Electrostatic versus Resonance Interactions in Photoreceptor Proteins: The Case of Rhodopsin

Riccardo Guareschi,† Omar Valsson,‡ Carles Curutchet,§ Benedetta Mennucci,*§ and Claudia Filippi*†

†MESA+ Institute for Nanotechnology, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands
‡Department of Chemistry and Applied Bioscience, ETH Zurich and Facoltà di Informatica, Instituto di Scienze Computazionali, Università della Svizzera italiana, Via Giuseppe Buffi 13, CH-6900 Lugano, Switzerland
§Departament de Fisicoquímica, Facultat de Farmàcia, Universitat de Barcelona, Av. Joan XXIII, s/n 08028 Barcelona, Spain
* Dipartimento di Chimica e Chimica Industriale, Università di Pisa, Via Giuseppe Moruzzi 13, 56124 Pisa, Italy

Supporting Information

ABSTRACT: Light sensing in photoreceptor proteins is subtly modulated by the multiple interactions between the chromophoric unit and its binding pocket. Many theoretical and experimental studies have tried to uncover the fundamental origin of these interactions but reached contradictory conclusions as to whether electrostatics, polarization, or intrinsically quantum effects prevail. Here, we select rhodopsin as a prototypical photoreceptor system to reveal the molecular mechanism underlying these interactions and regulating the spectral tuning. Combining a multireference perturbation method and density functional theory with a classical but atomistic and polarizable embedding scheme, we show that accounting for electrostatics only leads to a qualitatively wrong picture, while a responsive environment can successfully capture both the classical and quantum dominant effects. Several residues are found to tune the excitation by both differentially stabilizing ground and excited states and through nonclassical “inductive resonance” interactions. The results obtained with such a quantum-in-classical model are validated against both experimental data and fully quantum calculations.

Photoreceptor proteins play a fundamental role in the function of many biological systems. Specific to every system is how the initial step of light absorption is subtly tuned by the protein matrix which acts quite differently than a more homogeneous environment like a solvent. In particular, the residues closely surrounding the chromophoric unit can significantly affect its response to light, but also longer-range effects, well beyond the first shell of interacting residues, are non-negligible. Rationalizing the essence of the regulatory mechanism of absorption is of fundamental importance not only to understand a specific photoreceptor but also to get a more general appreciation of how these light-sensing systems have optimized their structure and molecular composition. To this aim, three precise questions have to be answered: (i) Are the chromophore–protein interactions purely classical (mostly electrostatic and induction), or do we need to account for quantum effects? (ii) Which residues are active in the excitation process, and what is the nature of their contribution? (iii) Can we define a unique theoretical framework for all these interactions?

The best computational strategy to answer these questions might seem to use an expensive quantum chemical description for an extended region beyond the chromophore’s boundaries. This would certainly allow us to establish reference target values for excited-state properties at the chosen level of theory but, ultimately, hinders their analysis because of the impossibility of disentangling the real nature of the interactions. A more effective strategy both in terms of computational cost and of ability to further our understanding is represented by a focused approach in which the quantum description is limited to the photoexcited site while the embedding protein environment is treated classically but still preserving the atomistic details. Most previous applications of such a strategy to photoreceptor proteins have employed embedding schemes that account for only electrostatic effects, especially through the use of static point charges. These approaches, however, are not only unsuitable to recover the complete essence of the chromophore–protein coupling but also prone to lead to incorrect interpretations. Here, we demonstrate that atomistic but polarizable classical embeddings can be successfully used to obtain a correct description of electrostatics and induction interactions and also to account for nonclassical effects as previously found for continuum representations of solvated systems. Moreover, we show that the same analysis can be employed to fingerprint the role of the surrounding residues in determining the excitation of the embedded chromophore.

Received: September 7, 2016
Accepted: October 27, 2016
Published: October 27, 2016
We focus on a well-known archetype of photoreceptor protein, namely, one of the visual opsins, the bovine rhodopsin. Visual opsins regulate the process of vision at the molecular level and have become a prime playground for the understanding and further manipulation of how a protein environment controls the light-activated response of an embedded photosensitive species.4–6 It is remarkable that most members of this large family of receptors share the same photosensitive component, the 11-cis retinal protonated Schiff base (RPSB) chromophore (Figure 1) and that relatively minor variations in the protein surroundings can tune its absorption over a spectral range as wide as 420–570 nm. This is the result of multiple interactions between the chromophore and the protein whose fundamental origin, however, has so far eluded our understanding.

