Induction effects on the absorption maxima of photoreceptor proteins

Church, Jonathan R.; Olsen, Jógvan Magnus Haugaard; Schapiro, Igor

Published in:
Biophysics and physicobiology

Link to article, DOI:
10.2142/biophysico.bppb-v20.s007

Publication date:
2023

Document Version
Publisher's PDF, also known as Version of record

Vol. 20, e201007 (2023)
doi: 10.2142/biophysico.bppb-v20.s007

Biophysics and Physicobiology
https://www.jstage.jst.go.jp/browse/biophysico/

Special Issue: Recent Advances in Retinal Protein Research

Regular Article (Invited)

Induction effects on the absorption maxima of photoreceptor proteins

Jonathan R. Church[1], Jógvan Magnus Haugaard Olsen[2], Igor Schapiro[1]

[1] *Fritz Haber Center for Molecular Dynamics Research, Institute of Chemistry, The Hebrew University of Jerusalem, Jerusalem 9190401, Israel*
[2] *DTU Chemistry, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark*

Received December 12, 2022; Accepted January 20, 2023;
Released online in J-STAGE as advance publication January 24, 2023
Edited by Yuki Sudo

Multiscale simulations have been established as a powerful tool to calculate and predict excitation energies in complex systems such as photoreceptor proteins. In these simulations the chromophore is typically treated using quantum mechanical (QM) methods while the protein and surrounding environment are described by a classical molecular mechanics (MM) force field. The electrostatic interactions between these regions are often treated using electrostatic embedding where the point charges in the MM region polarize the QM region. A more sophisticated treatment accounts also for the polarization of the MM region. In this work, the effect of such a polarizable embedding on excitation energies was benchmarked and compared to electrostatic embedding. This was done for two different proteins, the lipid membrane-embedded jumping spider rhodopsin and the soluble cyanobacteriochrome Slr1393g3. It was found that the polarizable embedding scheme produces absorption maxima closer to experimental values. The polarizable embedding scheme was also benchmarked against expanded QM regions and found to be in qualitative agreement. Treating individual residues as polarizable recovered between 50% and 71% of the QM improvement in the excitation energies, depending on the system. A detailed analysis of each amino acid residue in the chromophore binding pocket revealed that aromatic residues result in the largest change in excitation energy compared to the electrostatic embedding. Furthermore, the computational efficiency of polarizable embedding allowed it to go beyond the binding pocket and describe a larger portion of the environment, further improving the results.

Key words: photoreceptor proteins, QM/MM, embedding scheme, rhodopsin, cyanbacteriochrome

Significance

Multiscale QM/MM simulations are a power tool for studying systems which cannot be modelled using a full QM treatment. Often interactions between the QM and MM regions are treated using electrostatic embedding. In this embedding scheme the environment is modelled as electrostatic point charges surrounding the QM region. A weakness of this scheme is that it cannot account for mutual polarization between the QM and MM regions as well as interactions within the environment itself. In this work we show how employing an embedding scheme which accounts for induction effects can yield a more complete physical picture.

Introduction

Photoreceptor proteins are light-sensitive proteins that play a role in light sensing for a variety of organisms [1]. These systems gain their light sensitivity through highly conjugated chromophores embedded within the protein which can

Corresponding author: Igor Shapiro, Fritz Haber Center for Molecular Dynamics Research, Institute of Chemistry, The Hebrew University of Jerusalem, Jerusalem 9190401, Israel. ORCID iD: https://orcid.org/0000-0001-8536-6869, e-mail: Igor.Schapiro@mail.huji.ac.il

absorb light at a variety of wavelengths. Following light absorption, these proteins produce a biological response in the host organism and have been found to play roles in several biological functions, including regulation of circadian rhythms, phototaxis, and light-oriented growth in plants. Photoreceptor proteins are used in biomedical applications, such as optogenetics, where they can be used as light switches in living organisms [2]. Further utilization of these proteins for various applications can be achieved by tailoring the optical properties for specific applications. The optical properties are determined by the chromophore-protein interactions; however, an atomistic understanding of the precise spectral tuning mechanism is not fully understood. Multiscale simulations can contribute to the understanding of these mechanisms [3]. In such simulations, the photoreceptor protein is divided into two different regions which are treated at different levels of theory. A quantum mechanical (QM) description is applied to the chromophore and important surrounding amino acids. The remaining part is treated using molecular mechanics (MM) which can describe hundreds of thousands of atoms by parametrized force fields. This particular multiscale method is called hybrid quantum mechanics/molecular mechanics (QM/MM) [4]. This hybrid method allows one to accurately model the chromophore and its surrounding environment while remaining computationally feasible [5–7].

Setting up a QM/MM simulation requires several technical considerations. These include how to appropriately partition the QM and MM regions and what type of treatment should be used when describing the interactions between them. The effect of the MM environment on the electron density in the QM region is included through an embedding potential. For these potentials there are two important aspects: (1) the type of embedding, i.e., how is the interaction between the QM and MM region treated, and (2) the quality of the embedding parameters that describe the MM region. The simplest form of embedding is mechanical embedding where the QM/MM interactions are treated purely classically using either force-field or QM-derived charges for the QM region. Thus, in this scheme, the electron density of the QM region is not directly affected by the MM region making it unsuitable for calculating optical properties. Electrostatic embedding (EE) is a popular embedding scheme that uses static atom-centered point charges, typically from one of the popular force fields [3,8,9]. Here the charges in the MM region polarize the electron density in the QM region. However, a weakness of this scheme is that it cannot account for induction effects within the environment or between the QM and MM regions, i.e., the MM region is static. These induction effects are included in the more advanced polarized embedding (PE) schemes [3,8–11]. Mennucci et al. have suggested that inclusion of induction effects through polarizable schemes may be the future gold standard when modelling complex biochemical systems [12]. One approach is the fragment-based polarizable embedding (FBPE) method which was developed with a focus on accurate calculations of transition and response properties [13–16]. The FBPE method uses fragment-based atom-centered multipole expansions to model electrostatic interactions and through fragment-based atom-centered polarizabilities it allows mutual polarization between the QM and MM regions as well as within the MM region itself. Including these interactions were found to produce substantially better quality embedding potentials when compared to an MP2 reference than the classical AMBER and GAFF force fields for insulin [17]. For larger systems, one drawback of the FBPE method is that it requires additional QM calculations on the residues being considered in the environment. However, cost-effective standard potentials have recently been developed to help reduce the cost of these calculations [18].

