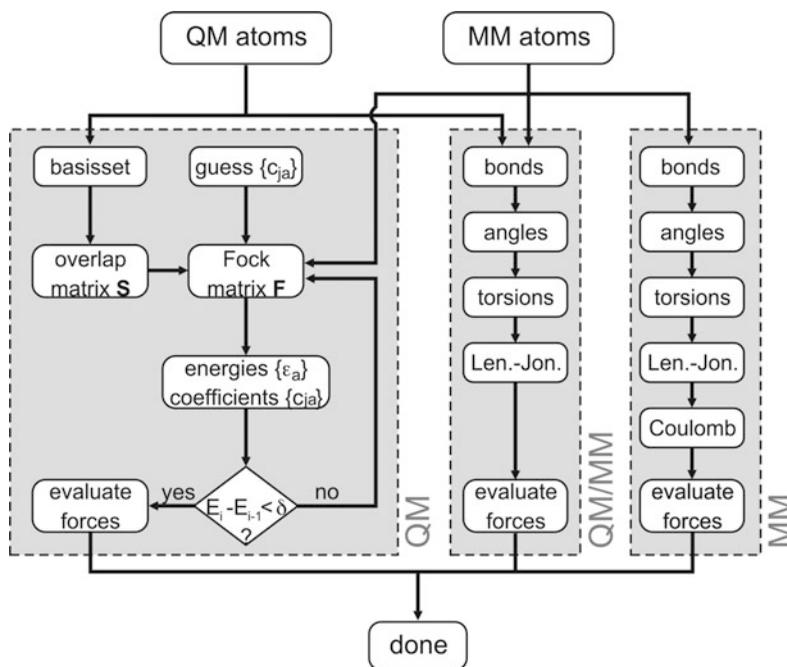


Fig. 4. Flow scheme of a QM/MM energy calculation within the electrostatic embedding scheme. Interactions between atoms in the MM subsystem are handled at the force field level (*third branch*). The QM atoms enter the self-consistent field routine, with the charged MM atoms included as point charges (*first branch*). Diagonalization of the augmented Fock matrix yields the energy of the QM atoms as well as the electrostatic interaction energy between the subsystems. All other interactions involving QM and MM atoms are described by the force field terms (*second branch*).

computation of the electronic wave function. The charged MM atoms enter the QM Hamiltonian as one-electron operators:

$$b_i^{\text{QM-MM}} = b_i^{\text{QM}} - \sum_J^M \frac{e^2 Q_J}{4\pi\epsilon_0 |\mathbf{r}_i - \mathbf{R}_J|}, \quad (6)$$

where  $\mathbf{r}_i$  and  $\mathbf{R}_J$  are the positions of electron  $i$  and MM atom  $J$ ,  $b_i^{\text{QM}}$  is the original one-electron operator for the kinetic and nuclear attraction energy of electron  $i$  (10) and  $M$  is the number of MM atoms that have a partial charge  $Q_J$ . Thus, the electrons “see” these MM atoms as special nuclei with non-integer and possibly negative charges. Since the electronic Hamiltonian contains extra terms, the electrostatic embedding model requires modifications of the quantum chemistry software. Martin Field and co-workers were among the first researchers to implement this scheme (11) and developed an interface between the molecular mechanics program, Charmm (12) and the *semi-empirical* quantum chemistry package Mopac (13). Figure 4 shows a schematic overview of how the QM and MM routines are

interconnected in a practical implementation of electrostatic embedding. In the electrostatic coupling approach, the MM atoms can polarize the electrons in QM subsystem. However, the atomic charges of the MM atoms have been parametrized to provide a realistic description of an MM system, rather than of a physically correct charge distribution. Therefore, the question arises whether polarization induced by these MM charges is realistic or not. In principle, one would need to re-derive the charges for use in QM/MM frameworks. In reality, interactions between the systems are not only due to electrostatics between charged atoms, but also due to polarization, exchange, charge transfer, dispersion and Pauli repulsion. In force fields, only the *combination* of atomic charges and Lennard-Jones parameters provides a reasonable description of all these effects taken together, albeit in an implicit manner. Part of the interaction due to polarization of the QM region is thus already accounted for by the Lennard-Jones potential. Therefore, not only the MM charges, but also the Lennard-Jones parameters would need to be reparametrized for use in electrostatic embedding QM/MM simulations. However, in practice this is hardly done, and most workers use default force field parameters.

A further problem that may arise when using standard MM atomic charges to describe the charge distribution in the MM system, is the risk of over-polarization near the boundary. The point charges on the MM side of the interface may attract (or repel) the electrons too strongly, which could lead to electron density spilling out into the MM region. Such artefacts can become serious if large flexible basis set (e.g., with polarization and diffuse functions), or plane waves are used in the QM calculations. The electron spill out can be avoided by using smeared-out charges instead of the traditional point charges (14). A convenient way for smearing the charges is to use a Gaussian distribution centred at the MM atom:

$$\Omega_J^{\text{MM}}(\mathbf{r}) = \sqrt{\frac{Q_J^{\text{MM}}}{\pi\alpha^3}} \exp\left[-\frac{|\mathbf{r} - \mathbf{R}_J|^2}{2\alpha^2}\right], \quad (7)$$

where  $|\Omega_J^{\text{MM}}(\mathbf{r})|^2$  is the charge density at position  $\mathbf{r}$ , due to MM atom  $J$  at position  $\mathbf{R}_J$  and charge  $Q_J$ . The factor  $\alpha$  controls the width of the distribution and is a parameter that needs to be calibrated. In contrast to the point charge model, the Coulomb interaction between the QM electrons and the Gaussian charge distributions does not diverge if the electrons approach the MM atoms:

$$h_{ij}(\mathbf{r}_1) = Q_J \int \phi_i^*(\mathbf{r}_1) \frac{\text{erf}(|\mathbf{r}_1 - \mathbf{R}_J|/\alpha)}{|\mathbf{r}_1 - \mathbf{R}_J|} \phi_j(\mathbf{r}_1) d\mathbf{r}_1, \quad (8)$$

with  $\phi_i$  the molecular orbital and  $h_{ij}$  the one-electron integral describing the interaction of a single electron with MM atom  $J$ . Such renormalization of the coulomb interaction avoids the

unphysical attraction of the electrons to charged atoms at the boundary between the two subsystems.

### 2.4.3. Polarization Embedding

The next step in increasing the level of sophistication is to include the polarizability of the MM atoms. In the polarization embedding scheme both regions can mutually polarize each other. Thus, not only is the QM region polarized by the MM atoms, the QM region can also induce polarization in the MM system. Different approaches have been developed to model polarization of MM atoms. Among the most popular methods are the charge-on-a-spring model (15), the induced dipole model (16) and the fluctuating charge model (17).

