# Advancing Computational Exploration into the Photochemistry of Vision

Shane Canfield*

Department of Biochemistry, Kenyon College, Gambier, OH 43022

## ABSTRACT:

Opsins, groups of proteins which the body uses to sense light, span the biological world with many sharing a retinal protonated Schiff base (RPSB) chromophore. Resulting from convergent evolution, these photo-sensitive molecules perform diverse functions in the animal kingdom. RPSB operates in animal G protein-coupled receptors (GPCRs) as well ion channels in bacteria. When light hits 11-*cis*-retinal (PSB11), the molecule becomes excited and isomerizes into the all-*trans*-retinal (PSBT), influenced by both electrostatic and steric effects. The photoisomerization reaction dynamics have been increasingly investigated with computational models like CASSCF, CASPT2, QM/MM, AMBER, and many more. This review examines methods proposed by recent computational and experimental investigations of opsin proteins that explore steric and electrostatic manipulations of RPSB and the binding pocket for advanced QM/MM simulations. The review specifically highlights the novel Automated Rhodopsin Modeling (ARM) protocol developed by computational photochemist Dr. Massimo Olivucci, Ph.D. The protocol assists experimental researchers who lack computational expertise, but desire accurate computational supplementation of rhodopsin models for experimental observations.

## I. INTRODUCTION

Computational chemistry seeks to solve chemical problems by adapting the underlying chemical theory and principles into high throughput computational algorithms. These algorithms help give insight into novel investigations of reaction mechanisms by using a complex combination of math equations rooted in classical and quantum mechanics. Traditional experimental approaches towards visualizing chemical reactions have required hyper-precise instrumentation which must overcome barriers such as chemical decomposition and unexpected reactions when done in novel systems. Computational chemistry stands to compliment experimental approaches and provide unique insights down to the quantum energy level. However, the potential of computational chemistry to model critical mechanisms in a reaction is bottlenecked as accuracy is not always guaranteed depending on the scope of the computation.

When running calculations with computational software, such as Gaussian, there are many basis sets that can be chosen by the computational chemist to optimize the calculation for either accuracy or speed. Methods like quantum mechanics/molecular mechanics (QM/MM) use QM to simulate high-accuracy bond breaking/formation reactions of key areas and leaves MM to explain larger protein-protein interactions. Similarly, careful attention is given to which molecular interfaces require the computational costly QM models. With these barriers in place, it's imperative that researchers focus on maximizing the accuracy of their calculations while minimizing the computational cost.

The historical struggle with simulating photochemical reaction mechanisms was the inability to calculate the energy states of the molecules before, during, and after excitation from a photon, so it was impossible to accurately model photoisomerization processes. Within the last three decades, the advent of cheaper and higher performing technological capabilities enabled photochemical reaction pathways of rhodopsins are to be increasingly understood.[1] Animal rhodopsins are one of the many molecules in the large family G protein-coupled receptors (GPCRs). GPCRs are plasma membrane bound proteins involved in signal transduction cascades and are made up of 7 α-helices that span the plasma membrane. Most of these GPCRs are activated by ligands which bind to membrane receptors inducing a conformational change. This conformational change results in a cascade of other essential biological mechanisms like vision.[2]

Bovine rhodopsin is one of the most studied visual pigments due to its ease of preparation and abundance. Operating in the retina's rod cells, the chromophore, 11-*cis*-retinal (PSB11), is bound via a Schiff-base linkage to a Lys296 side chain in helix VII and stabilized by a Glu113.[3] Interestingly, PSB11 acts as a reverse agonist. Only after the isomerization is the GCPR is activated. When light hits rhodopsin, PSB11 becomes excited into upper-level energy states and, on a picosecond timescale, relaxes isomerizes into all-*trans*-retinal (PSBT) (**Figure 1)**.

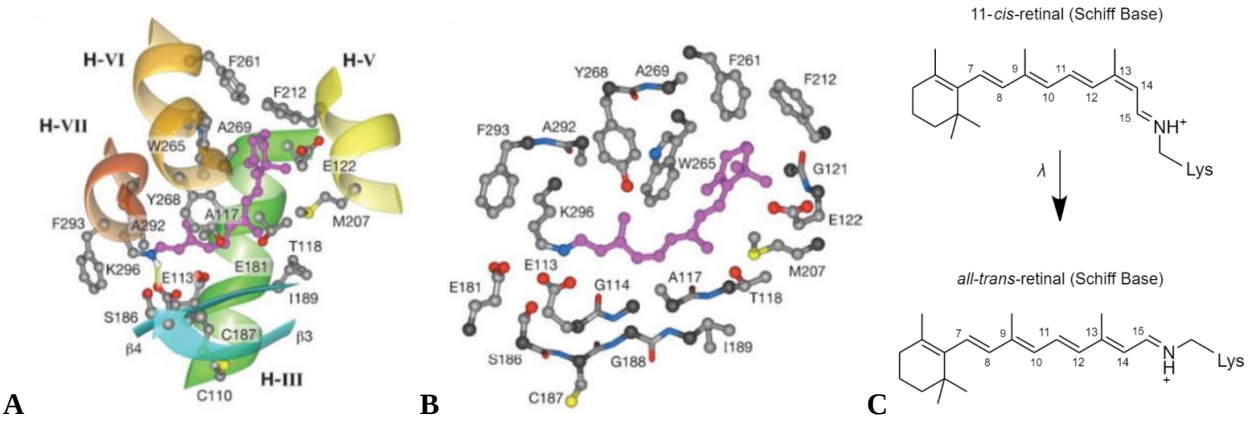



**Figure 1.** (**A)** Amino acid residues that exist in the pocket of the chromophore (pink)[3] (**B**) Residues within 5Å in distance from retinal.[3] (**C**) The photoisomerization of 11-*cis*-retinal (PSB11) diagram.

This adiabatic process results from potential energy surfaces of excited states guiding the molecule along the Franck-Condon (FC) trajectory to the conical intersection (CI), the point at which the potential energy surfaces of the ground state and excited state intersect (**Figure 2)**. On entry into the CI, the molecular decays through a "funneled" into a path either back to the GS or the photoproduct. The path taken by the molecule, and thus the shape of the CI, determines the outcome and timescale of the reaction. The most energetically favorable position in this section of space is described as the minimum energy conical intersection (MECI).

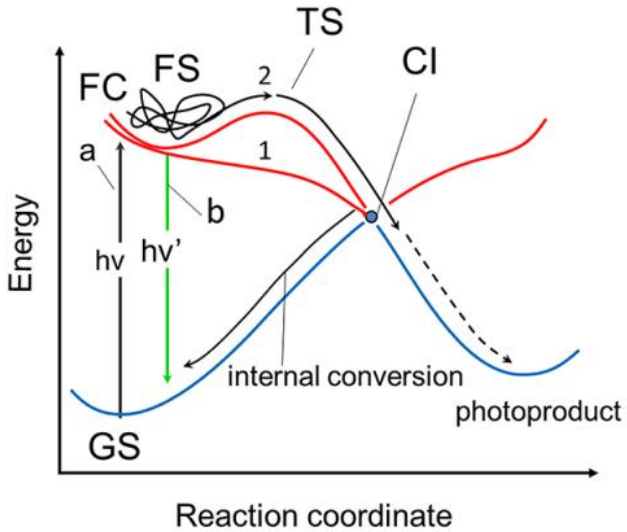



**Figure 2.** Schematic of barrierless (1) and barrier-controlled (2) pathways to either the ground state (GS) or the photoproduct.[4]

Quantum calculations of MECIs can be critical to photochemistry, but the chemical interpretation of their geometries is not so certain.[5] Retinal photoisomerization of other rhodopsins like microbial channelrhodopsins, which share similar relationships with animal rhodopsins, are also being examined as any structural insight on the process can provide chemical clues of photoisomerization mechanics. In photochemistry, QM/MM is used to model the FC trajectory. However, modeling the FC of many structures is computationally costly and additional parameters from other methods (like CASPT2) are needed to account for dynamic electron correlation. This lack of complete understanding has inspired many computational investigations into factors that affect the photoisomerization of the retinal protein within its protein environment.