In this work we will assess the effect of the FBPE scheme on two photoreceptor proteins (Figure 1). These include jumping spider rhodopsin-1 (JSR1, PDB ID: 6I9K) as well as the red-light absorbing reactant (Pr, PDB ID: 5DFX) and green-light absorbing photoproduct (Pg, PDB ID: 5M82) of Slr1393g3 cyanobacteriochrome (CBCR) [19–21]. For the rest of the paper, we refer to these two states of Slr1393g3 as Pg-Slr and Pr-Slr.

The chromophores of rhodopsins and CBCRs undergo an intramolecular charge transfer upon excitation, which makes them ideal candidates to study the effect of using a PE scheme. Rhodopsins contain a positively charged retinal chromophore linked to the opsin through a lysine side chain, which together form a retinal protonated Schiff base (RPSB). These proteins have been found to absorb over a wide range of wavelengths, ranging from 400 nm to 690 nm [22–26]. There is a redistribution of the positive charge in the chromophore upon excitation to the first excited state. This charge redistribution enables the protein environment to greatly influence the absorption maximum of RPSB because the ground and the excited state have different charge distributions [27–30]. The extent of this spectral tuning is controlled by the protein environment, which includes the protonation states of titratable residues and the interactions with nearby residues within the binding pocket of the protein. JSR1 has been experimentally measured to absorb at 505 nm while a hybrid QM/MM simulation on JSR1 using the EE scheme with TD-CAM-B3LYP/cc-pVDZ resulted in excitation energy of 484 nm [21,31–33].

CBCRs also have the ability to absorb light over a wide range of wavelengths, from near-infrared all the way to near-ultraviolet [34–38]. Here the spectral tuning is thought to be due to structural changes in the phycocyanobilin (PCB) chromophore as well as from interactions with its local environment [19,39,40]. The red-light absorbing Pr-Slr has been experimentally measured to absorb at 649 nm [41]. Upon illumination the chromophore undergoes an isomerization around the C-D rings leading to Pg-Slr, which absorbs at 537 nm [41]. The calculated absorption maxima were determined to be 607 nm and 546 nm respectively using EE with RI-ADC(2)/cc-pVDZ [40]. The experimental shift of 0.40 eV

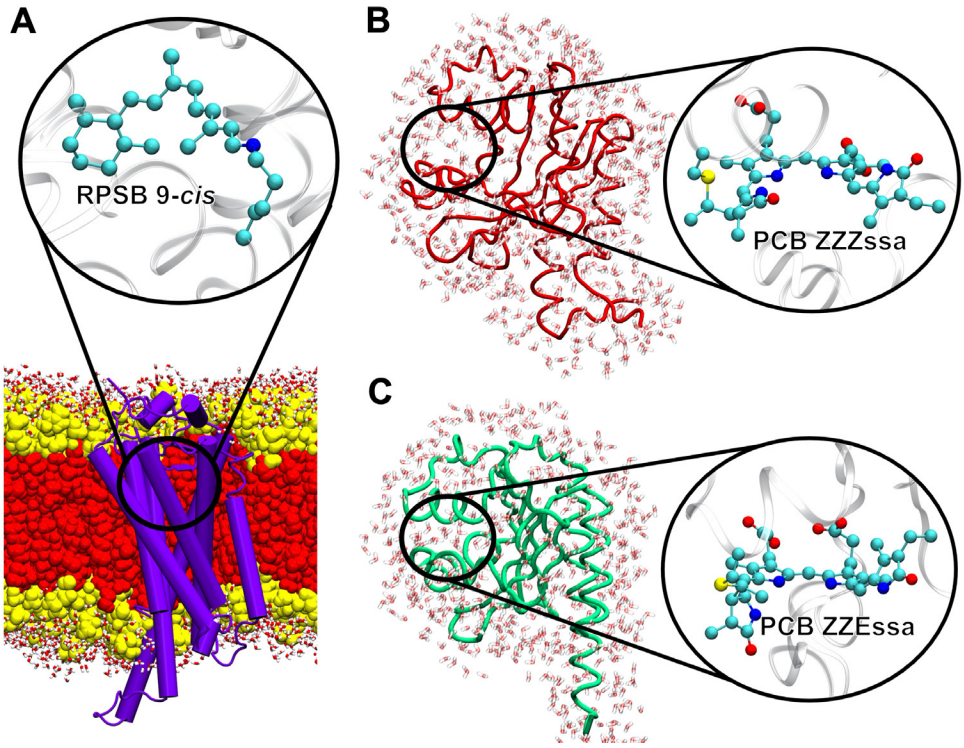

Figure 1 Photoreceptor proteins that were studied. **(A)** JSR1, **(B)** Pr-Slr, and **(C)** Pg-Slr. The inset shows the retinal protonated Schiff base (RPSB) for JSR1, while the phycocyanobilin (PCB) chromophore is presented for Pg-Slr and Pr-Slr.

between the Pg-Slr and Pr-Slr absorption maxima was partially reproduced using simulations with a value of 0.23 eV [40].

The results from using FBPE applied to these proteins are presented below. How this scheme improves the calculated absorption maxima is examined and compared to previous EE results. In particular, the amino acids that benefit the most from this improved description are identified.

Materials and Methods

Spectra Generation

The structures for the FBPE calculations were adapted from previous studies, JSR1, Pg-Slr, and Pr-Slr [31,39,40]. The full details can be found in the respective references. Briefly, 100 structures were generated for each protein by running a QM/MM trajectory for 1 ns and taking snapshots every 10 ps. The excitation energies of these snapshots were then calculated at the TD-CAM-B3LYP/cc-pVDZ level of theory using either the EE scheme with the AMBER ff14SB forcefield or the FBPE scheme with potentials generated using PyFraME and the Dalton program [42–45]. The QM-MM partitioning of JSR1 was done by cutting the retinal binding lysine residue between the C_α-C_β bond. For Pg-Slr and Pr-Slr, the boundary was placed at the C_α-C_β bond of the cysteine that binds the PCB chromophore. The polarizable embedding potentials were generated using PyFraME by explicitly deriving the atom-centered multipoles (up to and including quadrupoles) and dipole-dipole polarizabilities from individual fragment calculations [42]. The multipoles and polarizabilities were derived for each snapshot by using a fragmentation scheme as outlined in ref [16]. The individual fragment calculations were performed using CAM-B3LYP together with an ANO-type recontracted 6-31+G* basis set (named loprop-6-31+G* in Dalton). For each system, water molecules within 10 Å of the chromophore were calculated explicitly. Convergence analysis of the excitation energy with respect to the size of the included environment were performed on the first snapshot of JSR1, Pr-Slr, and Pg-Slr due to the large external environments of solvent surrounding the proteins. Here the additional solvent, including lipids, ions and water were treated with the appropriate standard potentials found within the PyFraME package [18,46,47]. The radius of the solvent treated as polarizable was gradually increased while the remaining solvent up to 50 Å was treated with point charges only (Supplementary Figure S1). It was