To obtain the total QM/MM energy in the polarizable embedding approach, the MM polarizations need to be computed at every step of the self-consistent-field iteration of the QM wave function. Since the polarization is computed in a self-consistent manner as well, the QM/MM computation can become very cumbersome and demanding. As a compromise, Zhang and co-workers have suggested to include polarization only in a small shell of MM atoms around the QM region (18).

Although polarization embedding offers the most realistic coupling between the QM and MM regions, polarizable force field for biomolecular simulations are not yet available. Therefore, despite progress in the development of such force fields, QM/MM studies with polarizable MM regions have so far been largely restricted to non-biological systems (19).

## 2.5. Capping Bonds at the QM/MM Boundary

If the QM and MM subsystems are connected by chemical bonds, care has to be taken when evaluating the QM wave function. A straightforward cut through the QM/MM bond creates one or more unpaired electrons in the QM subsystem. In reality, these electrons are paired in bonding orbitals with electrons belonging to the atom on the MM side. A number of approaches to remedy the artefact of such open valences have been proposed.

### 2.5.1. Link Atoms

The most easy solution is to introduce a monovalent link atom at an appropriate position along the bond vector between the QM and MM atoms (Figs. 3e and 5). Hydrogen is most often used, but there is no restriction on the type of the link atom and even complete fragments, such as methyl groups, can be used to cap the QM subsystem. The link atoms are present only in the QM calculation, and are invisible for the MM atoms. In principle each link atom introduces three additional degrees of freedom to the system. However, in practice the link atom is placed at a fixed position along the bond in every step of the simulation, so that these additional degrees of freedom are removed again. At each step, the force acting on the link atoms are distributed over the QM and MM atoms of the bond according to the lever rule.

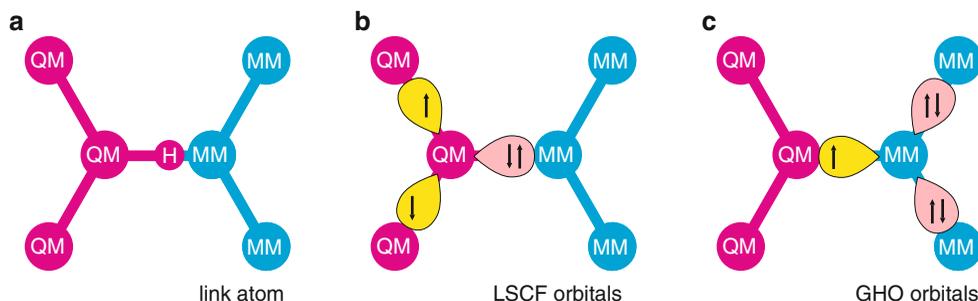


Fig. 5. Different approaches to cap the QM region: link atoms (a) and frozen orbitals (b,c). The hydrogen link atom (a) is placed at an appropriate distance along the QM/MM bond vector and is present only in the QM calculation. In the localized SCF method (b), a set of localized orbitals is placed on the QM atom. During the SCF iterations, the orbital pointing towards the MM atom is double-occupied and frozen, while the other orbitals are single-occupied and optimized. In the generalized hybrid orbital approach (c), a set of localized orbitals is placed on the MM atom. During the SCF interaction, the orbitals pointing towards the other MM atoms are double occupied and frozen, while the orbital pointing towards the QM atom is single-occupied and optimized.

### 2.5.2. Localized Orbitals

A popular alternative to the link atom scheme is to replace a chemical bond between the QM and MM subsystem by a double-occupied molecular orbital. This idea, which dates back to the pioneering work of Warshel and Levitt (5), assumes that the electronic structure of the bond is insensitive to changes in the QM region. The two most widely used approaches are the localized hybrid orbital method (20), which introduces orbitals at the QM atom (Fig. 5b), and the generalized hybrid orbital approach (21), which places additional orbitals on the MM atom (Fig. 5c).

In the localized self-consistent field (LSCF) method by Rivail and co-workers (20), the atomic orbitals on the QM atom of the broken bond are localized and hybridized. The hybrid orbital pointing towards the MM atom is occupied by two electrons. The other orbitals are each occupied by a single electron. During the SCF optimization of the QM wave function, the double-occupied orbital is kept frozen, while the other hybrid orbitals are optimized along with all orbitals in the QM region. The parameters in this method are the molecular orbital coefficients of the hybrid orbitals. In the original approach, these parameters are obtained by localising orbitals in smaller model systems. This procedure thus assumes that the electronic structure of a chemical bond is transferable between different systems.

Alternatively, the coefficients of the frozen orbital can be obtained by performing a single point QM/MM calculation with a slightly enlarged QM subsystem. Any further broken bonds between the larger QM subsystem and the MM region are capped by link atoms in this calculation. The advantage of this so-called frozen orbital approach (22) is that no assumption is made on the electronic structure of the chemical bond. The disadvantage is that an electronic structure calculation has to be performed on a larger QM subsystem.

In the generalized hybrid orbital approach (GHO) of Gao and co-workers, hybrid orbitals are placed on the MM atom of the broken bond (21). In contrast to the LSCF scheme, the orbital pointing to the QM atom is optimized, while the others are kept double-occupied and frozen (Fig. 5).

In all localized orbital approaches, one or more parametrization steps are required. For this reason, the link atom is still the most widely adopted procedure for capping the QM region. Furthermore, studies that compared the accuracy of both methods showed that there is little advantage in using a localized orbitals rather than link atoms (23, 24).

In addition to capping the QM subsystem, one also needs to be careful if the MM atom at the other side of the chemical bond is charged. Since this atom is very near the QM subsystem, artefacts can easily arise due to the over-polarization effect, as discussed above. The easiest way to avoid this problem is to set the charges of that MM atom to zero. Alternatively, the charge can be shifted to MM atoms further away from the bond. The latter solution keeps the overall charge of the system constant.

---

### 3. QM/MM Applications

#### 3.1. Molecular Dynamics Simulations

The QM/MM method provides both potential energies and forces. With these forces, it is possible to perform a molecular dynamics simulation. However, because of the computational efforts required to perform *ab initio* calculations, the timescales that can be reached in QM/MM simulations is rather limited. At the *ab initio* or DFT level, the limit is in the order of few hundreds of picoseconds. With *semi-empirical* methods (e.g., AM1 (25), PM3 (26, 27), or DFTB (28)) for the QM calculation, the limit is roughly 100 times longer. Therefore, unless the chemical process under consideration is at least an order of magnitude faster than the timescale that can be reached, an unrestrained MD simulation is not the method of choice to investigate that process. Although the lack of sampling can be overcome by using enhanced sampling techniques, most researchers rely on energy minimization techniques to study chemical reactivity in condensed phases.