I explore recent investigations into improving the current understanding of retinal's photoisomerization process using modified retinal proteins placed in the rhodopsin cavity. Then, I will introduce a cutting-edge protocol for modeling rhodopsins with high-throughput algorithms along with a few applications that are paving the way for rhodopsin research. Examining current questions will aid in the advancements of fields such as optogenetics which seek to use light-catalyzed reactions to control cellular events, seen in emerging prototype biomolecular devices.

## II. MODIFICATIONS OF RETINAL FOR IMPROVED PHOTOISOMERIZATION

### Methylation/Demethylation of Carbons Along the Polyene Backbone

The isomerization of PSB11 to PSBT is known to be one of the fastest[6] (200-250 fs) and most efficient[6,7] with a quantum yield (QY) of 0.65 observed. It is known that the photoisomerization of PSB11 occurs much faster in the rhodopsin cavity than in solution[8], so recent investigations of rhodopsins have put heavy emphasis on interactions between PSB11 and the protein environment to understand influences on photoisomerization. Specifically, non-binding interactions between the C10-H and 13-methyl along with other methyl group interactions with the protein cavity have been under investigation as a potential catalyst. This method is especially applicable to PSB11 as hypotheses have emerged that photoisomerization dynamics are facilitated by torsional distortion in the C10-C13 region in two main nonbonded interactions: C10-H/13-methyl interactions and methyl group/opsin cavity interactions.[9] Specifically, steric interactions of the 13-methyl group with Ala117 and the 9-methyl group with Thr118, Gly188, and Ile189. General electrostatic interactions of the Glu113 counterion with Schiff Base is also found to mediate this distortion seen in the C10-C13 regions.[10] Investigations into these interactions found strong changes in QY of the reaction. It was further discovered that C9 is critical in retinal G protein-coupled receptor (RGR) docking following photo-excitation, as the removal of the methyl group leads to a loss in RGR protein activation.[9] Similarly, the removal of the C13 methyl group induces weaker Van der Waals interactions between the chromophore and the protein.[9]

Andruniów and co-workers (2015) wanted to disentangle the role of the C9 and C13 polyene chain positions in the photoisomerization pathway by investigating the impact of retinal methylation, or a lack thereof, on the photoisomerization of PSB11 to its first thermally equilibrated photoproduct intermediate bathorhodopsin (Batho) within the protein cavity. A picosecond vibrational relaxion following PSB11 to PSBT isomerization leads to the formation of the protein intermediate.

*Ab initio* multiconfigurational quantum chemical methodology (QM) is combined with molecular mechanics (CASPT2//CASSCF/AMBER) (denoted QM/MM) to investigate the photoisomerization of an array of methylated and demethylated RPSBs at different carbon positions (**Figure 3A)**. Bond length alternation (BLA) patterns are constructed for native Rh and the altered RPSBs, which find bond lengths to be decreased by 0.01-0.025 Å (0.01-0.015 Å with B3LYP) and dihedral angles changes up to 28° (18° with B3LYP) (**Figure 3B)**. However, these results were considered insignificant. The QM/MM

environment (**Figure 3C)** was based on the protein cavity, and protein residues within 4 Å radius from the QM region were free to relax. The remaining portion were set at its original X-ray derived coordinate.

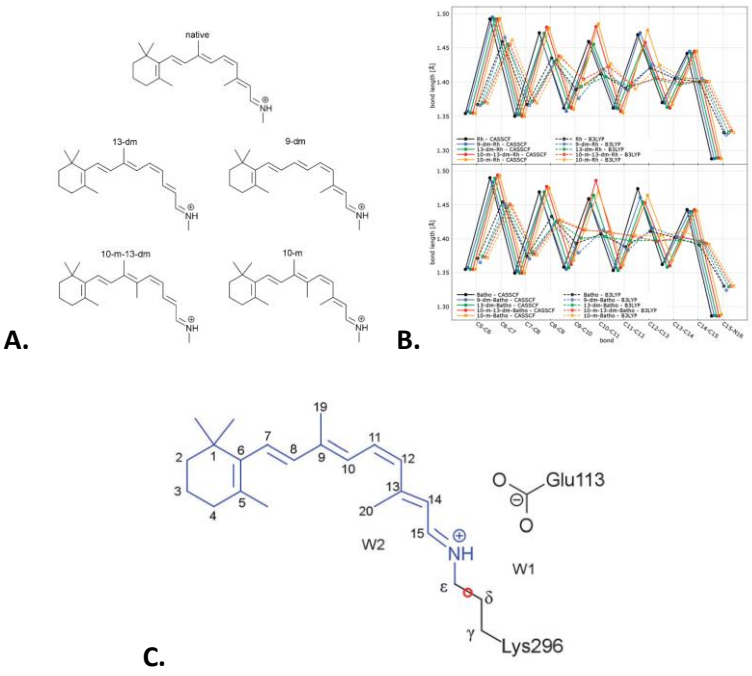



**Figure 3.** (A) Variations of RPSB used for CASPT2/CASSCF/AMBER calculations. (B) Bond length alternation (BLA) patterns of RPSB (Right-top) and all-trans-RPSB (Right-bottom) calculated at different levels of theory.[9] (C) QM region (blue) of the QM/MM calculations.

The CASPT2/CASSCF/MM energy profile calculations for the native and artificial RPSBs with different points of methylation **(Figure 4)** support that methylation has no effect on the photoisomerization reaction coordinate in the $S_1$ state.[9] It appears the steric effects of the methyl groups weren't strong enough to significantly affect the energetic pathway to Batho.

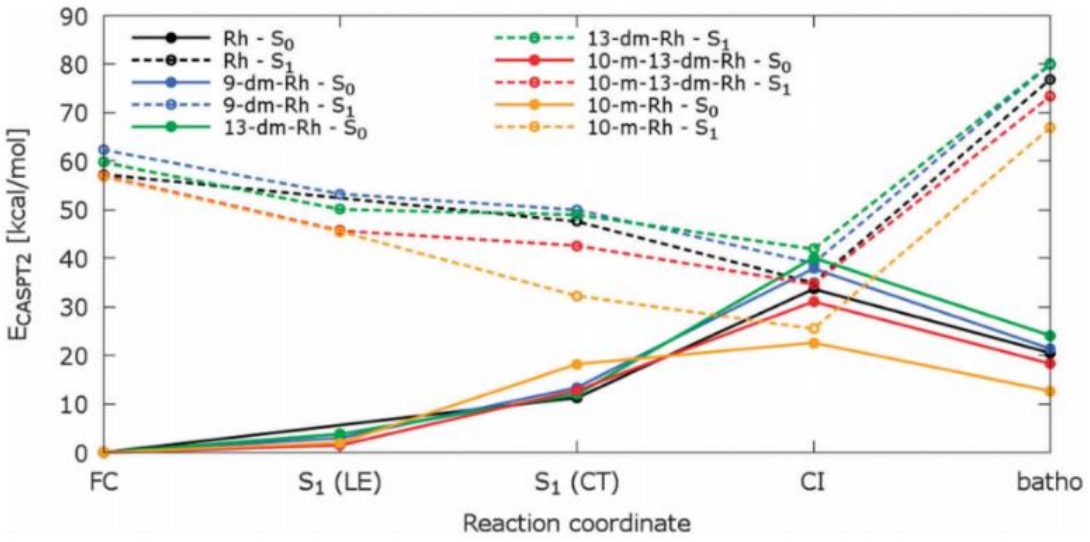



**Figure 4.** CASPT2/CASSCF/MM energy profiles for the native and modified RPSBs with different methyl group positions.[9]