found that the excitation energy of Pg-Slr and Pr-Slr converged within 20 Å, while JSR1 did not converge until close to 50 Å. The difference in convergence is likely due to the heterogenous and more polarizable environment of JSR1 compared to Slr1393g3. Because rhodopsins are transmembrane proteins, JSR1 was modelled inside a lipid bilayer with water and ions on either side of the membrane. Slr1393g3 on the other hand is solvated in water. The heterogeneity and polarizability of the environment in the JSR1 model made the calculation more sensitive to the amount of environment included in the FBPE potential. For each photoreceptor, the excitation energies of the 100 snapshots were then determined by treating the environment using electrostatic embedding and then again using the polarizable embedding potentials with the cutoffs determined above. The absorption maximum of each system was determined by broadening the stick spectra of each snapshot assuming a Gaussian lineshape with a σ value of 0.15 eV and then averaging over the 100 configurations at each wavelength.

Subset Analysis

A subset of 10 evenly spaced snapshots were then selected by sampling every 100 ps. The side chains of residues within 3.5 Å of the chromophore were then treated individually as polarizable while the remaining environment was treated as static with point charges. The excitation energies of each subset when treating these residues as polarizable were then determined and compared to the values obtained using pure EE (Supplementary Tables S1-S3). Finally, these same residues were then placed inside the QM region for comparison to the FBPE results (Supplementary Tables S4-S6). For a graphical representation of the various embedding treatments of the environment please see Supplementary Figure S2.

Side chains that were found to change the excitation energies by more than ± 0.01 eV for more than one of the 10 snapshots were then grouped and treated together with the PE scheme (denoted expanded FBPE) and the excitation energies were calculated. These same residues were again treated together in an expanded QM calculation for comparison. The character of the excited states was found to change in some cases when using the standard CAM-B3LYP functional due to the expanded size of the QM region. For range-separated hybrid functionals, such as CAM-B3LYP, the amount of Hartree-Fock exchange included depends on the distance (r_{12}) in the exchange operator. The artifacts in the description of the excited states were alleviated by setting the fraction of Hartree-Fock (HF) exchange at long-range (LR) to 1.0 (corresponding to β=0.81) instead of the default of 0.65 (corresponding to β=0.46). This modified functional is sometimes referred to as fully LR corrected CAM-B3LYP.

Results and Discussion

Absorption Maxima

The results of the computed absorption maxima from using the EE and FBPE schemes, as defined in the methodology, are summarized in Table 1 and Figure 2. Two trends emerged from the analysis. First, the absorption maximum of each photoreceptor protein calculated with the EE scheme was higher in energy compared to those obtained with the FBPE scheme (Table 1 and Figure 2). Second, the fully long-range corrected CAM-B3LYP functional (β=0.81) produced absorption maxima that are higher in energy than those calculated with the standard amount of HF exchange at long-range (β=0.46). The absorption maxima obtained with the FBPE scheme were found to be in closer agreement with their corresponding experimental value. The shift between the absorption maxima of Pg-Slr and Pr-Slr from using FBPE was determined to be 0.25 eV and 0.27 eV with β=0.46 and β=0.81, respectively. When using EE with standard CAM-B3LYP, the deviation in the calculated absorption maxima to the experimental values ranged from 0.11 eV to 0.36 eV, while the differences from FBPE were lower and spanned from 0.04 eV to 0.22 eV.

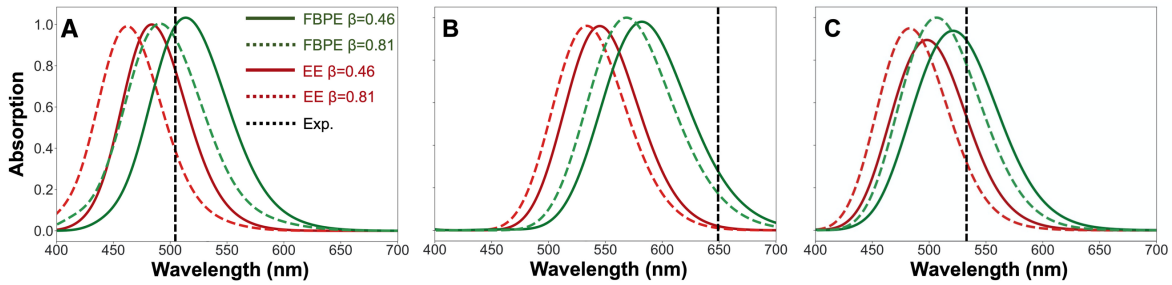

Figure 2 Computed absorption spectra for **(A)** JSR1, **(B)** Pr-Slr, and **(C)** Pg-Slr calculated using 100 snapshots with either the FBPE (green) or EE (red) schemes at the TD-CAM-B3LYP/cc-pVDZ level of theory. For each embedding method, results from two LR HF exchange fractions are included, β=0.46 (solid) and β=0.81 (dashed). The experimental absorption maximum as a black dashed line.

Table 1 Absorption maxima of the two photoreceptor proteins studied in this work

Method	JSR1		Pr-Slr		Pg-Slr		(Pg-Pr)-Slr	
	nm	eV	nm	eV	nm	eV	nm	eV
EE ($\beta = 0.46$)	484 ± 4 (-21)	2.56 ± 0.02 (0.11)	545 ± 4 (-104)	2.28 ± 0.02 (0.36)	498 ± 5 (-39)	2.49 ± 0.02 (0.18)	-47 ± 7 (65)	0.21 ± 0.03 (-0.19)
EE ($\beta = 0.81$)	462 ± 5 (-43)	2.68 ± 0.03 (0.23)	534 ± 4 (-115)	2.32 ± 0.02 (0.41)	483 ± 5 (-54)	2.57 ± 0.02 (0.26)	-51 ± 7 (61)	0.25 ± 0.03 (-0.15)
FBPE ($\beta = 0.46$)	**513 ± 5 (8)**	**2.42 ± 0.02 (-0.04)**	**582 ± 6 (-67)**	**2.13 ± 0.02 (0.22)**	**521 ± 6 (-16)**	**2.38 ± 0.03 (0.07)**	-61± 9 (51)	0.25± 0.04 (-0.15)
FBPE ($\beta = 0.81$)	491 ± 6 (-13.6)	2.52 ± 0.03 (0.07)	569 ± 5 (-82)	2.18 ± 0.02 (0.28)	506 ± 5 (-31)	2.45 ± 0.02 (0.14)	**-63± 7 (49)**	**0.27 ± 0.03 (-0.13)**
Experimental	505[20]	2.46	649[41]	1.91	537[41]	2.31	-112	0.40