#### 3.2. Geometry Optimization

The traditional approach to study reactivity on a computer has been to characterize stationary points on the potential energy surface of the isolated system. The minima are identified as reactants and products, whereas the lowest energy saddle points that connect these minima are interpreted as the transition states. Extending this approach to QM/MM potential energy surfaces, however, is difficult, due to the much higher dimensionality of a typical QM/MM system. Since there are many more degrees of freedom that have to be optimized,

the geometry optimizer needs to be very efficient. Furthermore, the number of local minima in high dimensional systems is usually very large. At temperatures above zero, many of these minima are populated and there are also many paths connecting them. Therefore, even if the optimization can be carried out successfully, it may be difficult, if not impossible, to characterize all reaction pathways that are relevant for the process under study (29).

Despite these problems, optimizing the stationary points on the QM/MM potential energy surface is often the first step in exploring the reaction pathway. It usually gives important insights into the mechanism of the reaction, and the way by which it is controlled by the environment.

### 3.3. Free Energy Computation

To understand reactivity, one rather needs the *free energy* surface of the process. Computing free energies requires sampling of the underlying potential energy surface to generate ensembles. In equilibrium, the free energy difference  $\Delta G_{A \rightarrow B}$  between the reactant state ( $A$ ) and the product state ( $B$ ), both defined as areas on the free energy landscape, is determined by their populations  $p$ :

$$\Delta G_{A \rightarrow B} = -k_B T \ln \frac{p_B}{p_A}, \quad (9)$$

with  $k_B$  the Boltzmann constant, and  $T$  the temperature. However, for chemical reactions, the barriers separating the states  $A$  and  $B$  are high, and transitions are rare events. Therefore, it is not likely that both states are sampled sufficiently in a normal MD simulation to provide a reasonable estimate for  $\Delta G_{A \rightarrow B}$ .

#### 3.3.1. Umbrella Sampling

Equal sampling of  $A$  and  $B$  can be enforced by introducing a biasing potential that drives the system from state  $A$  into state  $B$ . After correcting for such biasing potential, the free energy can, in principle, be calculated with sufficient accuracy (30). A *single* simulation with a bias potential is not very efficient. Therefore, in practice, *several* independent simulations are carried out, each with a different biasing potential. These potentials are called umbrellas and are placed at different points along the reaction pathway. In each simulation, or window, the sampling is enhanced around the centre of the umbrella potential. Umbrella sampling yields a set of biased probability distributions. To generate the free energy profile for the entire pathway, the results of the various windows are combined and unbiased (31).

In QM/MM simulations, even the sampling of the windows can pose a problem due to the high demand on the computational resources for computing the wave function. As an approximation, the QM subsystem can be kept frozen in the windows. If also the charges on the QM atoms are kept fixed at each umbrella, no QM calculations are needed during the sampling of the remaining MM degrees of freedom. Thus, within this approach, the QM and

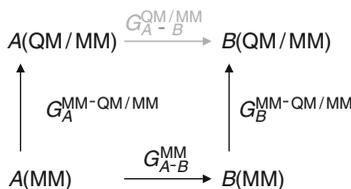


Fig. 6. Thermodynamic cycle for computing the free energy difference between states  $A$  and  $B$  at the QM/MM level ( $\Delta G_{A \rightarrow B}^{\text{QM/MM}}$ ). In the first step, the free energy difference between  $A$  and  $B$  is determined at the MM level ( $\Delta G_{A \rightarrow B}^{\text{MM}}$ ), either by thermodynamic integration or free energy perturbation. In the second step, the free energy required to transform the MM ensemble of  $A$  and  $B$  into the QM/MM ensemble ( $\Delta G_A^{\text{MM} \rightarrow \text{QM/MM}}$  and  $\Delta G_B^{\text{MM} \rightarrow \text{QM/MM}}$ ) are computed by free energy perturbation. The QM/MM free energy of converting  $A$  into  $B$  is calculated by adding up the free energy differences in going around the cycle from  $A(\text{QM/MM})$  to  $B(\text{QM/MM})$ . This procedure avoids computing the  $\Delta G_{A \rightarrow B}^{\text{QM/MM}}$  directly.

MM degrees of freedom are assumed to be uncoupled. Whether such assumption is valid, depends on the process at hand. Another issue concerns finding a suitable reaction path along which the umbrella sampling will be carried out.

### 3.3.2. Free Energy Perturbation

An alternative approach for extracting the free energy associated with the conversion between two states from QM/MM simulations is to use a combination of thermodynamic integration (32) and free energy perturbation (33). In thermodynamic integration (TI), the Hamiltonian is interpolated between the two states with a coupling parameter  $\lambda$ :

$$H(\mathbf{q}, \mathbf{p}, \lambda) = (1 - \lambda)H_A(\mathbf{q}, \mathbf{p}) + \lambda H_B(\mathbf{q}, \mathbf{p}), \quad (10)$$

where  $\mathbf{q}$  and  $\mathbf{p}$  are the positions and momenta of all atoms in the system. To obtain the free energy difference between state  $A$ , when  $\lambda = 0$ , and state  $B$ , when  $\lambda = 1$ , the system is sampled at fixed values of  $\lambda$  between 0 and 1, followed by integration over the ensemble averages of  $\langle \partial H / \partial \lambda \rangle_\lambda$  at these  $\lambda$  values with respect to  $\lambda$ :

$$\Delta G = \int_0^1 \left\langle \frac{\partial H}{\partial \lambda} \right\rangle_\lambda d\lambda. \quad (11)$$

An advantage of the TI approach is that the pathway connecting the two states does neither have to be physically meaningful nor possible. For example, the free energy cost of changing or even disappearing atoms, can be computed efficiently this way. Such non-physical transformations are usually only possible at the MM level. To get the free energy change at the QM/MM level, an additional step is required (34).

Because the free energy is a state function, its magnitude does not depend on the pathway taken. Therefore, one can always construct a so-called thermodynamic cycle, as shown in Fig. 6. For the free energy of a transformation at the QM/MM level, the quantity of interest is the

free energy associated with the top process ( $\Delta G_{A \rightarrow B}^{\text{QM/MM}}$ ). Since the cycle is closed (i.e. the  $\Delta G$ s add up to zero upon completing the cycle), this quantity can be computed as:

$$\Delta G_{A \rightarrow B}^{\text{QM/MM}} = \Delta G_{A \rightarrow B}^{\text{MM}} + \Delta G_B^{\text{MM} \rightarrow \text{QM/MM}} - \Delta G_A^{\text{MM} \rightarrow \text{QM/MM}}, \quad (12)$$

with the free energies defined in Fig. 6.