Bovine rhodopsin, in particular, has been extensively studied both experimentally4–8 and theoretically.9–25 This protein exhibits a strong absorption centered at 2.49 eV, dominated by the $\pi \rightarrow \pi^*$ transition on the RPSB chromophore, and associated with electronic charge transfer from the $\beta$-ionone ring toward the terminal side of the conjugated chain. From qualitative considerations, one expects that the electrostatic interaction of the positive chromophore with its Glu113 counterion (Figure 1) induces a blue-shift in absorption with respect to the isolated system. The size of this shift and the role played by the rest of the protein in compensating this shift has been much debated over the years with estimates for both effects often differing considerably in magnitude.9,12–16,19

Given the complexity of the system, multiscale approaches to bridge between the smaller scales of the chromophore and the larger protein environment have been commonly employed to study the absorption properties of opsins, in particular, the quantum-mechanics-in-molecular-mechanics (QM/MM) approaches where the environment is treated as a sea of classical static point charges. However, when opting for such a multiscale scheme, one has implicitly acknowledged that the environment has not an active role in the photoduced process. To overcome these limitations, recent theoretical studies have extended the quantum treatment to larger regions and suggested that at least one shell of amino acids in proximity of the RPSB chromophore participates in the photoduced process.20,24 These findings are qualitatively in line with earlier studies26–28 on prototypical opsins by Elstner and co-workers, who found the environment to respond electrostatically to the photoexcitation of the chromophore when relaxed in a state-specific manner using either a quantum-in-quantum partition or through polarizable force fields. In all these cases, the observed effects were classified as “polarization”, but is the polarization captured by a responsive (quantum or classical) treatment of the environment the same one recovered in the quantum calculation of a large photoexcited region?

To answer this fundamental question, we begin our analysis of rhodopsin from the TDDFT excitation energies (with the CAM-B3LYP exchange-correlation functional29) calculated using an electrostatic (QM/MM) or a polarizable embedding (based on atomic charges in combination with induced dipoles in a so-called QM/MMpol approach).30 In the latter case, three different schemes are employed to selectively introduce different types of interactions between the RPSB chromophore and the protein. First, we polarize the dipoles to the ground-state density of RPSB and keep them frozen in the calculation of the excitation energy (polGS). Then, state-specific dipoles are used to account for possible differential electrostatic and induction effects between ground and excited states (polSS). Given the large transfer of electronic charge following photoexcitation, one would expect a strong electrostatic interaction of RPSB with its surroundings and a sizable polSS correction on the polGS excitation energies (a rather small shift, however, was previously reported for one structural realization of the system17). In the last scheme (polRes), we adopt a computational framework where the induced dipoles respond to the transition density of the photoexcited species. This embedding condition accounts for the nonclassical part of the coupling between the chromophore excitation and the environment which, in previous analyses within a continuum formulation of the classical subsystem, has been described as a dispersion-like term, an excitonic coupling, or an “inductive resonance”.31 This last definition will be adopted also in the present study.

To account for temperature effects and avoid the risk of limiting the analysis to a single and possibly not representative structure, we consider here different sets of frames extracted from QM/MM molecular dynamics (MD) trajectories where the QM part is treated within DFT using the PBE exchange-correlation functional32 (see the Supporting Information for the details). Frames 1–6 are the central frames of the most populated clusters obtained in a cluster analysis of a trajectory we generated with the chromophore and the Glu113 counterion in the QM region. We then analyze a second set (frames 7–11) obtained in a previous study33 from a different QM/MM trajectory and selected from the region of phase space which gives excitation energies close to the maximum of the theoretical absorption band computed at the ZINDO level. Finally, we include two more structures, an annealed structure from the same trajectory of frames 1–6 and frame 12 from another trajectory previously obtained with also the anionic residue Glu181 included in the QM region33 and selected with the same recipe used to obtain frames 7–11. All the results are summarized in Figure 2.