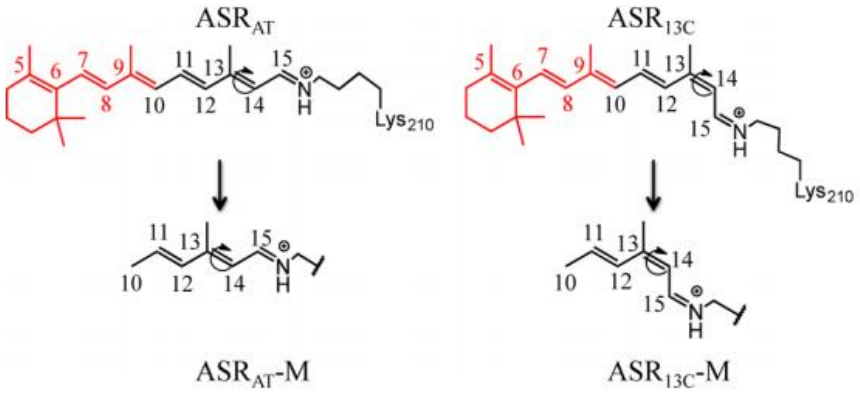
The steric interactions between the methyl and hydrogen groups on RPSB are worth investigating because it allows researchers to separate photoisomerization influences from both steric and electrostatic interactions. Laying out the BLA analysis with QM/MM simulation gives insight to just how dominating these interactions might be in the photoisomerization of RPSB. Further research into these methylation effects, specifically on QY and timescale of PSBT production would be beneficial to further understanding steric effects on this specific mechanism.

## Decreasing QM/MM Computational Complexity with Abridged Retinal Chromophores

In computational photochemistry, the beauty of computer simulation is the possibility for high throughput screening of the molecular pathways. One could theoretically map out each possible molecular trajectory from its excited state (ES) back to its ground state (GS), or photoproduct. This calculation requires great computational power as the algorithms needed are made up of equations stemming from highly complex quantum mechanics. As mentioned earlier, one of the greatest barriers in computational photochemistry are the algorithm's time complexities. As time complexities hinder the number of trajectories that can be simulated, the statistical accuracy of the experiment is threatened. Computational chemists have found ways to improve their time complexities through different combinations of basis sets that find a practice balance between speed and accuracy. Other computational chemists like Manathunga et al. (2016) chose a different approach. In the paper authored by Olivucci himself, as opposed to employing different computational methods the group chose to reduce the size and therefore complexity of the rhodopsin. In the case of Manathunga et al.'s experiment, they investigate the sensory rhodopsin found in the cyanobacterium *Anabaena* PCC7120 (ASR) and modify the molecule from its six conjugated double bond system to only three. It should be noted that a key difference between ASR and other rhodopsin molecules is the presence of a second chromophore. $ASR_{AT}$ (all-*trans)* and $ASR_{13C}$ (13-*cis)* undergo photoisomerization together to function (**Scheme 1**). From here on out, the two-chromophore system will be referred to as PSB3.

It is understood that photoisomerization of rhodopsins occurs mostly at a single carbon double bond, so excluding unnecessary parts of the molecule that aren't significant to the energies of the ES could be a promising approach. Theoretically, it could return more valuable information from the highly accurate and computationally expensive QM method than other cheaper, faster, and less-accurate approaches.

Indeed, it has been reported[11] that in the gas phase the PSB3 chromophores behave much like the full, non-reduced chromophore. This in turn has allowed PSB3 to be a molecule of interest for investigating performance of other QM methods[12–14], and I would argue that due to its success likely prompted Manathunga et al. (2016) to investigate PSB3 chromophores not only in the gas phase but also in a simulated protein environment. If the team was to find that the PSB3 chromophores behave similarly to that of the non-abridged counterparts, then this experiment would advance and optimize computational methods that lead to faster computational approaches to rhodopsin research like ones seen later in this paper.



**Scheme 1.** Isomerization of the two chromophores required for ASR to properly function. Reduced models do not include the red part of the chromophore.[15]

The team took a few routes in this investigation, but the most relevant experiment was in comparing CASSCF/6-31G*/AMBER (denotated as "CASSCF" further in the paper) and CASPT2/6-31G*/AMBER (denoted as "CASPT2") photoisomerization dynamics for QM/MM models between the abridged chromophore against the full chromophore within a protein environment. Calculations were performed using the MOLCAS[16] software package. The QM/MM model and cavity used by Manathunga et al. (2016) is shown in **Figure 5.**

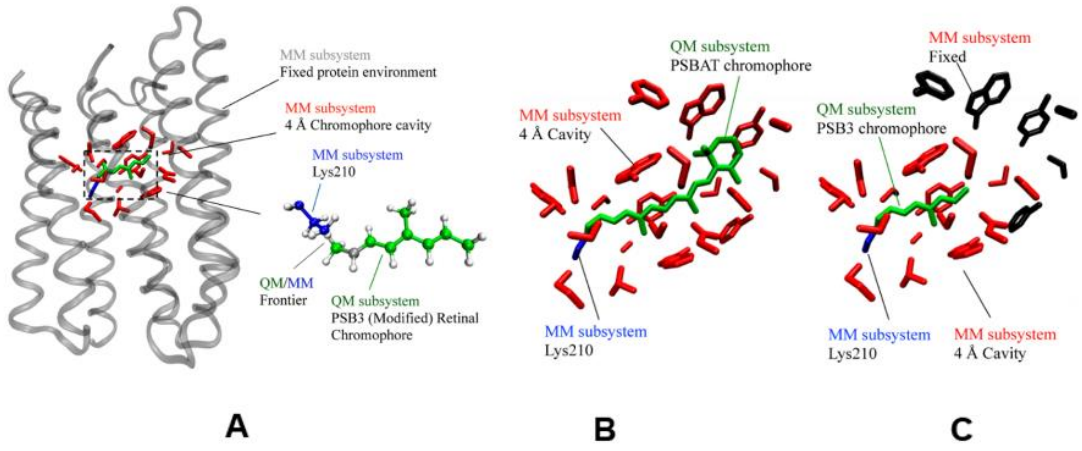



**Figure 5.** (A) The complete QM/MM model of the unabridged PSB3 retinal chromophore. Cavities of the full ASR (B) and PSB3 (C) reduced chromophore.[15]

The results of this experiment indicate that the protein environment leads to an ca. 10-15 degree twisted central double bond in the PSB3 models and the profiles of $S_0$ and $S_1$ potential energies yield similar trends to that of the unabridged ASR.[15] The team reported a qualitative difference between the CASSCF and CASPT2 models expected with the systematic underestimation of the $S_1$ CASPT2 potential energy surface relative to CASSCF and seen in the literature for gas-phase PSB3 models.[14] The team concludes that the PSB3 abridged models and the un-abridged ASR models follow the same reaction coordinate of the full chromophore. This discovery supports the idea that overall photoisomerization reaction coordinate of ASR is not dependent on the electronic structure of the β-ionone ring within the protein cavity. This conclusion can inspire other computational chemists interested in modeling FC reaction coordinates of highly complex protein systems to cut-down their QM/MM calculations with abridged versions, speeding up their research and lowering their computational costs.