Thus, instead of calculating  $\Delta G_{A \rightarrow B}^{\text{QM/MM}}$  directly, which is often impossible, one can evaluate this free energy in three steps. First, the free energy of the process is calculated at the MM level, by means of thermodynamic integration (Eq. 11). In the second and third steps, the free energy associated with changing the potential energy landscape from MM to QM/MM is computed for the end states of the TI process ( $\Delta G^{\text{MM} \rightarrow \text{QM/MM}}$ ). One way of obtaining these two quantities is to make use of the free energy perturbation formalism that describes the free energy difference between two states as the overlap between the ensembles (33):

$$\begin{aligned} \Delta G^{\text{MM} \rightarrow \text{QM/MM}} &= G^{\text{QM/MM}} - G^{\text{MM}} \\ &= -k_B T \ln \left\langle \exp \left( -\frac{V^{\text{QM/MM}} - V^{\text{MM}}}{k_B T} \right) \right\rangle_{\text{MM}}, \end{aligned} \quad (13)$$

with  $k_B$  the Boltzmann constant,  $T$  the temperature,  $V^{\text{MM}}$  and  $V^{\text{QM/MM}}$  the potential energy at the QM/MM and MM levels, respectively. The Boltzmann factor is averaged over the ensemble generated at the MM level. Since many MM snapshots may be required to get a converged Boltzmann factor, sampling remains a critical issue.

### 3.4. Computational Spectroscopy

Spectroscopy in the visible and infrared spectral regions are among most important experimental techniques to probe the structure and dynamics of sub-picosecond photochemical processes. However, the interpretation of the spectra requires knowledge about the structure and dynamics of the system under study. Therefore, the full potential of this technique can only be realized when it is complemented by computational modelling. Many spectroscopic quantities can be computed accurately with quantum chemistry methods, but mostly for small model systems in isolation. Including the environment, as in QM/MM methods, therefore, may be required to obtain spectra that can be compared to experiment.

#### 3.4.1. UV/vis Absorption Spectra

This class of spectroscopic techniques probes the energy gaps between the different electronic states of the system. The absorption (or stimulated emission) spectra are sensitive to the structure, and structural changes can be traced in real time by time-resolved spectroscopic measurements (e.g. pump-probe). For small systems, the energy levels of the electronic states can be computed accurately

with high-end *ab initio* methods. Suitable methods are based on the complete active space self-consistent field method, such as CASSCF, RASSCF, and CASPT2 (35). However, these methods are too time and memory consuming for larger systems. Therefore, computing spectra of condensed phase systems requires a QM/MM approach. A realistic spectrum is obtained by evaluating the excitation energies in snapshots taken from classical MD trajectories. After averaging the excitation energies over the ensemble, the computed spectrum can be compared directly to the experimental spectrum (36).

#### 3.4.2. IR Absorption Spectra

Infrared spectroscopy probes transitions between vibrational states. The spectra provide a wealth of information about the structure of the system under study, but the assignment of the observed vibrational bands often requires modelling. A popular computational approach for computing vibrational spectra is the normal mode analysis (NMA). In this technique the matrix of second derivatives of the energy with respect to the nuclear displacements is evaluated and diagonalized. The resulting eigenvalues and eigenvectors are the intensities and vibrational modes of the system, respectively. Because this procedure is preceded by a rigorous energy minimization, the spectra are effectively calculated at zero Kelvin. Therefore, the width of the absorption bands, which reflects thermal averaging, are not accessible in the NMA approach.

An alternative approach to extract IR spectra from QM/MM simulations is to take the Fourier transform of the dipole-dipole autocorrelation function:

$$I(\omega) \propto \int_{-\infty}^{\infty} \langle \mu(t) \cdot \mu(0) \rangle \exp(-i\omega t) dt, \quad (14)$$

with  $I$  the intensity at the vibrational frequency  $\omega$ ,  $\mu(t)$  the system's dipole at time  $t$ . The major drawback of this method, however, is that the dipole moment needs to be sampled sufficiently. Therefore, this approach is most often used in conjunction with semi-empirical methods (37).

---

## 4. Case Study: QM/MM Simulations of a Photochemical Process

Photoactive yellow protein (PYP) is believed to be the primary photoreceptor for the photo-avoidance response of the salt-tolerant bacterium *Halorhodospira halophila* (38). PYP contains a deprotonated 4-hydroxycinnamic acid (or *p*-coumaric acid, PCA) chromophore linked covalently to the  $\gamma$ -sulphur of Cys69 via a thioester bond (Fig. 7). Upon absorbing a blue-light photon, PYP enters a fully reversible photocycle involving several intermediates on timescales ranging from a few hundred femtoseconds to seconds (38). Before the QM/MM work that was done to elucidate

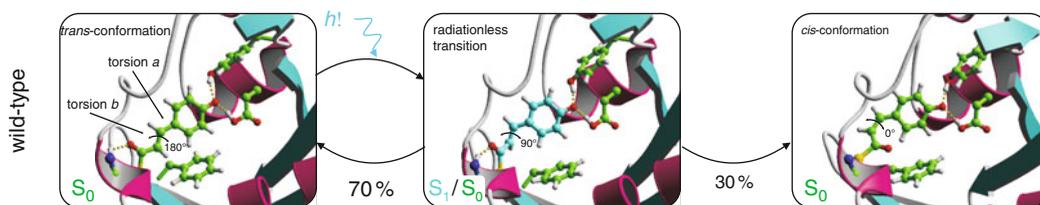


Fig. 7. Snapshots from excited-state trajectories of wild-type PYP, showing the chromophore (pca) in the active site pocket. The first snapshot is at the excitation. The second shows the configuration at the radiationless transition from  $S_1$  to  $S_0$ . The third snapshot shows the photoproduct, in which the carbonyl oxygen of the thioester linkage has flipped and is no longer hydrogen bonded to the backbone of Cys69.