The excitation energies of all snapshots follow precisely the same trend: the inclusion of static point charges causes a significant blue shift of the excitation energy (about 0.2 eV) with respect to the values computed on the isolated chromophore at the geometry extracted from the protein. This shift can be mostly attributed to the presence of the counterion stabilizing the ground state and inhibiting the charge transfer in the excited state. A very similar shift is obtained when the chromophore is embedded in an environment with its polarization being kept frozen to the ground state of RPSB (polGS): the excitation energies are only somewhat
more blue-shifted than the MM values as a consequence of the further stabilization of the ground state. The effect of this improved ground-state description is rather small at the TDDFT level, being at most 0.06 eV and in average 0.02 eV. If we further enhance the description of the chromophore–protein coupling by allowing the induced dipoles to relax according to the excited state (polSS), we recover a red-shift in the excitation energies with respect to the polGS values, which is rather small and consistently between −0.03 and −0.02 eV. The presence of a responsive environment has a much larger effect only when the induced dipoles are calculated within the resonance regime (polRes), namely, in response to the transition density associated with the electronic excitation. The nonclassical part of the coupling captured in this model appears therefore to play a dominant role, lowering the excitation energies by as much as 0.16–0.18 eV with respect to the polGS value and largely recovering the so-called counterion quenching which was lost in an electrostatic treatment of the protein. We finally note that the trend of excitation energies with respect to the different models of the environment is similar in all frames, though the latter are selected based on very different criteria. In particular, while frames 7–11 can be considered as biased toward the close neighborhood of the chromophores as they have been selected on the basis of the excitation energies, the choice of frames (1–6) has been based on a cluster analysis performed over the complete protein, and consequently, they are expected to yield a larger spread of excitation energies (see the Supporting Information for further discussion).

To better understand the origin of the calculated shifts we recall that, perturbatively to linear order, the polSS and polRes corrections with respect to the polGS excitation energies are proportional to the square of the variation of the molecular dipole between ground and excited states and to the transition dipole moment, respectively, so that

\[
\frac{|\Delta E_{\text{exc}}(SS-GS)|}{|\Delta E_{\text{exc}}(Res-GS)|} \approx \frac{|\Delta \mu|^2}{2 |\mu_0|^2}
\]

with \(\Delta E_{\text{exc}}(SS-GS)\) and \(\Delta E_{\text{exc}}(Res-GS)\). If we first focus on the state-specific contribution, we observe that we would have expected a more sizable reduction of the excitation energy due to the inclusion of differential polarization effects: the excited state of RPSB displays a strong charge redistribution from the β-ionone ring to the nitrogen terminus, which should be greatly stabilized by the relative optimization of the chromophore surroundings. The difference between the ground- and excited-state molecular dipole moment is in fact experimentally rather large (ca. 12–13 D), \(^{34,35}\) The underestimation of this electrostatic differential contribution can be explained as resulting from a limitation of TDDFT/CAM-B3LYP, which tends to underestimate the dipole moment changes in charge-transfer excitations. \(^{36}\) Our calculations for rhodopsin are affected by the same problem and yield values of \(|\Delta \mu|\) (computed in polGS embedding) that are too low by about a factor of 2. This incorrectly predicted change in the dipole moment translates in an underestimation by about a factor of 4 of the electrostatic differential polarization contributions in the TDDFT excitation energies, as further discussed below in comparison with wave function-based calculations.

The polRes corrections to the polGS excitation energies are instead proportional to the magnitude square of the transition dipole moment which, for the bright \(\pi \rightarrow \pi^*\) transition of RPSB in rhodopsin, is quite large. The TDDFT transition dipole moments of about 12 D (see Table 1) lead in fact to sizable polRes contributions to the excitation energies. In this case, the estimates can be considered reliable as TDDFT/CAM-B3LYP has generally been shown to produce reasonably good oscillator strengths. \(^{37}\)

The polRes and polSS schemes account therefore for different components of the chromophore–protein interaction.