Results for the EE and FBPE schemes are reported in nm and eV. Results from the fully LR corrected CAM-B3LYP (β=0.81) and standard version (β=0.46) are tabulated. The confidence intervals at 95% are shown next to each absorption maximum value. Values in parenthesis show the absolute error to the experimental absorption maximum. Bolded terms show the method with the lowest error in comparison to the experimental value for each photoreceptor protein.

Although the FBPE scheme was found to produce absorption maxima in closer agreement to experiment, the origin behind the change in the excitation energies was not immediately clear. To analyze this, the average energy (E_{avg}) of the ground (S_0) and first singlet excited state (S_1) were computed (see Table 2). Because the FBPE scheme accounts for additional interactions which are neglected in EE, E_{avg} was determined only from the polarized QM density to make the results from the two embedding schemes comparable. Taking the difference between E_{avg} obtained from both embedding schemes (ΔE (FBPE-EE)) shows that both electronic states are destabilized upon switching to the FBPE scheme. However, the destabilization of S_1 was found to be less than that of S_0, leading to the lower energy absorption maxima in Figure 2. The destabilization of both electronic states originates from missing interactions not described in EE. These interactions can have a greater influence on the ground or excited state depending on their respective charge distribution. For example, upon excitation RPSB undergoes a rapid charge redistribution. Here the positive charge moves from the Schiff base towards the β-ionone ring [28–30,48,49]. Using the EE scheme, Fujimoto and coworkers found that the electronic states of RPSB are stabilized to different degrees due to this charge redistribution and electrostatic interactions with the environment [48,49]. Likewise, here it was found that the additional induction effects of FBPE change the stability of the electronic states to different degrees for each chromophore studied.

Table 2 Average energies (E_{avg}) in Hartree when neglecting the embedding energy, for the S_0 and S_1 singlet states from the 100 structures sampled for each photoreceptor

	JSR1		Pr-Slr		Pg-Slr	
$\beta = 0.46$	S_0	S_1	S_0	S_1	S_0	S_1
E_{avg}-EE (a.u.)	-991.260 ± 0.002	-991.166 ± 0.002	-2388.309 ± 0.003	-2388.225 ± 0.003	-2388.313 ± 0.004	-2388.221 ± 0.004
E_{avg}-FBPE (a.u.)	-991.229 ± 0.002	-991.140 ± 0.002	-2388.236 ± 0.007	-2388.159 ± 0.008	-2388.242 ± 0.005	-2388.154 ± 0.005
ΔE (FBPE-EE) (eV)	0.85 ± 0.09	0.71 ± 0.08	1.99 ± 0.21	1.81 ± 0.23	1.93 ± 0.17	1.82 ± 0.16
$\beta = 0.81$	S_0	S_1	S_0	S_1	S_0	S_1
E_{avg}-EE (a.u.)	-990.443 ± 0.002	-990.344 ± 0.002	-2386.770 ± 0.004	-2386.685 ± 0.004	-2386.778 ± 0.005	-2386.683 ± 0.005
E_{avg}-FBPE (a.u.)	-990.412 ± 0.002	-990.319 ± 0.002	-2386.693 ± 0.005	-2386.613 ± 0.005	-2386.707 ± 0.004	-2386.617 ± 0.004
ΔE (FBPE-EE) (eV)	0.84 ± 0.09	0.68 ± 0.08	2.10 ± 0.16	1.95 ± 0.16	1.91 ± 0.16	1.79 ± 0.16

The difference between E_{avg} obtained using both embedding schemes is also presented in Hartree to show how switching between the embedding scheme impacts the chromophore. Also shown are the confidence intervals at 95%.

Single Side Chain Effects

The FBPE scheme provides a more physically complete description of chromophore-protein interactions in comparison to EE. However, to understand this improvement on the atomistic level, the effect of individual amino acid side chains was probed. For a subset of 10 equally spaced snapshots, the individual side chains of the chromophore binding pocket were treated as polarizable while the remaining protein was modeled by static point charges only. The residues for this in-depth analysis were selected if at least one atom was within 3.5 Å of any atom of the chromophore. These residues are considered to be part of the chromophore binding pocket. The change in the excitation energy relative to a full EE treatment was then determined for each residue (Supplementary Tables S1-S3). Since the binding pockets are comprised of 9-21 amino acids, the contributions were grouped into four categories: aromatic, charged, polar, and non-polar residues (Figure 3). The same procedure was then repeated treating the same residues as part of the QM region. The results are tabulated in Table 3.