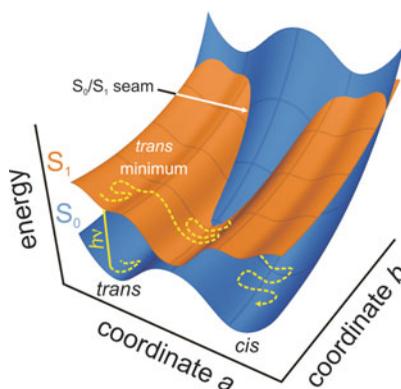


Fig. 8. Schematic overview of a photochemical reaction pathway (*dashed line*). After photon absorption, evolution takes place on the excited-state potential energy surface (*red*) until the system hits the  $S_1/S_0$  intersection seam. At the intersection, a radiationless transition to the ground state occurs (*blue*). After the decay, the system continues evolving in the ground state.

the mechanism by which photon absorption induces signalling, we briefly introduce the basic concepts of photochemistry.

#### 4.1. Photochemical Reactions

The central mechanistic feature in a photochemical reaction is the intersection seam between the potential energy surfaces of the excited ( $S_1$ ) and ground states ( $S_0$ , Fig. 8). Any point on the seam provides a funnel for efficient radiationless decay to the ground state. Just as a transition state separates reactants and products in ground-state chemistry, the seam separates the excited-state branch from the ground-state branch in a photochemical reaction. The crucial difference, however, is that while a transition state connects a reactant to a *single* product via a *single* reaction path, the seam connects the excited state and reacts to *several* products on the ground state via *several* paths. Just as ground-state reactivity is enhanced by a stabilization of the transition state, photoreactivity is also enhanced by stabilization of the seam.

##### 4.1.1. MD Simulations of Photochemical Processes

To model the dynamics of a photochemical reaction, the ground-state and excited-state potential energy surfaces must be described

accurately. After light absorption, the reaction starts in the excited state ( $S_1$ ), but ends in the ground state ( $S_0$ ). Therefore, it is essential to model the radiationless transition between the excited-and ground-state surfaces in a manner that is consistent with a quantum mechanical treatment of the complete system. Because we use Newton's equation of motion to compute molecular dynamics trajectories, the quantum mechanical character of the nuclei is ignored. As a consequence, population transfer from  $S_1$  to  $S_0$  cannot occur, and the classical trajectory is restricted to a single potential energy surface. Thus, in contrast to a full quantum mechanical approach, radiationless transitions do not take place spontaneously. Instead, a binary decision to jump to a different electronic surface must be made at every timestep in a single trajectory. The criterion for switching between electronic states must result in a distribution of state populations, whose average can be compared to observable quantities, such as quantum yield, lifetimes, etc.

In our simulations we allow hopping only at the intersection seam. In principle, this strict diabatic hopping criterion could lead to an underestimation of the population transfer probability. However, because of the high dimensionality of the seam, most trajectories can usually encounter such regions of high probability. The diabatic hopping model is clearly an approximation, but helps one to keep a proper physical insight, which is crucial in understanding complex systems.

#### **4.2. Chromophore in Vacuum**

To understand the intrinsic photochemical properties of the PYP chromophore, we have performed geometry optimizations of an isolated chromophore analogue at the CASSCF level of *ab initio* theory (39). In these optimizations, the complete  $\pi$  system of the chromophore was included in the active space, which thus consisted of 12 electrons in 11  $\pi$  orbitals. In addition to optimizing the local minima on the  $S_1$  potential energy surface and the barriers that separate them, we also located conical intersections in the vicinity of these minima. The optimizations revealed that there are two minima on  $S_1$ : the single-bond twisted minimum, in which the bond adjacent to the phenol ring is rotated by  $90^\circ$ , and the double-bond twisted minimum, in which the ethylenic bond is twisted at  $90^\circ$  (Fig. 9). In the isolated chromophore, there is almost no barrier for reaching the single-bond twisted  $S_1$  minimum from the Franck-Condon region, whereas there is a significant barrier to double-bond rotation. Thus, after photon absorption in vacuum, the main relaxation channel on  $S_1$  involves rotation of the single bond to  $90^\circ$ . We furthermore found that the  $S_1/S_0$  intersection seam lies very far away from this minimum. As a consequence, radiationless decay is not very efficient in vacuum. In subsequent QM/MM simulations, we have probed the effect of different environments on the photochemistry of the chromophore.

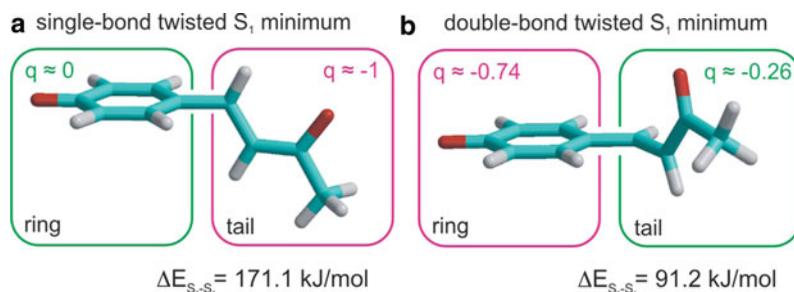


Fig. 9. Excited-state minimum energy configurations of a chromophore analogue. In both the single-bond twisted  $S_1$  minimum (**a**) and the double-bond twisted  $S_1$  minimum (**b**) there is a substantial energy gap between the ground and excited state. The distribution of the negative charge in these minima is opposite.

### 4.3. Chromophore in Water

To examine the effect of an aqueous environment, we have performed 91 QM/MM excited-state dynamics simulations of a chromophore analogue in water (39). The chromophore was described at the CASSCF(6,6)/3-21G level of theory, while the water molecules were modelled by the SPC/E force field (40). The results of these simulations demonstrate that in water, radiationless decay is very efficient (39). The predominant excited-state decay channel involves twisting of the single bond (88%) rather than the double bond (12%). In contrast to vacuum, decay takes place very near these minima. Inspection of the trajectories revealed that decay is mediated by specific hydrogen bond interactions with water molecules. These hydrogen bonds are different for the single- and double-twisted  $S_1$  minima, which reflects the difference in charge distribution between these minima (41). In the single-bond twisted  $S_1$  minimum, the negative charge resides on the alkene moiety of the chromophore (Fig. 9). Three strong hydrogen bonds to the carbonyl oxygen stabilize this charge distribution to such an extent that the seam almost coincides with the single-bond twisted  $S_1$  minimum (Fig. 10). In the double-bond twisted  $S_1$  minimum, the negative charge is localized on the phenolate ring (Fig. 9). Transient stabilization of this charge distribution by two or more strong hydrogen bonds to the phenolate oxygen brings the seam very close to this  $S_1$  minimum (Fig. 10). Thus, in water, the ultrafast excited-state decay is mediated by hydrogen bonds.