---

### Table 1. Excitation Energies (Electronvolts) of Rhodopsin Computed with Nonpolarizable QM/MM and the QM/polGS and Extrapolated QM/pol(Res+SS) Schemes Using TDDFT/CAM-B3LYP and CASPT2/ANO-L-VDZP

| Frame | \(|\mu_0|\) | \(|\Delta \mu|\) | \(E_{\text{exc}}^{\text{QM/MM}}\) | \(E_{\text{exc}}^{\text{polGS}}\) | \(\Delta E_{\text{exc}}^{(SS-GS)}\) | \(\Delta E_{\text{exc}}^{(Res-GS)}\) | \(E_{\text{exc}}^{\text{pol(Res+SS)}}\) |
|-------|----------------|-----------------|-----------------|----------------|-----------------|-----------------|----------------|
| 4     | 11.84          | 8.29            | 2.652           | 2.680          | −0.034          | −0.160          | 2.486          |
| 8     | 11.88          | 5.73            | 2.630           | 2.627          | −0.018          | −0.158          | 2.451          |
| 12    | 11.85          | 7.18            | 2.653           | 2.662          | −0.028          | −0.160          | 2.475          |
| 4     | 9.35           | 14.68           | 2.645           | 2.784          | −0.136          | −0.110          | 2.537          |
| 8     | 10.64          | 13.50           | 2.609           | 2.686          | −0.129          | −0.160          | 2.397          |
| 12    | 11.38          | 15.41           | 2.724           | 2.809          | −0.113          | −0.124          | 2.571          |

The polSS and polRes corrections to the polGS values are also listed. The polRes results are computed at the TDDFT level and estimated in CASPT2 by rescaling the CASPT2 polSS corrections according to eq 1. The polGS transition dipole moment, \(|\mu_0|\), and the variation of the molecular dipole moment, \(|\Delta \mu|\), are given in debye.
that coexist in rhodopsin. We can thus extrapolate the combined effects of these polarization treatments by adding the two corrections on the polGS excitation energies as shown in Figure 2. As discussed below, such an additive approach appears to be appropriate for rhodopsin, being supported by the comparison with TDDFT calculations on large clusters. The extrapolated excitation energies are reduced by about 0.2 eV with respect to the polGS values and, for the snapshots energetically close to the annealed one, in good agreement with the experimental absorption maximum. Nevertheless, we should not attribute too much significance to this agreement: we must recall that the state-specific contribution is underestimated in TDDFT, resulting in a blue-shift which is fortuitously compensated by the use of an RPSB geometry computed with a generalized gradient approximation, yielding a smaller bond length alternation and a consequent red-shift. Furthermore, we are neglecting vibronic effects which are not negligible in in photoactive rhodopsin and should result in the computed vertical excitation energy being higher than the absorption maximum.

The TDDFT/MMpol calculations allows us to further deepen our analysis and identify which residues respond more strongly and whether they polarize to the ground state, to the excited state, or in a resonance fashion. In Figure 3, we show for the annealed frame the dominant contributions to the ground-state polarization ($E_{pol}^{GS}$) and the contributions induced at the polSS and polRes level relative to polGS

$$\delta E_{pol}^{XX} = \frac{1}{2} \sum_a \mu_a (\rho_{XX}^a) \cdot \mathbf{E}_a (\rho_{XX}^a)$$

where the sum is over the classical polarizable sites and the dipoles are induced by either the transition density, $\rho_{pol}^{Res}$ or by the density difference, $\rho_{pol}^{SS} = \Delta \rho$, between the ground and excited states of RPSB. For a set of frames, we then determine the residues with contributions to $\delta E_{pol}$ larger than about 0.09 kcal/mol. Because the polSS correction to polGS is underestimated by about a factor of 4 as discussed above, we identify the residues whose polSS contributions rescaled by the same factor are above the chosen threshold. In Figure 4, we show the position of the common set of surviving residues, which nicely embrace the RPSB chromophore.