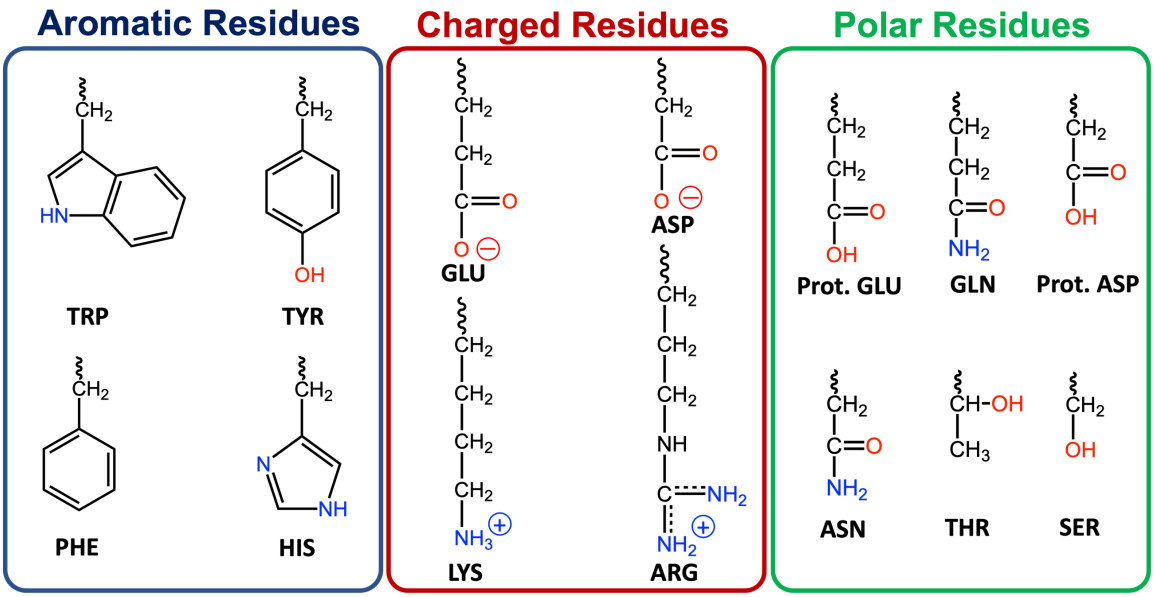

Figure 3 Schematic showing how side chains were grouped as either aromatic, charged, or polar to analyze the impact of each group on the excitation energy when using the polarizable embedding scheme.

Table 3 Effect of the aromatic, charged, polar, and non-polar amino acid side chains on the excitation energy in eV

FBPE	**JSR1**	**Pr-Slr**	**Pg-Slr**
Aromatic	-0.074 ± 0.006	-0.061 ± 0.006	-0.037 ± 0.006
Charged	0.009 ± 0.002	-0.013 ± 0.009	-0.004 ± 0.004
Polar	-0.020 ± 0.005	-0.013 ± 0.002	-0.007 ± 0.002
Non-polar	-0.005 ± 0.006	-0.016 ± 0.002	-0.010 ± 0.002
QM	**JSR1**	**Pr-Slr**	**Pg-Slr**
Aromatic	-0.072 ± 0.006	-0.092 ± 0.012	-0.063 ± 0.010
Charged	0.002 ± 0.001	-0.016 ± 0.004	-0.014 ± 0.004
Polar	-0.026 ± 0.007	-0.014 ± 0.002	-0.007 ± 0.001
Non-polar	-0.025 ± 0.006	-0.014 ± 0.001	-0.012 ± 0.002

Confidence intervals at 95% are also reported.

When treating individual residues either as polarizable or quantum mechanically, the largest change in the excitation energies relative to the EE scheme originated from aromatic residues. These amino acids have larger polarizabilities that upon being polarized can induce larger polarization effects in the chromophore as well as the surrounding protein environment. The side chains accounting for the largest changes in the excitation energy for each photoreceptor were: Y293 and W290 in JSR1, H529 and W496 in Pr-Slr, as well as H529 and F474 in Pg-Slr. These residues are in proximity to the conjugated π system of the chromophores (see Figure 4) enabling stronger polarizing interactions. This highlights an important effect neglected in the EE scheme. Treating the polarizable side chains of the binding pocket with FBPE and including them in the QM region produced similar changes in the excitation energies from those obtained with the EE scheme.

Expanded FBPE and QM Regions

A simultaneous treatment of all the residues comprising the chromophore binding pocket using the FBPE scheme is desirable to capture as many of the chromophore/protein interactions as possible. However, the reference value from an expanded QM region would then be computationally intractable. Therefore, a subset of side chains was selected which comprised residues that were found to alter the excitation energy by more than ± 0.01 eV. This subset of residues was included in the expanded FBPE and QM calculations (Figure 4). Selecting the important side chains of the binding pockets amounted to 21 residues (322 atoms including the RPSB) for JSR1 as well as 9 residues and 11 residues (205 atoms and 236 atoms including PCB) for Pr-Slr and Pg-Slr, respectively.

The resulting changes in the excitation energies for the expanded QM region and expanded FBPE region are presented in Figure 5. Collective treatment of the binding pocket residues with FBPE and QM produced the same trend, namely the lowering of the excitation energies for each snapshot (Supplementary Tables S7-S9). The lowering of the excitation energy was consistently larger for the expanded QM region than for the expanded FBPE. The FBPE treatment recovered 71%, 70% and 50% of the change for JSR1, Pr-Slr and Pg-Slr, respectively. The only exception was found for a single snapshot of Pg-Slr. However, in this snapshot the QM correction was the smallest among each photoreceptor protein and the correction from FBPE was close to zero. Therefore, induction effects in this snapshot can be considered small.