Both papers in this section use modified structures of a RPSB inside the protein cavity to understand their direct effects on photoisomerization. Methylation of the backbone appeared to have no effect on the photoisomerization reaction coordinate in the $S_1$ state, however, abridged chromophores follow similar trends to that of their unabridged counterparts. These conclusions provide an interesting perspective on the driving forces of retinal photoisomerization. Both manipulate the sterics between the chromophore and the protein cavity, and both seem to have no effect on the larger trends of the energy state calculations.

## III. ADVANCEMENTS IN COMPUTATIONAL TOOLS WITH AUTOMATIC RHODOPSIN MODELING (ARM)

As computational research of rhodopsin, and other photo-responsive proteins, continues to advance, there lay a few barriers that distance its application from those who are not experts in computational chemistry. For one, input generation for programs like Gaussian and QM/MM modeling requires hours of user implementation and preparation. This barrier inhibits computational chemistry acceptance and growth, however there are a few specific solutions to this problem that allow non-computational scientists to use these tools to speed up their own research.

**a-ARM: Automatic Rhodopsin Modeling**

In a step towards general photobiological automated tools, Olivucci and co-workers have designed what they call the Automatic Rhodopsin Modeling (ARM) protocol.[17] As rhodopsins are of extensively studied in optogenetics to control neuron activity through interactions with light,[18,19] there is monumental incentive to be able to collect large databases of variable rhodopsin configurations to search for beneficial mutations with key properties.

The original concept of the ARM protocol has been briefly reviewed[20] back when the protocol was more accurately described as a "semi-automatic" rhodopsin protocol. At the time, this version had been tested for absorption maximum wavelength ($\lambda_{max}$) of various set of rhodopsins in agreement with experimental data, but there were a few concerns regarding its usage. The first version of ARM required the following inputs: X-ray crystallographic structure or comparative model of the protein in PDB format, a list of residues forming the chromophore cavity, the protonation states of the ionizable side chains, and the position of extracellular and intracellular counterions.[17] However, it was described as "semi-automatic" due to a requirement from the user to manually manipulate the input template structure to accommodate for the requirements above. By manually manipulating the input files, not only was reproducibility a concern, but also human error due to the scalability of the protocol for large arrays of Rhodopsin models. The problems were corrected with the introduction of the automatic input generation functions **(Figure 6)**.

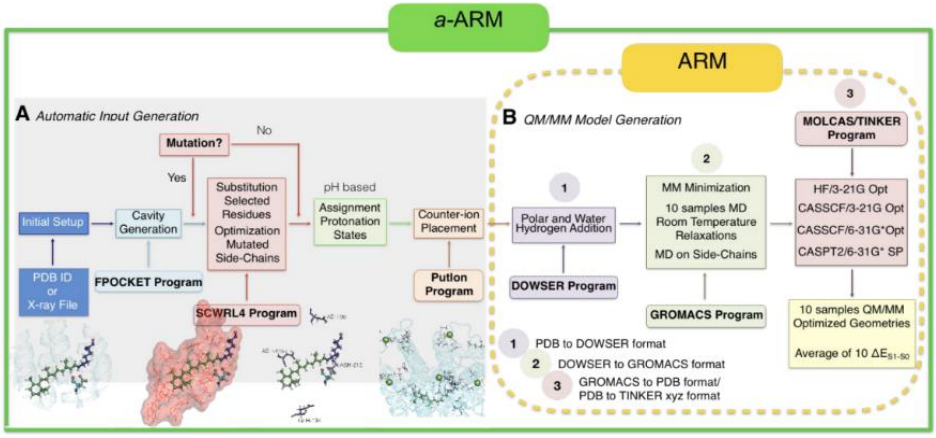



**Figure 6.** General workflow of the new a-ARM protocol for generation of QM/MM models for rhodopsins.[17]

With the updated ARM protocol, the calculation performed with a 32/39 success ratio in reproducing the observed $\lambda_{max}$ trend.[17] With ARM, investigators have been able to speed up their efforts in QM/MM model generation. Today, researchers may go onto Olivucci's "Web-ARM Interface" website to use the tool in their own research. The benefits of efficient ARM QM/MM model generation are only just beginning to be seen.

**ARM for Red-shifting mutations of light-driven sodium-pump microbial rhodopsin**

It is understood that rhodopsins that function with longer wavelengths of light have lower phototoxicities and higher tissue penetration. Previous work in this field includes investigations searching for natural red-shifting rhodopsins[21], screening of different amino acids mutants[22], and reconstitution with retinal analogues.[23] Adjusting the $\lambda_{max}$ is achieved by altering the steric and electrostatic interaction between the chromophore and the protein cavity.[24] It is known[25,26] that mutations that twist the retinylidene backbone shorten the pi-conjugated system resulting in a decreased $\lambda_{max}$. Microbial rhodopsins, however, have a higher planarity which makes shifting the $\lambda_{max}$ more difficult. The Schiff-base linkage of the chromophore has a localized positive charge in the $S_0$ state, which becomes delocalized, and shifts towards the β-ionone ring upon vertical excitation into the $S_1$ state (**Figure 7**).[27] This difference in localization enables selective stabilization of the $S_1$ and $S_0$ state induced by a negative charge near the ring or the Schiff-base, respectively.[28] It turns out that this highly sensitive charge difference imposed by rhodopsin on its chromophore is responsible for the wavelengths of light that are detected by the chromophore.

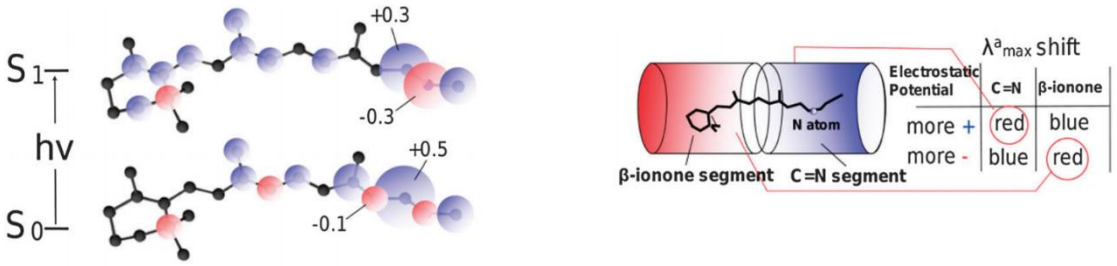



**Figure 7.** (Left) The electrostatic effects upon photoexcitation from $S_0$ to $S_1$ in retinal. The blue and red representing more positive or more negative charge, respectively, and the size of the sphere representing minimum and maximum values. (Right) A mutated retinal cavity with a more positive potential on the Schiff base segment or a more negative potential on the β-ionone ring will red shift the $\lambda_{max}$.[28] An inversion would result in a blue shift.[28]

In animal rhodopsins, several color-regulating residues of varying polarity within the retinal cavity have been investigated. Polar residues around the β-ionone ring have red-shifted both animal[29] and microbial[21] rhodopsins. Inoue and co-workers take this idea and apply it to KR2. KR2 is an outward sodium-pump rhodopsin which transports Na+ from the cytoplasm to the extracellular environment (**Figure 8)**.[24]