### 4.4. Chromophore in the Protein

To find out how the protein mediates the photochemical process, we also carried out a series of QM/MM simulations of wild-type PYP (42). Fig. 7 shows the primary events after photoexcitation in the simulation. The chromophore rapidly decays to the ground state via a  $90^\circ$  rotation of the double bond (Fig. 7), rather than the single bond. During this photo-isomerization process, the hydrogen bonds between the chromophore's phenolate oxygen atom and the side chains of the highly conserved Tyr42 and

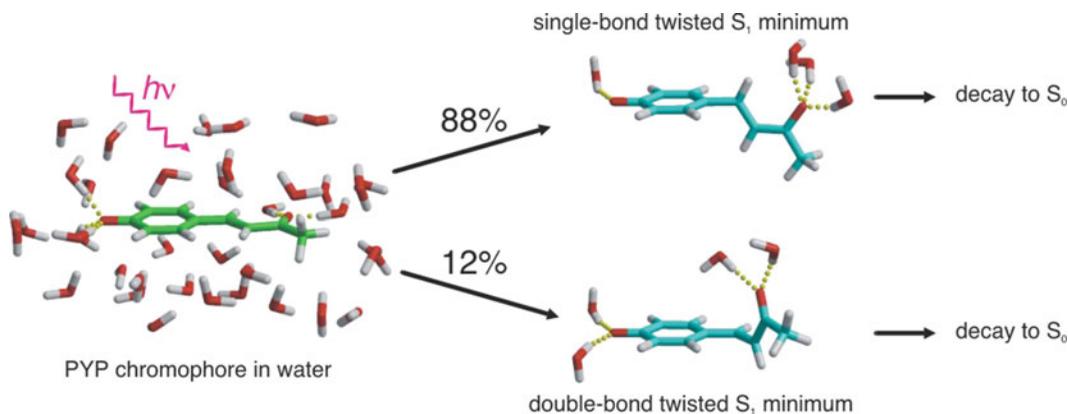


Fig. 10. In water the chromophore undergoes both single- and double-bond isomerization. Excited-state decay from these minima is very efficient due to stabilization of the chromophore's  $S_1$  charge distribution by specific hydrogen bond interactions.

Glu46 residues remain intact. Just as in water, these hydrogen bonds enhance excited-state decay from the double-bond twisted minimum.

Upon returning to the ground state, the chromophore either relaxes back to the original *trans* conformation ( $180^\circ$ ) or it continues isomerizing to a *cis* conformation ( $0^\circ$ ). In the latter case, the relaxation also involves a flip of the thioester linkage, which causes the carbonyl group to rotate  $180^\circ$ . During this rotation, the hydrogen bond between the carbonyl oxygen and the Cys69 backbone amino group is broken (Fig. 7). In total, 14 MD simulations were carried out, initiated from different snapshots from a classical ground-state trajectory. The fluorescence lifetime (200 fs) and isomerization quantum yield (30%) in the simulations agree well with experiments (400 fs (43) and 35% (38), respectively).

In the wild-type protein, no single-bond isomerization was observed. Thus, the protein not only provides the hydrogen bonds required for ultrafast decay but also controls which of the chromophore's bonds isomerizes upon photoexcitation. We identified the positive guanidinium moiety of Arg52 located just above the chromophore ring as the "catalytic" residue that enforces double-bond isomerization. The preferential electrostatic stabilization of the double-bond twisted  $S_1$  minimum by the positive Arg52 strongly favors double-bond isomerization over single-bond isomerization.

To elucidate the role of this arginine in the activation process in more detail, we performed excited-state dynamics simulations on the Arg52Gln mutant of PYP (44). This mutant can still enter the photocycle, albeit with a lower rate and quantum yield (45, 46). Without the positive Arg52, the predominant excited-state reaction in the mutant involves isomerization of a single bond in the

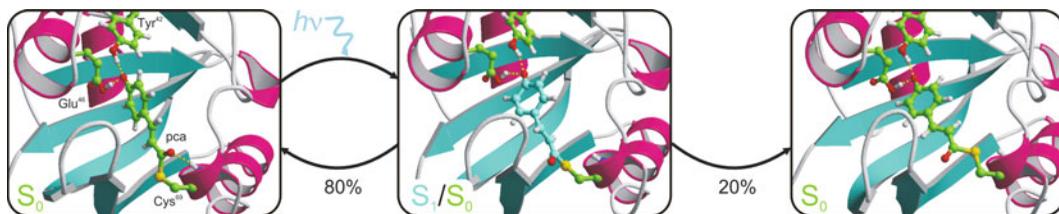


Fig. 11. Snapshots from an excited-state trajectory of the Arg52Gln mutant of PYP, showing the chromophore (pca) in the active site pocket. The first snapshot is at the excitation. The second snapshot shows the configuration at the radiationless transition from  $S_1$  to  $S_0$ . The third snapshot shows the photoproduct. In the mutant, isomerization takes place around the single bond. Like in the wild-type protein, the carbonyl oxygen of the thioester linkage flips, causing the break of the hydrogen bond to the backbone of Cys69.

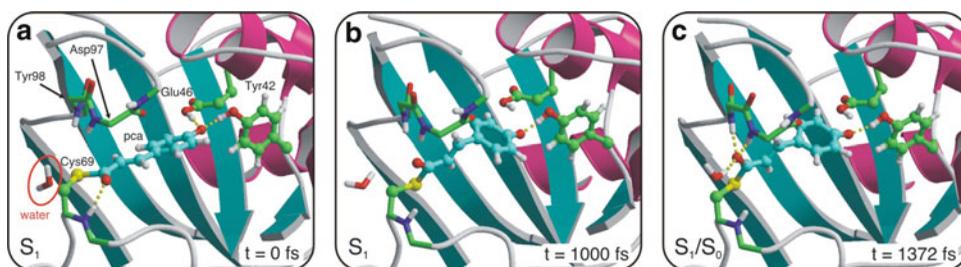


Fig. 12. Snapshots from an excited-state trajectory of the Arg52Gln mutant of PYP, demonstrating that three hydrogen bonds to the carbonyl moiety are essential for  $S_1$  decay at the single-bond twisted minimum. The first snapshot is at the excitation to  $S_1$ . The second snapshot shows the twisted configuration without hydrogen bonds to the carbonyl. The gap between  $S_1$  and  $S_0$  is far too high for decay at this configuration. However, the third snapshot shows two backbone amino groups and a bulk water that has moved into the chromophore pocket during the excited-state dynamics, donating the three hydrogen bonds that are required for efficient decay from the  $S_1$  minimum.

chromophore, rather than the double bond (Fig. 11) (47). This observation confirms that the role of Arg52 is to steer the initial events after photoabsorption to ensure rotation of the double rather than the single bond in the chromophore.