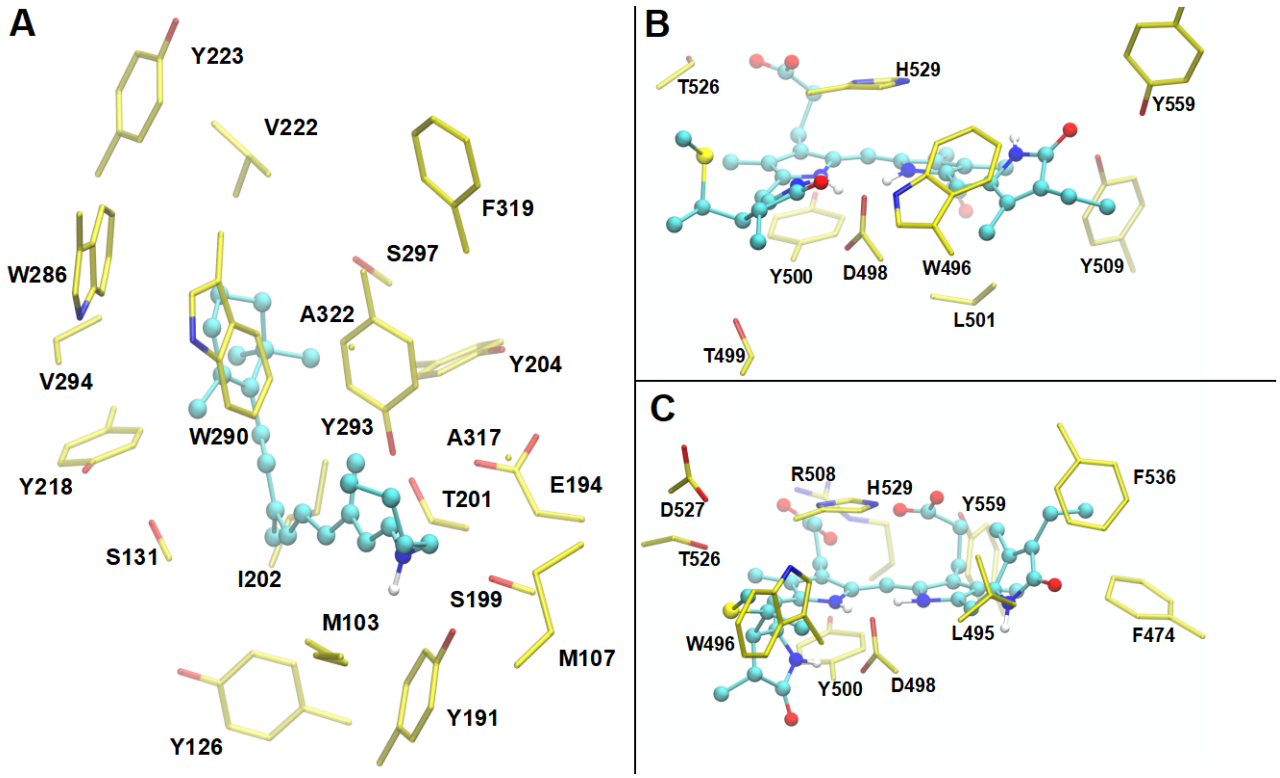

Figure 4 The side chains that were determined to cause a change in the excitation energy for at least one snapshot **(A)** JSR1, **(B)** Pr-Slr, and **(C)** Pg-Slr. Here the RPSB and PCB chromophores are shown in cyan and side chains in yellow.

Unfortunately, it is not currently feasible to study long-range effects from treating the entire protein using QM methods. Instead, the expanded FBPE results were compared to treating the full system with FBPE. The cutoff radius for the full treatment was converged at a distance of 50 Å for JSR1 and 20 Å for the two forms of Slr1393g3 (Supplementary Figure S1). The full FBPE scheme (Figure 5) produced larger changes in the excitation energies than both expanded FBPE and expanded QM. The average change in the expanded FBPE excitation energies were found to reproduce 49%, 45%, and 28% of the change from the full FBPE treatment for JSR1, Pr-Slr and Pg-Slr, respectively. These results highlight two important features. First, the inclusion of induction effects is an important feature missed when treating the environment with a static point charge model. These results are in agreement with other studies in the literature [15,50]. Second, long-range effects are important as they further decrease the excitation energies in our three examples.

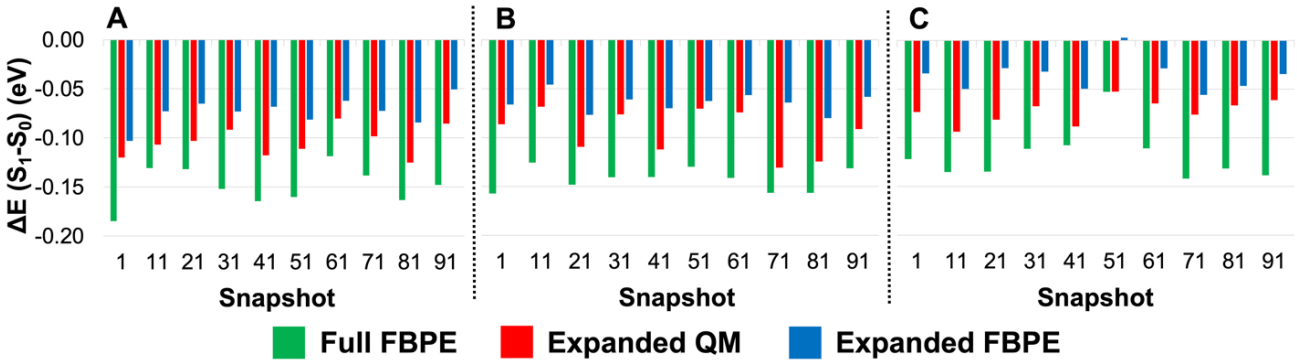

Figure 5 Changes in the excitation energies of the ten snapshots from **(A)** JSR1, **(B)** Pr-Slr, and **(C)** Pg-Slr relative to full EE when treating the entire system with FBPE (green), an expanded QM region (red), and an expanded FBPE region (blue).

Conclusion

In this work we assessed the effect of FBPE by calculating the excitation energy of two different photoreceptor proteins. The FBPE scheme provides results in better agreement with experimental values compared to the EE scheme. In the FBPE scheme, residues are treated as polarizable which yields a more realistic environment surrounding the chromophore compared to the static point charges in EE. Analysis of individual side chains revealed that aromatic residues have the largest effect on the excitation energy. These findings were corroborated by treating these aromatic residues in an expanded QM region. Overall, the FBPE scheme was able to produce absorption maxima that were more consistent with treating the environment near the chromophore with higher level QM methods. Moreover, the computational efficiency of the FBPE scheme allows treating the amino acids beyond the binding pocket which is not feasible at the QM level. The results with full FBPE show that an improved description of residues outside the binding pocket brings the calculated absorption maxima in closer agreement to experimental values. The systematic comparison of the FBPE results to a QM treatment show that the improvement is not a coincidence.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

I.S. initially conceived this project and acted in a supervising role. J.R.C. and J.M.H.O generated the potential files used to generate the absorption spectra. J.R.C. wrote the initial draft of the paper, and all authors contributed to discussions and editing of the final draft.

Data Availability

The data generated and analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgements

I.S. thanks the DFG Collaborative Research Center 1078, project C6 for support. I.S. gratefully acknowledges funding by the Israel Science Foundation (Research Center grant no. 3131/20). J.R.C thanks the Zuckerman STEM Leadership Program for their support. J.M.H.O. thanks VILLUM FONDEN for financial support (grant no. 29478).