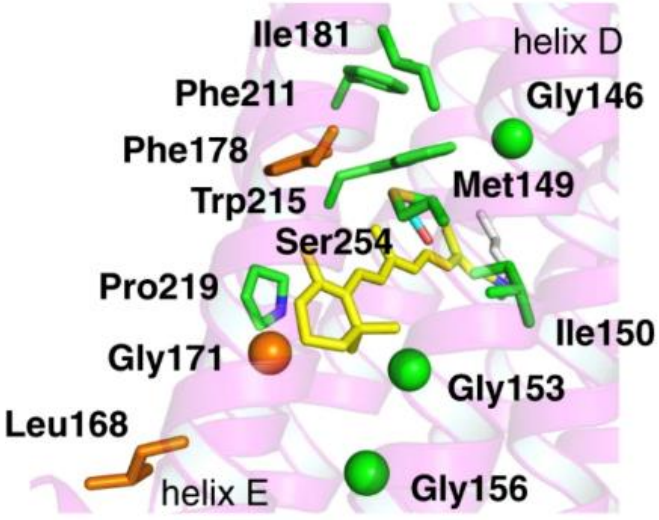



**Figure 8.** The protein cavity of KR2 and the surrounding amino acid residues.[24]

The team screened a total of 12 mutations of hydrophobic residues around the β-ionone ring, and the visible absorption of the mutants and WT was determined with bleached spectra.[24] Although seven failed to express significant red-shift in the absorption, they found that the mutation of Pro219 to Thr successfully returned an ~17 nm ($\lambda_{max}$ = 542nm) red-shift. This is expected as Thr increases the polarity around the β-ionone ring. Another mutation, Ser254 to Ala also returned an approximate ~21 nm red-shift. The hydrophobicity of Ala is ideal over the polar residues Ser perhaps due to some kind of charge disruption around the β-ionone ring in the $S_1$ state. This observation prompted an additional P219T/S254A double-mutation which resulted in an astounding red-shift of ~40 nm without disruption the Na+ transport activity of the WT, which was determined by monitoring pH changes using an ion-transport assay.

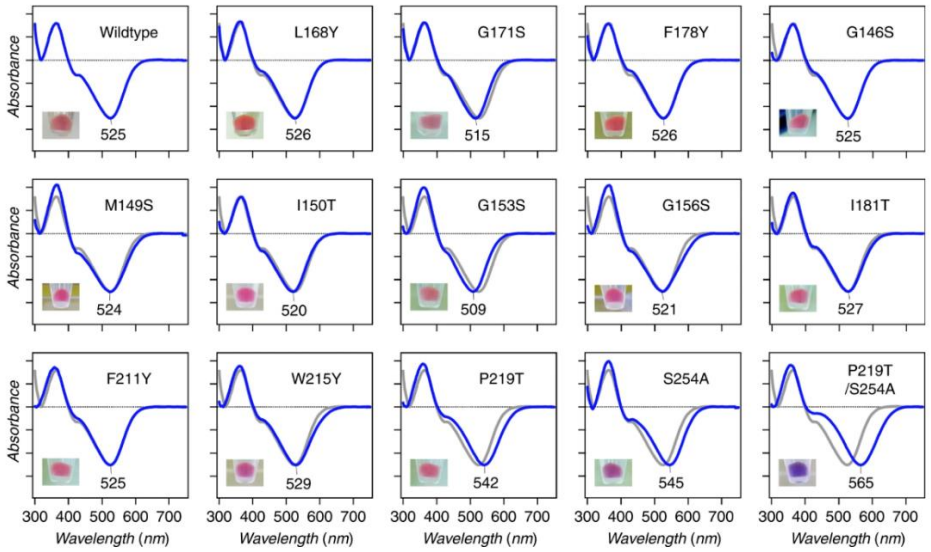



**Figure 9.** The $\lambda_{max}$ calculations for the twelve KR2 mutants (Blue) and WT (Gray). Cultures were bleached with hydroxylamine[24] and visible absorption maxima was analyzed on both KR2. Protein pellets of *E. coli* are also displayed in the lower left of each box.[24]

Although significant, the results from the mutations were insufficient to fully determine the mechanism as to why this shift was happening. To gain structural insights to this mechanism and disentangle the electrostatic and steric effects potentially responsible for the observed red-shift, the team deployed QM/MM calculations using the ARM protocol. Models were generated using the X-ray crystal structure of KR2 as a template for both WT and the KR2 mutants, and vertical excitation energy ($\Delta E_{S1-S0}$) values reflecting changes in chromophore geometry induced by the mutations were calculated for the proteins in isolation (**Figure 10)**.[24]

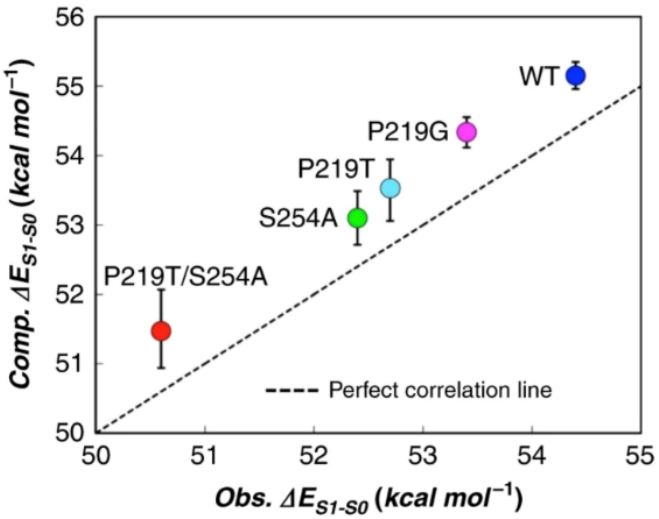



**Figure 10.** Computed vs. observed of the QM/MM models built with the ARM protocol using CASPT2//CASSCF(12,12)/6-31G*/AMBER methods. Error bars are shown in black.[24]

It should be noted that the values returned by the ARM protocol show a minor kcal $mol^{-1}$ systematic error consistent with a blue-shift, but this was ignored as the computational trend is consistent with the observed energy shifts. However, the red-shifting electrostatic contribution induced by the protein environment, specifically in the changes in the dipoles at 219 and 254 positions, ends up dominating over the chromophore steric effect resulting in a net red-sifted change.[24] The small change in excitation energy observed in the P219T/S254 mutation is supported by the experimentally derived $\lambda_{max}$ results, and supports the idea that further design of red-shifting proteins like KR2 would benefit from emphasizing modifications of these dipole moments. The insights provided by Inoue et al. show that the absorption of the rhodopsins can be red-shifted by mutating the protein cavity, specifically the polar residues around the β-ionone ring. The results of this experiment serve as an experimental basis for further investigations into *in-vivo* red-shifting rhodopsins that could one day lead to novel optogenetic tools.

**a-ARM for Testing Infrared Light Absorbance on Bovine Rhodopsin as Human Models**

Since the late 1970s, it has been understood[30] that the human retina can perceive infrared (IR) light in the 800-1355 nm range. Although a few mechanisms have been suggested since, the exact mechanism for photoreceptor activation of IR light has remained undefined. Olivucci and co-workers (2019) accepted this challenge and used the newest version of ARM, a-ARM, to compute the two-photon absorption (TPA) spectra of Bovine rhodopsin (Rh) using ten automated QM/MM models based on a multistate multiconfigurational second-order perturbation (MS-MC-PT2) theory **(Figure 11)**.[31] The investigators chose to use a Rh photoreceptor with a 93% identity to that of the human rhodopsin to serve as a human model for IR absorption.