During the rotation of the single bond, the hydrogen bond between the carbonyl oxygen and Cys69 backbone amino group is broken. As shown in Fig. 12, new hydrogen bonds are rapidly formed between the carbonyl oxygen atom and the backbone amino groups of Tyr98 and Asp97. A water molecule from outside enters the chromophore pocket to donate a third hydrogen bond. With these three hydrogen bonds stabilizing the negative charge on the alkene moiety, the chromophore rapidly decays to  $S_0$ . Thus, the decay mechanism in the Arg52Gln mutant and in water are essentially the same.

Although single-bond isomerization does not result in the formation of the *cis* chromophore, a  $180^\circ$  flip of the thioester group and a rupture of the hydrogen bond to Cys69 was observed with a 20% quantum yield (Fig. 12). Together with the experimental

observation that the mutant has a photoactivation quantum yield of about 21% (46), this suggests that the key step to enter the photocycle is the oxygen flip rather than the double-bond isomerization.

To summarize, the simulations are consistent with experimental observations and have provided detailed structural and dynamic information at a resolution well beyond that achievable by other means. From the simulations, key amino acids have been identified and the mechanism by which they control the primary events in the photocycle of PYP. These are (i) double-bond photoisomerization, and (ii) the break of a hydrogen bond between the chromophore and the protein backbone. These events trigger a proton transfer from the protein to the chromophore, which ultimately leads to the signalling state of PYP (48).

---

## 5. Conclusion and Outlook

In this contribution we have reviewed the basic concepts of hybrid QM/MM simulation techniques. More elaborate discussions on the QM/MM method are available as review articles, see for instance references (49–53). In principle, QM/MM simulations can provide detailed structural information of chemical reactions in the condensed phase at an atomic resolution. In practice, the QM/MM methods still suffer from limitations in computational hardware, which restrict both system size and timescale of the processes that can be studied today. However, the expected increase in computer power, complemented by the development of more efficient electronic structure methods and new algorithms may soon enable the investigation of reactions in larger systems and at longer timescales. Therefore, QM/MM simulation has the potential to lead to a better understanding of chemical reactions, and the mechanisms by which in particular protein environments control these reactions. Ultimately, these simulations may not only enable accurate predictions of chemical properties but also become a standard tool for rational design of artificial molecular devices.

---

## Acknowledgements

The Volkswagenstiftung and Deutsche Forschungsgemeinschaft (SFB755) are acknowledged for their financial support. I am grateful to Dr. Mehdi Davari and Pedro Valiente for critically reading the manuscript.

## References

- Jensen F (2001) Introduction to computational chemistry. Wiley, New York
- Berendsen HJC (2001) Bio-molecular dynamics comes of age. *Science* 271:954–955
- McCammon JA, Gelin BR, Karplus M, Wolynes PG (1976) Hinge-bending mode in lysozyme. *Nature* 262:325–326
- Shaw D, Maragakis P, Lindorff-Larsen K, Piana S, Dror R, Eastwood M, Bank J, Jumper J, Salmon J, Shan Y, Wriggers W, (2010) Atomic-level characterization of the structural dynamics of proteins. *Science* 330:341–346
- Warshel A, Levitt M (1976) Theoretical studies of enzymatic reactions: dielectric, electrostatic and steric stabilization of carbonium ion in the reaction of lysozyme. *J Mol Biol* 103:227–249
- Maseras F, Morokuma K (1995) IMOMM—a new integrated ab-initio plus molecular mechanics geometry optimization scheme of equilibrium structures and transition-states. *J Comput Chem* 16:1170–1179
- Svensson M, Humbel S, Froese RDJ, Matsubara T, Sieber S, Morokuma K (1996) ONIOM: a multilayered integrated MO + MM method for geometry optimizations and single point energy predictions. A test for Diels–Alder reactions and Pt(*t*-Bu)<sub>3</sub><sub>2</sub> + H<sub>2</sub> oxidative addition. *J Phys Chem* 100:19357–19363
- Bayly C, Cieplak P, Cornell W, Kollman P (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges—the RESP model. *J Phys Chem* 97:10269–10280
- Besler B, Merz K, Kollman P (1990) Atomic charges derived from semiempirical methods. *J Comput Chem* 11:431–439
- Szabo A, Ostlund NS (1989) Modern quantum chemistry. Dover Publications, New York
- Field MJ, Bash PA, Karplus M (1990) A combined quantum-mechanical and molecular mechanical potential for molecular-dynamics simulations. *J Comp Chem* 11:700–733
- Brooks B, Karplus M (1983) Harmonic dynamics of proteins—normal-modes and fluctuations in bovine pancreatic trypsin-inhibitor. *Proc Natl Acad Sci USA* 80:6571–6575
- Dewar M (1983) Development and status of MINDO/3 and MNDO. *J Mol Struct* 100:41–50
- Amara P, Field MJ (2003) Evaluation of an ab initio quantum mechanical/molecular mechanical hybrid-potential link-atom method. *Theor Chem Acc* 109:43–52
- Lamoureux G, Roux B (2003) Modeling induced polarization with classical Drude oscillators: theory and molecular dynamics simulation algorithm. *J Chem Phys* 119:3025–3039
- Warshel A, Sharma PK, Kato M, Xiang Y, Liu HB, Olsson MHM (2006) Electrostatic basis for enzyme catalysis. *Chem Rev* 106:3210–3235
- Rappe AK, Goddard III WA (1991) Charge equilibration for molecular dynamics simulations. *J Phys Chem* 95:3358–3363
- Zhang Y, Lin H, Truhlar D (2007) Self-consistent polarization of the boundary in the redistributed charge and dipole scheme for combined quantum-mechanical and molecular-mechanical calculations. *J Chem Theory Comput* 3:1378–1398
- Hillier I (1999) Chemical reactivity studied by hybrid QM/MM methods. *J Mol Struct (Theochem)* 463:45–52
- Assfeld X, Rivail J (1996) Quantum chemical computations on parts of large molecules: the ab initio local self consistent field method. *Chem Phys Lett* 263:100–106
- Gao J, Amara P, Alhambra C, Field M (1998) A generalized hybrid orbital (GHO) method for the treatment of boundary atoms in combined QM/MM calculations. *J Phys Chem A* 102:4714–4721
- Philipp DM, Friesner RA (1999) Mixed ab initio QM/MM modeling using frozen orbitals and tests with alanine dipeptide and tetrapeptide. *J Comput Chem* 20:1468–1494
- Nicoll R, Hindle S, MacKenzie G, Hillier I, Burton N (2001) Quantum mechanical/molecular mechanical methods and the study of kinetic isotope effects: modelling the covalent junction region and application to the enzyme xylose isomerase. *Theor Chem Acc* 106:105–112, 10th International Congress of Quantum Chemistry, Nice, France, June 13–15, 2000
- Rodriguez A, Oliva C, Gonzalez M, van der Kamp M, Mulholland A, (2007) Comparison of different quantum mechanical/molecular mechanics boundary treatments in the reaction of the hepatitis C virus NS3 protease with the NS5A/5B substrate. *J Phys Chem B* 111:12909–12915
- Dewar MJS, Zoebisch EG, Healy EF, Stewart JJP (1985) The development and use of quantum-mechanical molecular models AM1—a new general-purpose quantum-mechanical molecular-model. *J Am Chem Soc* 107:3902–3909