References

[1] Van Der Horst, M. A., Hellingwerf, K. J. Photoreceptor Proteins, 'Star Actors of Modern Times': A Review of the Functional Dynamics in the Structure of Representative Members of Six Different Photoreceptor Families. Acc. Chem. Res. 37, 13–20 (2004). https://doi.org/10.1021/ar020219d

[2] Hegemann, P., Nagel, G. From channelrhodopsins to optogenetics. EMBO Mol. Med. 5, 173–176 (2013). https://doi.org/10.1002/emmm.201202387

[3] Senn, H. M., Thiel, W. QM/MM methods for biomolecular systems. Angew. Chem. Int. Ed. Engl. 48, 1198–1229 (2009). https://doi.org/10.1002/anie.200802019

[4] Warshel, A., Levitt, M. Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. J. Mol. Biol. 103, 227–249 (1976). https://doi.org/10.1016/0022-2836(76)90311-9

[5] Warshel, A. Multiscale modeling of biological functions: From enzymes to molecular machines (nobel lecture). Angew. Chem. Int. Ed. Engl. 53, 10020–10031 (2014). https://doi.org/10.1002/anie.201403689

[6] Karplus, M. Development of multiscale models for complex chemical systems: From $H+H_2$ to biomolecules (nobel lecture). Angew. Chem. Int. Ed. Engl. 53, 9992–10005 (2014). https://doi.org/10.1002/anie.201403924

[7] Levitt, M. Birth and future of multiscale modeling for macromolecular systems (Nobel Lecture). Angew. Chem. Int. Ed. Engl. 53, 10006–10018 (2014). https://doi.org/10.1002/anie.201403691

[8] Brunk, E., Rothlisberger, U. Mixed quantum mechanical/molecular mechanical molecular dynamics simulations of biological systems in ground and electronically excited States. Chem. Rev. 115, 6217–6263 (2015). https://doi.org/10.1021/cr500628b

[9] Lipparini, F., Mennucci, B. Hybrid QM/classical models: Methodological advances and new applications. Chem. Phys. Rev. 2, 041303 (2021). https://doi.org/10.1063/5.0064075

[10] Lipparini, F., Cappelli, C., Barone, V. Linear response theory and electronic transition energies for a fully polarizable QM/classical hamiltonian. J. Chem. Theory Comput. 8, 4153–4165 (2012). https://doi.org/10.1021/ct3005062

[11] Gurunathan, P. K., Acharya, A., Ghosh, D., Kosenkov, D., Kaliman, I., Shao, Y., et al. Extension of the Effective Fragment Potential Method to Macromolecules. J. Phys. Chem. B 120, 6562–6574 (2016). https://doi.org/10.1021/acs.jpcb.6b04166

[12] Bondanza, M., Nottoli, M., Cupellini, L., Lipparini, F., Mennucci, B. Polarizable embedding QM/MM: The future gold standard for complex (bio)systems? Phys. Chem. Chem. Phys. 22, 14433–14448 (2020). https://doi.org/10.1039/d0cp02119a

[13] Olsen, J. M., Aidas, K., Kongsted, J. Excited states in solution through polarizable embedding. J. Chem. Theory Comput. 6, 3721–3734 (2010). https://doi.org/10.1021/ct1003803

[14] Olsen, J. M. H., Kongsted, J. Molecular properties through polarizable embedding. Adv. Quantum Chem. 61, 107–143 (2011). https://doi.org/10.1016/B978-0-12-386013-2.00003-6

[15] List, N. H., Olsen, J. M. H., Kongsted, J. Excited states in large molecular systems through polarizable embedding. Phys. Chem. Chem. Phys. 18, 20234–20250 (2016). https://doi.org/10.1039/c6cp03834d

[16] Steinmann, C., Reinholdt, P., Nørby, M. S., Kongsted, J., Olsen, J. M. H. Response properties of embedded molecules through the polarizable embedding model. Int. J. Quantum Chem. 119, e25717 (2019). https://doi.org/10.1002/qua.25717

[17] Olsen, J. M. H., List, N. H., Kristensen, K., Kongsted, J. Accuracy of protein embedding potentials: An analysis in terms of electrostatic potentials. J. Chem. Theory Comput. 11, 1832–1842 (2015). https://doi.org/10.1021/acs.jctc.5b00078

[18] Reinholdt, P., Kjellgren, E. R., Steinmann, C., Olsen, J. M. H. Cost-effective potential for accurate polarizable embedding calculations in protein environments. J. Chem. Theory Comput. 16, 1162–1174 (2020). https://doi.org/10.1021/acs.jctc.9b00616

[19] Xu, X., Port, A., Wiebeler, C., Zhao, K. H., Schapiro, I., Gärtner, W. Structural elements regulating the photochromicity in a cyanobacteriochrome. Proc. Natl. Acad. Sci. U.S.A. 117, 2432–2440 (2020). https://doi.org/10.1073/pnas.1910208117

[20] Varma, N., Mutt, E., Mühle, J., Panneels, V., Terakita, A., Deupi, X., et al. Crystal structure of jumping spider

rhodopsin-1 as a light sensitive GPCR. Proc. Natl. Acad. Sci. U.S.A. 116, 14547–14556 (2019). https://doi.org/10.1073/pnas.1902192116

[21] Ran, T., Ozorowski, G., Gao, Y., Sineshchekov, O. A., Wang, W., Spudich, J. L., et al. Cross-protomer interaction with the photoactive site in oligomeric proteorhodopsin complexes. Acta Crystallogr. Sect. D Biol. Crystallogr. 69, 1965–1980 (2013). https://doi.org/10.1107/S0907444913017575

[22] Wang, W., Nossoni, Z., Berbasova, T., Watson, C. T., Yapici, I., Lee, K. S. S., et al. Tuning the electronic absorption of protein-embedded all-trans-retinal. Science 338, 1340–1343 (2012). https://doi.org/10.1126/science.1226135

[23] Broser, M., Spreen, A., Konold, P. E., Peter, E., Adam, S., Borin, V., et al. NeoR, a near-infrared absorbing rhodopsin. Nat. Commun. 11, 1–10 (2020). https://doi.org/10.1038/s41467-020-19375-8

[24] Béja, O., Spudich, E. N., Spudich, J. L., Leclerc, M., DeLong, E. F. Proteorhodopsin phototrophy in the ocean. Nature 411, 786–789 (2001). https://doi.org/10.1038/35081051

[25] Bogomolni, R. A., Spudich, J. L. The photochemical reactions of bacterial sensory rhodopsin-I. Flash photolysis study in the one microsecond to eight second time window. Biophys. J. 52, 1071–1075 (1987). https://doi.org/10.1016/S0006-3495(87)83301-5

[26] Karasuyama, M., Inoue, K., Nakamura, R., Kandori, H., Takeuchi, I. Understanding colour tuning rules and predicting absorption wavelengths of microbial rhodopsins by data-driven machine-learning approach. Sci. Rep. 8, 1–11 (2018). https://doi.org/10.1038/s41598-018-33984-w

[27] Toker, Y., Langeland, J., Gruber, E., Kjær, C., Nielsen, S. B., Andersen, L. H., et al. Counterion-controlled spectral tuning of the protonated Schiff-base retinal. Phys. Rev. A 98, (2018). https://doi.org/10.1103/PhysRevA.98.043428