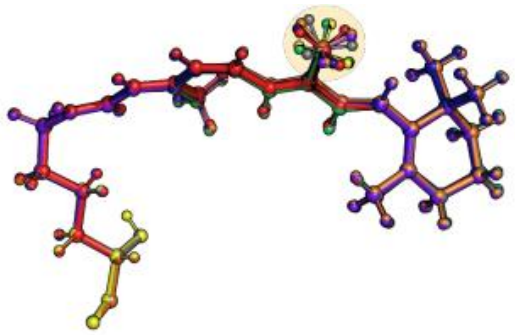


**Figure 11.** The superposition of the Lys-QM system of 10 models generated by the a-ARM protocol. (Highlighted) The significant structural difference is the position of the C13 methyl group.[31]

Ultimately, the TPA calculations of $\lambda_{max}$ and cross-section, $\sigma_{max}$, were performed using a XMCQDPT2/cc-pVTZ//CASSCF/6-31G(d)/AMBER protocol. TPA spectroscopy was computed because it often provides a complementary and more detailed spectroscopy than one-photon absorption (OPA).[32] As a result, the protocol returned transition dipole moments, transition energies, and $\sigma_{TPA}$ values comparable to experimental data.[32] An average $\sigma_{TPA}$ of 472 GM at $\lambda_{max,TPA}$ = 950nm for the $S_0$ to $S_1$ transition (**Figure 12)** was calculated which supports the hypothesis that the cattle eye and likely a human eye, can detect IR light.[31]

The method used for simulating TPA spectrum of the dim-light receptor Rh was validated with computational and experimental values of *trans*-stilbene and its acceptor-pi-donor derivative 4-Dimethylamino-4'-nitrostilbene (ACCD). This comparison was analyzed because ACCD contains a noncentrosymmetric pi-conjugation similar to that of the 11-*cis*-retinal chromophore (rPSB11)[31] and has been used in respectable amount of computational[33–35] and experimental[32] research. In fact, their TPA calculations were in best agreement with that of experimentalist de Wergifosse et al. It should be noted that the computed $\sigma_{TPA}$ value is different than reported[36] in Palczewska et al. (2006). A difference likely due to the use of TDDTP to calculate transition dipole moments or to the macroscopic conversion prefactor, according to Olivucci et al. (2019).

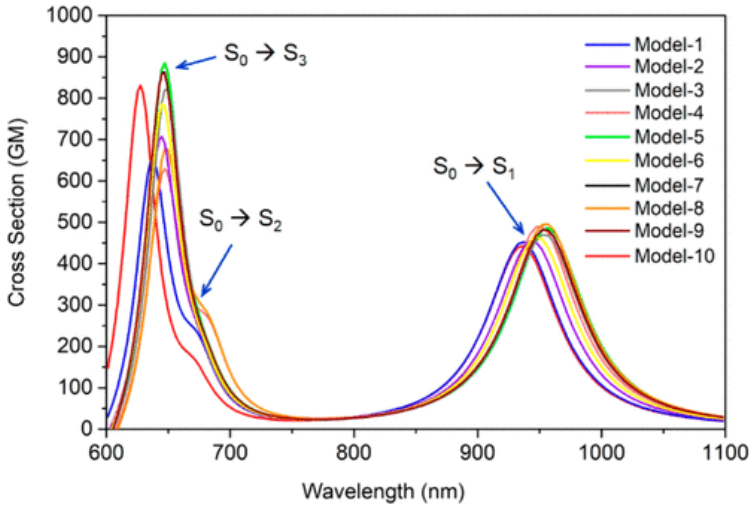



**Figure 12.** TPA spectrum of Rh (rPSB11 embedded in the protein cavity) of 10 a-ARM averaged QM/MM models using the XMCQDPT2/cc-pVTZ//CASSCF/6-31G(d)/AMBERS protocol. [31]

Using ARM enabled Olivucci et al. to further expand the computational evidence of a highly computational studied protein. Not only do the results of this paper support the claim that cattle eye can respond to IR, but specifically that 950 nm light can be perceived by Rh at the TPA level. Interestingly, a significant $S_0$ to $S_2$ transition with an average $\sigma_{TPA}$ of 231 GM at $\lambda_{max,TPA}$ = 678 nm is also observed. Past research by Birge et al. (1985) with OPA[37] supports the $S_0$ to $S_1$ transition seen by Olivucci et al., but not the $S_0$ to $S_2$ transition. The population of the $S_2$ state due to excitation from the ground state is known, however its role in the reaction channel for photoisomerization isn't believed to be a determining factor. Olivucci et al., however, provides the evidence that TPA-based perception of red-light around 678 nm is significant and should be further investigated. Thanks to ARM, new mechanistic insights of relatively well-known rhodopsin proteins are still being discovered.

Although the ARM protocol is very new, non-computational chemists have been using computational tools to further their understanding for quite a while. The Protein Data Base was created to store thousands of protein's digital information to be accessed and downloaded by anyone. Online tools like Benchling uses cloud-based informatics to serve the life sciences. With Benchling, hundreds of protein sequences can be aligned, visualized, and edited in one convenient browser window. More complex protein structure prediction can be achieved with the Rosetta software suite. Computational tools for the life sciences are already integrated and will only increase as technological accessibility and computational efficiency continue to grow.

## V. CONCLUSION AND OUTLOOK

As seen from this paper, the photoisomerization of retinal is being increasingly understood. Computational chemists interested in rhodopsin proteins have explored different approaches to develop a refined understanding of the energetic pathways, intermolecular torsion angles, and electrostatic/steric effects that influence the photoisomerization reaction. Andruniów et al. (2015) explored the effects of methylation, or demethylation, of different carbon bonds along the retinal backbone, and now evidence that methylation does not directly contribute to differences in photoisomerization reaction coordinate in the $S_1$ state. Manathunga et al. (2016) explored reducing the retinal molecule prior to QM/MM calculations and found that abridged molecules follow similar trends to their unabridged counterparts, avoiding the barriers that are a consequence in calculations on sizeable arrays composed of larger molecules. Both researchers independently concluded that changes in steric effects do not seem to be detrimental to photoisomerization, at least in the $S_1$ state.

For non-computationally inclined academics, the ARM protocol has already been used by experimental researchers like Inoue et al. (2019), supporting their experimentally determined observations and providing them a quantum structure basis for further speculation. Already, we see that the electrostatic effects of the protein residues within the environment drastically outweigh steric effects induced by additional/removed methyl groups.

As tools like ARM become refined and applicable to photobiology outside of rhodopsin, many fields will see improvements due to the automation of QM/MM modeling. Researchers like Martínez et al. (2016) have begun to advanced the computational power of graphics processing units (GPU) for GPU-accelerated state-averaged complete active space self-consistent field (SA-CASSCF) calculations.[38] Additionally, insights into the amino acid sequences of rhodopsin and their effect on its functionality have been of significance to photobiology and the design of synthetic mimics and molecular devices.[17] Additionally, optogenetics has seen many advancements due to advancements in rhodopsin research. Rhodopsins are currently being used to improve motor behaviors in *Drosophila* models of Parkinson's disease[39] and even optogenetic activation of the nervous system without fiber-optic implantation through machine learning-guided Gaussian models[40]. The scientific community is approaching the convergence of experimental and highly accurate computational methods which will inevitably boost research productivity in a wide range of biochemical sciences.