26. Stewart JJP (1989) Optimization of parameters for semiempirical methods. 1. Method. *J Comput Chem* 10:209–220
27. Stewart JJP (1989) Optimization of parameters for semiempirical methods. 2. Applications. *J Comput Chem* 10:221–264
28. Elstner M, Porezag D, Jungnickel G, Elsner J, Haugk M, Frauenheim T, Suhai S, Seifert G (1998) Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. *Phys Rev B* 58:7260–7268
29. Klahn M, Braun-Sand S, Rosta E, Warshel A (2005) On possible pitfalls in ab initio quantum mechanics/molecular mechanics minimization approaches for studies of enzymatic reactions. *J Phys Chem B* 109:15645–15650
30. Torrie GM, Valle JP (1977) Non-physical sampling distributions in Monte-Carlo free energy estimation—umbrella sampling. *J Comput Phys* 23:187–199
31. Roux B (1995) The calculation of the potential of mean force using computer-simulations. *Comp Phys Comm* 91:275–282
32. Kirkwood J (1935) Statistical Mechanics of Fluid Mixtures. *J Chem Phys* 3:300–313
33. Zwanzig R (1954) High-temperature equation of state by a perturbation method. I. Nonpolar gases. *J Chem Phys* 22:1420–1426
34. Muller R, Warshel A (1995) Ab-initio calculations of free energy barriers for chemical-reactions in solution. *J Phys Chem* 99:17516–17524
35. Roos BO (1999) Theoretical studies of electronically excited states of molecular systems using multiconfigurational perturbation theory. *Acc Chem Res* 32:137–144
36. Schäfer LV, Groenhof G, Klingner AR, Ullmann GM, Boggio-Pasqua M, Robb MA, Grubmüller H (2007) Photoswitching of the fluorescent protein asFP595: mechanism proton pathways, and absorption spectra. *Angew Chemie Int Ed* 46:530–536
37. Kaminski S, Gaus M, Phatak P, von Stetten D, Elstner M, Mrogiński M (2010) Vibrational Raman spectra from the self-consistent charge density functional tight binding method via classical time-correlation functions. *J Chem Theory Comput* 6:1240–1255
38. Hellingwerf KJ, Hendriks J, Gensch T (2003) Photoactive yellow protein, a new type of photoreceptor protein: will this “yellow lab” bring us where we want to go? *J Phys Chem A* 107:1082–1094
39. Boggio-Pasqua M, Robb M, Groenhof G (2009) Hydrogen bonding controls excited-state decay of the photoactive yellow protein chromophore. *J Am Chem Soc* 131:13580
40. Berendsen H, Grigera J, Straatsma T (1987) The missing term in effective pair potentials. *J Phys Chem* 91:6269–6271
41. Gromov EV, Burghardt I, Hynes JT, Köppel H, Cederbaum LS (2007) Electronic structure of the photoactive yellow protein chromophore: ab initio study of the low-lying excited singlet states. *J Photochem Photobiol A* 190:241–257
42. Groenhof G, Bouxin-Cademartory M, Hess B, De Visser S, Berendsen H, Olivucci M, Mark A, Robb M (2004) Photoactivation of the photoactive yellow protein: why photon absorption triggers a trans-to-cis isomerization of the chromophore in the protein. *J Am Chem Soc* 126:4228–4233
43. Mataga N, Chosrowjan H, Shibata Y, Imamoto Y, Tokunaga F (2000) Effects of modification of protein nanopore structure and change of temperature on the femtosecond to picosecond fluorescence dynamics of photoactive yellow protein. *J Phys Chem B* 104:5191–5199
44. Shimizu N, Kamikubo H, Yamazaki Y, Imamoto Y, Kataoka M (2006) The crystal structure of the R52Q mutant demonstrates a role for R52 in chromophore pK(a) regulation in photoactive yellow protein. *Biochemistry* 45:3542–3547
45. Chagnenet-Barret P, Plaza P, Martin MM, Chosrowjan H, Taniguchi S, Mataga N, Imamoto Y, Kataoka M (2007) Role of arginine 52 on the primary photoinduced events in the PYP photocycle. *Chem Phys Lett* 434:320–325
46. Takeshita K, Imamoto Y, Kataoka M, Mihara K, Tokunaga F, Terazima M (2002) Structural change of site-directed mutants of PYP: new dynamics during pR state. *Biophys J* 83:1567–1577
47. Groenhof G, Schäfer LV, Boggio-Pasqua M, Grubmüller H, Robb MA (2008) Arginine 52 controls photoisomerization in photoactive yellow protein. *J Am Chem Soc in press JACS* 130: 3250–3251
48. Groenhof G, Lensink MF, Berendsen HJC, Mark AE (2002) Signal transduction in the photoactive yellow protein. II. Proton transfer initiates conformational changes. *Proteins* 48:212–219
49. Gao J (1996) Hybrid quantum and molecular mechanical simulations: an alternative avenue to solvent effects in organic chemistry. *Acc Chem Res* 29:298–305
50. Monard G, Merz K (1999) Combined quantum mechanical/molecular mechanical

- methodologies applied to biomolecular systems. *Acc Chem Res* 32:904–911
51. Gao J, Truhlar D (2002) Quantum mechanical methods for enzyme kinetics. *Annu Rev Phys Chem* 53:467–505
52. Friesner R, Guallar V (2005) Ab initio quantum chemical and mixed quantum mechanics/molecular mechanics (QM/MM) methods for studying enzymatic catalysis. *Annu Rev Phys Chem* 56:389–427
53. Senn H, Thiel W (2009) QM/MM methods for biomolecular systems. *Angew Chem Int Ed* 48:1198–1229