It is first important to appreciate that the dominating contributions to the ground-state polarization are mostly carried by a few residues in proximity of the nitrogen end of RPSB, in particular, the covalently linked Lys296, the counterion Glu113, a water molecule occasionally hydrogen-bonded to it, and Ser186, while the response to the excitation is spread over a larger number of residues surrounding also the β-ionone ring. As expected, the polSS and polRes responses correlate rather well: the residues more strongly interacting with the chromophore in the excited state are also the ones that predominantly interact at the polRes level. In addition to the three obvious candidates close to the positive terminus of RPSB (Lys296, Glu113, Ser186), the residues displaying a sizable π electronic system (Trp265, Tyr268, Phe293) contribute the most both to the polSS and polRes responses, with Trp265 π-stacked to one end of the chromophore giving the absolute largest value in several frames. A few residues close to the central part of the conjugated skeleton of the chromophore (Ala117, Tyr118) further stabilize the excitation. We finally note that the validity of this analysis, based on eq 2, to pinpoint the role of the various residues is strongly supported by the close equality recovered between the total $\delta E_{pol}^{Res}$ and the excitation energy difference $\Delta E_{exc}^{Res-GS}$ (see the Supporting Information).

We can also exploit this analysis to corroborate our approach and perform supermolecular TDDFT reference calculations on realistic cluster models of the system constructed from the most responsive residues we have identified above. We follow such a procedure for the annealed frame and build two clusters comprising the RPSB chromophore and either the counterion only or all important residues. As shown in Figure 5, the resulting supermolecular TDDFT excitation energies are in very good agreement with the pol(Res+SS) values, differing by roughly the same small amount, and further support the validity of our approach. Moreover, this analysis further details the role of the different residues and their influence on the chromophore: the introduction of the nearby Glu113 counterion leads to a large blue-shift because the electrostatics is there dominant, stabilizing the charged ground state. By enlarging the number of interacting residues through the cluster, an opposite
trend is observed showing that their presence is necessary to allow the stabilization of the excited state. This stabilization is further enhanced when the full system is accounted for within the MMPol approach.

The understanding gained so far appears to be sound (and internally consistent) within the limitations of TDDFT, which we already know, for instance, to underestimate differential polarization effects. To further validate our observations, we also employ a wave function-based description of the QM region and select three frames (4, 8, and 12) to be treated with the complete-active-space self-consistent-field (CASSCF) approach followed by second-order perturbation theory (CASPT2). The comparison between the CASPT2 and TDDFT excitation energies is shown in Figure 6 for the three selected frames and is summarized in Table 1. The MM dipoles are induced in a CASSCF calculation targeting either the ground or the excited state and are used unchanged in the PT2 computation of the polGS and polSS excitation energies.

The dependence of the CASPT2 excitation energies on the choice of electrostatic model for the environment follows a trend similar to the one observed at the TDDFT level, but the magnitude of the shifts induced by the different descriptions of the environment is much larger. The use of a ground-state-polarized protein raises the excitation energies with respect to point-charge embedding by more than 0.1 eV and, accounting for differential polarization effects in the polSS scheme, yields a correction of the same magnitude and opposite sign. The change in the correlated excitation energies computed with state-specific dipoles with respect to the polGS values is much larger than in the TDDFT case as a consequence of the more accurate estimation of $|\Delta \mu|$ at the CASPT2 level. In fact, the CASPT2 differences, $\Delta E_{\text{exc}}^{(\text{SS}-\text{GS})}$, can be remarkably well estimated by simply rescaling the corresponding TDDFT values by the ratio squared between the CASPT2 transition dipole moment and the variation of the molecular dipoles obtained in a polGS calculation. These estimates are in very good agreement with the corresponding TDDFT linear-response corrections, as shown for the three frames in Table 1, where the CASPT2 resonance contributions consistently obtained as in eq 1 are listed.

The analysis carried out within TDDFT and CASPT2 reveals a qualitatively similar behavior. Furthermore, the quantitative differences can be rationalized if one accounts for the underestimated electrostatic response of TDDFT in the excited state. From both analyses, it is clear that an embedding scheme based on fixed point charges not only is too simplistic to capture the complexity of the protein-chromophore interactions but also can lead to qualitatively wrong results. The comparison with the different polarizable models has here given.

Figure 4. RPSB chromophore covalently linked to Lys296 and the residues with the largest contributions $\delta E_{\text{pol}}$ to the relative polarization energies in the (a) polSS and (b) polRes schemes. The residues are visualized up to their C$_{\alpha}$ atoms.

Figure 5. Comparison of the excitation energies (in electronvolts) computed for the annealed frame with TDDFT/CAM-B3LYP in a supermolecular calculation on the complete system or within QM/MMpol with only RPSB in the QM region. We consider the isolated chromophore and add the counterion, a cluster of residues (see text), and the whole protein.