## ACKNOWLEDGEMENTS

I would like to thank Dr. John Payton for his help and availability during the production of this review paper. I would also like to thank the Department of Chemistry and Biochemistry for their seemingly limitless resources and motivation during my time at Kenyon College.

## REFERENCES

(1) Migani, A.; Robb, M. A.; Olivucci, M. Relationship between Photoisomerization Path and Intersection Space in a Retinal Chromophore Model. *J. Am. Chem. Soc.* **2003**, *125* (9), 2804–2808. https://doi.org/10.1021/ja027352l.

(2) YOSHIZAWA, T.; WALD, G. Pre-Lumirhodopsin and the Bleaching of Visual Pigments. *Nature* **1963**, *197* (4874), 1279–1286. https://doi.org/10.1038/1971279a0.

(3) Palczewski, K. G Protein–Coupled Receptor Rhodopsin. *Annu. Rev. Biochem.* **2006**, *75* (1), 743–767. https://doi.org/10.1146/annurev.biochem.75.103004.142743.

(4) Laricheva, E. N.; Gozem, S.; Rinaldi, S.; Melaccio, F.; Valentini, A.; Olivucci, M. Origin of Fluorescence in 11- *Cis* Locked Bovine Rhodopsin. *J. Chem. Theory Comput.* **2012**, *8* (8), 2559–2563. https://doi.org/10.1021/ct3002514.

(5) Nakai, H.; Inamori, M.; Ikabata, Y.; Wang, Q. Unveiling Controlling Factors of the $S_0/S_1$ Minimum Energy Conical Intersection: A Theoretical Study. *J. Phys. Chem. A* **2018**, *122* (45), 8905–8910. https://doi.org/10.1021/acs.jpca.8b07864.

(6) Polli, D.; Altoè, P.; Weingart, O.; Spillane, K. M.; Manzoni, C.; Brida, D.; Tomasello, G.; Orlandi, G.; Kukura, P.; Mathies, R. A.; et al. Conical Intersection Dynamics of the Primary Photoisomerization Event in Vision. *Nature* **2010**, *467* (7314), 440–443. https://doi.org/10.1038/nature09346.

(7) Kim, J. E.; Tauber, M. J.; Mathies, R. A. Analysis of the Mode-Specific Excited-State Energy Distribution and Wavelength-Dependent Photoreaction Quantum Yield in Rhodopsin. *Biophysical Journal* **2003**, *84* (4), 2492–2501. https://doi.org/10.1016/S0006-3495(03)75054-1.

(8) Logunov, S. L.; Song, L.; El-Sayed, M. A. Excited-State Dynamics of a Protonated Retinal Schiff Base in Solution. *J. Phys. Chem.* **1996**, *100* (47), 18586–18591. https://doi.org/10.1021/jp962046d.

(9) Walczak, E.; Andruniów, T. Impacts of Retinal Polyene (de)Methylation on the Photoisomerization Mechanism and Photon Energy Storage of Rhodopsin. *Phys. Chem. Chem. Phys.* **2015**, *17* (26), 17169–17181. https://doi.org/10.1039/C5CP01939G.

(10) Yan, E. C. Y.; Ganim, Z.; Kazmi, M. A.; Chang, B. S. W.; Sakmar, T. P.; Mathies, R. A. Resonance Raman Analysis of the Mechanism of Energy Storage and Chromophore Distortion in the Primary Visual Photoproduct †. *Biochemistry* **2004**, *43* (34), 10867–10876. https://doi.org/10.1021/bi0400148.

(11) Borin, V. A.; Wiebeler, C.; Schapiro, I. A QM/MM Study of the Initial Excited State Dynamics of Green-Absorbing Proteorhodopsin. *Faraday Discuss.* **2018**, *207*, 137–152. https://doi.org/10.1039/C7FD00198C.

(12) Ping, Y.; Xu, T.; Momen, R.; Azizi, A.; Kirk, S. R.; Filatov, M.; Jenkins, S. Isomerization of the RPSB Chromophore in the Gas Phase along the Torsional Pathways Using QTAIM. *Chemical Physics Letters* **2017**, *685*, 222–228. https://doi.org/10.1016/j.cplett.2017.07.066.

(13) Bauer, C. A.; Hansen, A.; Grimme, S. The Fractional Occupation Number Weighted Density as a Versatile Analysis Tool for Molecules with a Complicated Electronic Structure. *Chem. Eur. J.* **2017**, *23* (25), 6150–6164. https://doi.org/10.1002/chem.201604682.

(14) Wang, Y.-T.; Liu, X.-Y.; Cui, G.; Fang, W.-H.; Thiel, W. Photoisomerization of Arylazopyrazole Photoswitches: Stereospecific Excited-State Relaxation. *Angew. Chem. Int. Ed.* **2016**, *55* (45), 14009–14013. https://doi.org/10.1002/anie.201607373.