[28] Barry, H., Uri, D., Koji, N., Valeria, B. N., Gawinowicz, M. A., Maria, A., et al. An external point-charge model for wavelength regulation in visual pigments. J. Am. Chem. Soc. 101, 7084–7086 (1979). https://doi.org/10.1021/ja00517a060

[29] Nakanishi, K., Balogh-Nair, V., Amaboldi, M., Tsujimoto, K., Honig, B. An external point-charge model for bacteriorhodopsin to account for its purple color. J. Am. Chem. Soc. 102, 7945–7947 (1980). https://doi.org/10.1021/ja00547a028

[30] Mordechai, S., Koji, N., Barry, H. Through-Space Electrostatic Effects in Electronic Spectra. Experimental evidence for the external point-charge model of visual pigments. J. Am. Chem. Soc. 101, 7086–7088 (1979). https://doi.org/10.1021/ja00517a061

[31] Church, J. R., Olsen, J. M. H., Schapiro, I. The impact of retinal configuration on the protein–chromophore interactions in bistable jumping spider rhodopsin-1. Molecules 27, 71 (2022). https://doi.org/10.3390/molecules27010071

[32] Ehrenberg, D., Varma, N., Deupi, X., Koyanagi, M., Terakita, A., Schertler, G. F. X., et al. The two-photon reversible reaction of the bistable jumping spider rhodopsin-1. Biophys. J. 116, 1248–1258 (2019). https://doi.org/10.1016/j.bpj.2019.02.025

[33] Church, J. R., Amoyal, G. S., Borin, V. A., Adam, S., Olsen, J. M. H., Schapiro, I. Deciphering the Spectral Tuning Mechanism in Proteorhodopsin: The Dominant Role of Electrostatics Instead of Chromophore Geometry. Chem. - A Eur. J. 28, e202200139 (2022). https://doi.org/10.1002/chem.202200139

[34] Ikeuchi, M., Ishizuka, T. Cyanobacteriochromes: A new superfamily of tetrapyrrole-binding photoreceptors in cyanobacteria. Photochem. Photobiol. Sci. 7, 1159–1167 (2008). https://doi.org/10.1039/b802660m

[35] Rockwell, N. C., Martin, S. S., Gulevich, A. G., Lagarias, J. C. Phycoviolobilin formation and spectral tuning in the DXCF cyanobacteriochrome subfamily. Biochemistry 51, 1449–1463 (2012). https://doi.org/10.1021/bi201783j

[36] Rockwell, N. C., Martin, S. S., Feoktistova, K., Lagarias, J. C. Diverse two-cysteine photocycles in phytochromes and cyanobacteriochromes. Proc. Natl. Acad. Sci. U.S.A. 108, 11854–11859 (2011). https://doi.org/10.1073/pnas.1107844108

[37] Enomoto, G., Hirose, Y., Narikawa, R., Ikeuchi, M. Thiol-based photocycle of the blue and teal light-sensing cyanobacteriochrome Tlr1999. Biochemistry 51, 3050–3058 (2012). https://doi.org/10.1021/bi300020u

[38] Buhrke, D., Battocchio, G., Wilkening, S., Blain-Hartung, M., Baumann, T., Schmitt, F. J., et al. Red, orange, green: Light- and temperature-dependent color tuning in a cyanobacteriochrome. Biochemistry 59, 509–519 (2020). https://doi.org/10.1021/acs.biochem.9b00931

[39] Wiebeler, C., Schapiro, I. QM/MM benchmarking of cyanobacteriochrome Slr1393g3 absorption spectra. Molecules 24, 1720 (2019). https://doi.org/10.3390/molecules24091720

[40] Wiebeler, C., Rao, A. G., Gärtner, W., Schapiro, I. The effective conjugation length is responsible for the red/green spectral tuning in the cyanobacteriochrome Slr1393g3. Angew. Chem. Int. Ed. Engl. 58, 1934–1938 (2019). https://doi.org/10.1002/anie.201810266

[41] Slavov, C., Xu, X., Zhao, K. H., Gärtner, W., Wachtveitl, J. Detailed insight into the ultrafast photoconversion of the cyanobacteriochrome Slr1393 from Synechocystis sp. Biochim. Biophys. Acta 1847, 1335–1344 (2015). https://doi.org/10.1016/j.bbabio.2015.07.013

[42] Olsen, J. M. H. PyFraME: Python framework for fragment-based multiscale embedding. (2020). https://doi.org/10.5281/ZENODO.3820471

[43] Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., Simmerling, C. ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. J. Chem. Theory Comput. 11, 3696–3713 (2015). https://doi.org/10.1021/acs.jctc.5b00255

[44] Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., et al. The Amber biomolecular simulation programs. J. Comput. Chem. 26, 1668–1688 (2005). https://doi.org/10.1002/jcc.20290

[45] Aidas, K., Angeli, C., Bak, K. L., Bakken, V., Bast, R., Boman, L., et al. The Dalton quantum chemistry program system. Wiley Interdiscip. Rev. Comput. Mol. Sci. 4, 269–284 (2014). https://doi.org/10.1002/wcms.1172

[46] Witzke, S., List, N. H., Olsen, J. M. H., Steinmann, C., Petersen, M., Beerepoot, M. T. P., et al. An averaged polarizable potential for multiscale modeling in phospholipid membranes. J. Comput. Chem. 38, 601–611 (2017). https://doi.org/10.1002/jcc.24718

[47] Beerepoot, M. T. P., Steindal, A. H., List, N. H., Kongsted, J., Olsen, J. M. H. Averaged solvent embedding potential parameters for multiscale modeling of molecular properties. J. Chem. Theory Comput. 12, 1684–1695 (2016). https://doi.org/10.1021/acs.jctc.5b01000

[48] Fujimoto, K., Hayashi, S., Hasegawa, J. Y., Nakatsuji, H. Theoretical studies on the color-tuning mechanism in retinal proteins. J. Chem. Theory Comput. 3, 605–618 (2007). https://doi.org/10.1021/ct6002687

[49] Fujimoto, K. J. Electronic couplings and electrostatic interactions behind the light absorption of retinal proteins. Front. Mol. Biosci. 8, (2021). https://doi.org/10.3389/fmolb.2021.752700

[50] Guareschi, R., Valsson, O., Curutchet, C., Mennucci, B., Filippi, C. Electrostatic versus resonance interactions in photoreceptor proteins: The case of rhodopsin. J. Phys. Chem. Lett. 7, 4547–4553 (2016). https://doi.org/10.1021/acs.jpclett.6b02043