Figure 6. Comparison of the TDDFT/CAM-B3LYP and CASPT2 excitation energies (in electronvolts) for frames 4, 8, and 12 for RPSB isolated and with different models of the protein environment.
us the needed clues on where a point-charge model goes amiss. We have also seen that, unfortunately, the use of a polarizable embedding frozen in the ground state does not ameliorate the point-charge results but further blue-shifts the excitation energy. An improved description of ground-state rhodopsin alone is therefore not sufficient, and the lack of environmental response to the electronic transition lies at the heart of the problem. The correction due to the relaxation of the protein polarization in the excited state is in fact expected to be large because of the density difference upon excitation resulting from the flux of electronic charge from the β-ionone ring to the nitrogen terminus of RPSB. When this physical feature is properly reproduced by the selected QM method, as in the case of CASPT2, the state-specific correction gives indeed a red-shift of more than 0.1 eV, which can be ascribed to the response of several residues in the protein pocket of the chromophore. Nevertheless, the inclusion of differential polarization effects does not account for the complete description of the chromophore–protein coupling. The same residues which more strongly polarize in the excited state of the chromophore also “resonate” to the transition density when treated within a linear-response scheme at the TDDFT level. The resulting correction to the excitation energies is of the same order of magnitude as the state-specific one if the QM method describes equally well the transition density and the change in the density going from the ground to the excited state. If these two contributions are finally added together, the resulting excitation energy is in good agreement with the measured value.

Our analysis on this prototypical photosensitive system demonstrates that the protein tuning of the excitation process is not of electrostatic nature only and is therefore not straightforwardly amenable to a description in terms of a multiscale quantum-in-classical approach. Consequently, the use of static point charges commonly adopted to treat photoreceptor proteins is generally bound to be inadequate. For the same reasons, also a quantum-in-quantum approach based on the product state of chromophore and protein wave functions is expected to fail. The process of light absorption is in fact modulated by the concurrent presence of induction and resonance effects. A computational option to recover them at once is to perform an explicit quantum calculation on an extended region surrounding the photoactive moiety, employing a technique able to describe coupled excitations between the chromophore and the many residues we have shown to participate in the excitation process. This is unfortunately very expensive and currently limited to the use of low-correlation approaches. Here, we have, however, outlined an alternative, practical route in which a classical embedding model is adapted to mimic both induction and resonant coupling and to provide key information on the origin of the interactions and the detailed quantitative role of specific amino acids in the photoexcitation of a complex chromophore–protein system. Importantly, the subtle interplay between electrostatic and resonance interactions we unveiled in absorption is expected to modulate also subsequent photo-induced processes which often characterize photoreceptor proteins and ultimately regulate their function. The computational framework presented here has the potential to capture the essence of the chromophore–protein coupling in all relevant steps of its photochemical evolution.

- **ASSOCIATED CONTENT**

  - **Supporting Information**
    The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpclett.6b02043.
    Computational details; TDDFT, CASSCF, and CASPT2 excitation energies; further analysis of resonance CASPT2 and TDDFT corrections; dominant contributions to polarization energies for selected frames; and dependence of TDDFT excitation energies on structural parameters (PDF)

- **AUTHOR INFORMATION**

  - **Corresponding Authors**
    E-mail: benedetta mennucci@unipi.it.
    E-mail: c.filippi@utwente.nl.

  - **Notes**
    The authors declare no competing financial interest.

- **ACKNOWLEDGMENTS**

  R.G. is supported by an ECHO Grant (712.012.005) of The Netherlands Organisation for Scientific Research (NWO), and we received support from NWO for the use of the SARA supercomputer facilities. B.M. acknowledges the European Research Council (ERC) in the framework of the Starting Grant (EnLight-277755) and the University of Pisa under Project PRA-2016-46 for financial support. C.C. acknowledges support from the Spanish Ministerio de Economía y Competitividad (RYC2011-08918) and the Agencia de Gestión d’Ajusts Universitaris i de Recerca from Generalitat de Catalunya (SGR2014-1189).

- **REFERENCES**

cis-, 9-cis-, or 13-cis-Retinal in the Eye?