(15) Manathunga, M.; Yang, X.; Luk, H. L.; Gozem, S.; Frutos, L. M.; Valentini, A.; Ferrè, N.; Olivucci, M. Probing the Photodynamics of Rhodopsins with Reduced Retinal Chromophores. *J. Chem. Theory Comput.* **2016**, *12* (2), 839–850. https://doi.org/10.1021/acs.jctc.5b00945.
(16) Aquilante, F.; De Vico, L.; Ferré, Nicolas, N.; Ghigo, G.; Malmqvist, P.; Neogrády, P.; Pedersen, T. B.; Pitoňák, M.; Reiher, M.; Roos, B. O.; et al. MOLCAS 7: The Next Generation. *J. Comput. Chem.* **2010**, *31* (1), 224–247. https://doi.org/10.1002/jcc.21318.
(17) Pedraza-González, L.; De Vico, L.; Marín, M. del C.; Fanelli, F.; Olivucci, M. *A*-ARM: Automatic Rhodopsin Modeling with Chromophore Cavity Generation, Ionization State Selection, and External Counterion Placement. *J. Chem. Theory Comput.* **2019**, *15* (5), 3134–3152. https://doi.org/10.1021/acs.jctc.9b00061.
(18) Gushchin, I.; Shevchenko, V.; Polovinkin, V.; Borshchevskiy, V.; Buslaev, P.; Bamberg, E.; Gordeliy, V. Structure of the Light-Driven Sodium Pump KR2 and Its Implications for Optogenetics. *FEBS J* **2016**, *283* (7), 1232–1238. https://doi.org/10.1111/febs.13585.
(19) Govorunova, E. G.; Sineshchekov, O. A.; Li, H.; Spudich, J. L. Microbial Rhodopsins: Diversity, Mechanisms, and Optogenetic Applications. *Annu. Rev. Biochem.* **2017**, *86* (1), 845–872. https://doi.org/10.1146/annurev-biochem-101910-144233.
(20) Gozem, S.; Luk, H. L.; Schapiro, I.; Olivucci, M. Theory and Simulation of the Ultrafast Double-Bond Isomerization of Biological Chromophores. *Chem. Rev.* **2017**, *117* (22), 13502–13565. https://doi.org/10.1021/acs.chemrev.7b00177.
(21) Schneider, F.; Grimm, C.; Hegemann, P. Biophysics of Channelrhodopsin. *Annu. Rev. Biophys.* **2015**, *44* (1), 167–186. https://doi.org/10.1146/annurev-biophys-060414-034014.
(22) Prigge, M.; Schneider, F.; Tsunoda, S. P.; Shilyansky, C.; Wietek, J.; Deisseroth, K.; Hegemann, P. Color-Tuned Channelrhodopsins for Multiwavelength Optogenetics. *J. Biol. Chem.* **2012**, *287* (38), 31804–31812. https://doi.org/10.1074/jbc.M112.391185.
(23) Takayama, R.; Kaneko, A.; Okitsu, T.; Tsunoda, S. P.; Shimono, K.; Mizuno, M.; Kojima, K.; Tsukamoto, T.; Kandori, H.; Mizutani, Y.; et al. Production of a Light-Gated Proton Channel by Replacing the Retinal Chromophore with Its Synthetic Vinylene Derivative. *J. Phys. Chem. Lett.* **2018**, *9* (11), 2857–2862. https://doi.org/10.1021/acs.jpclett.8b00879.
(24) Inoue, K.; del Carmen Marín, M.; Tomida, S.; Nakamura, R.; Nakajima, Y.; Olivucci, M.; Kandori, H. Red-Shifting Mutation of Light-Driven Sodium-Pump Rhodopsin. *Nat Commun* **2019**, *10* (1), 1993. https://doi.org/10.1038/s41467-019-10000-x.
(25) Sudo, Y.; Okazaki, A.; Ono, H.; Yagasaki, J.; Sugo, S.; Kamiya, M.; Reissig, L.; Inoue, K.; Ihara, K.; Kandori, H.; et al. A Blue-Shifted Light-Driven Proton Pump for Neural Silencing. *J. Biol. Chem.* **2013**, *288* (28), 20624–20632. https://doi.org/10.1074/jbc.M113.475533.
(26) Kato, H. E.; Kamiya, M.; Sugo, S.; Ito, J.; Taniguchi, R.; Orito, A.; Hirata, K.; Inutsuka, A.; Yamanaka, A.; Maturana, A. D.; et al. Atomistic Design of Microbial Opsin-Based Blue-Shifted Optogenetics Tools. *Nat Commun* **2015**, *6* (1), 7177. https://doi.org/10.1038/ncomms8177.
(27) Coto, P. B.; Strambi, A.; Ferre, N.; Olivucci, M. The Color of Rhodopsins at the Ab Initio Multiconfigurational Perturbation Theory Resolution. *Proceedings of the National Academy of Sciences* **2006**, *103* (46), 17154–17159. https://doi.org/10.1073/pnas.0604048103.
(28) Melaccio, F.; Ferré, N.; Olivucci, M. Quantum Chemical Modeling of Rhodopsin Mutants Displaying Switchable Colors. *Phys. Chem. Chem. Phys.* **2012**, *14* (36), 12485. https://doi.org/10.1039/c2cp40940b.
(29) van Hazel, I.; Dungan, S. Z.; Hauser, F. E.; Morrow, J. M.; Endler, J. A.; Chang, B. S. W. A Comparative Study of Rhodopsin Function in the Great Bowerbird ( *P Tilonorhynchus Nuchalis* ): Spectral Tuning and Light-Activated Kinetics: A Comparative Study of Rhodopsin Function. *Protein Science* **2016**, *25* (7), 1308–1318. https://doi.org/10.1002/pro.2902.
(30) Dmitriev, V. G.; Emel'yanov, V. N.; Kashintsev, M. A.; Kulikov, V. V.; Solov'ev, A. A.; Stel'makh, M. F.; Cherednichenko, O. B. Nonlinear Perception of Infrared Radiation in the 800–1355 Nm Range with Human Eye. *Soviet Journal of Quantum Electronics* **1979**, *9* (4), 475–479. https://doi.org/10.1070/qe1979v009n04abeh008913.

(31) Gholami, S.; Pedraza-González, L.; Yang, X.; Granovsky, A. A.; Ioffe, I. N.; Olivucci, M. Multistate Multiconfiguration Quantum Chemical Computation of the Two-Photon Absorption Spectra of Bovine Rhodopsin. *J. Phys. Chem. Lett.* **2019**, *10* (20), 6293–6300. https://doi.org/10.1021/acs.jpclett.9b02291.

(32) de Wergifosse, M.; Houk, A. L.; Krylov, A. I.; Elles, C. G. Two-Photon Absorption Spectroscopy of *Trans* -Stilbene, *Cis* -Stilbene, and Phenanthrene: Theory and Experiment. *The Journal of Chemical Physics* **2017**, *146* (14), 144305. https://doi.org/10.1063/1.4979651.

(33) Albota, M. Design of Organic Molecules with Large Two-Photon Absorption Cross Sections. *Science* **1998**, *281* (5383), 1653–1656. https://doi.org/10.1126/science.281.5383.1653.

(34) Morel, Y.; Irimia, A.; Najechalski, P.; Kervella, Y.; Stephan, O.; Baldeck, P. L.; Andraud, C. Two-Photon Absorption and Optical Power Limiting of Bifluorene Molecule. *J. Chem. Phys.* **2001**, *114* (12), 5391–5396. https://doi.org/10.1063/1.1351160.

(35) Nayyar, I. H.; Masunov, A. E.; Tretiak, S. Comparison of TD-DFT Methods for the Calculation of Two-Photon Absorption Spectra of Oligophenylvinylenes. *J. Phys. Chem. C* **2013**, *117* (35), 18170–18189. https://doi.org/10.1021/jp403981d.

(36) Palczewska, G.; Vinberg, F.; Stremplewski, P.; Bircher, M. P.; Salom, D.; Komar, K.; Zhang, J.; Cascella, M.; Wojtkowski, M.; Kefalov, V. J.; et al. Human Infrared Vision Is Triggered by Two-Photon Chromophore Isomerization. *Proc Natl Acad Sci USA* **2014**, *111* (50), E5445–E5454. https://doi.org/10.1073/pnas.1410162111.

(37) Birge, R. R.; Murray, L. P.; Pierce, B. M.; Akita, H.; Balogh-Nair, V.; Findsen, L. A.; Nakanishi, K. Two-Photon Spectroscopy of Locked-11-Cis-Rhodopsin: Evidence for a Protonated Schiff Base in a Neutral Protein Binding Site. *Proceedings of the National Academy of Sciences* **1985**, *82* (12), 4117–4121. https://doi.org/10.1073/pnas.82.12.4117.

(38) Snyder, J. W.; Curchod, B. F. E.; Martínez, T. J. GPU-Accelerated State-Averaged Complete Active Space Self-Consistent Field Interfaced with Ab Initio Multiple Spawning Unravels the Photodynamics of Provitamin D $_3$. *J. Phys. Chem. Lett.* **2016**, *7* (13), 2444–2449. https://doi.org/10.1021/acs.jpclett.6b00970.

(39) Imai, Y.; Inoshita, T.; Meng, H.; Shiba-Fukushima, K.; Hara, K. Y.; Sawamura, N.; Hattori, N. Light-Driven Activation of Mitochondrial Proton-Motive Force Improves Motor Behaviors in a Drosophila Model of Parkinson's Disease. *Commun Biol* **2019**, *2* (1), 424. https://doi.org/10.1038/s42003-019-0674-1.

(40) Bedbrook, C. N.; Yang, K. K.; Robinson, J. E.; Mackey, E. D.; Gradinaru, V.; Arnold, F. H. Machine Learning-Guided Channelrhodopsin Engineering Enables Minimally Invasive Optogenetics. *Nature Methods* **2019**, *16* (11), 1176–1184. https://doi.org/10.1038/s41592-019-0583-8.