



Wrocław University
of Science and Technology

DOCTORAL THESIS

Two-Photon Absorption in Isolated Retinal Protonated Schiff Base and Rhodopsin Studied by QM/MM methods

AUTHOR

M.Sc. Saruti Sirimatayanant

The PhD thesis was prepared in the Institute of Advanced
Materials of Wrocław University of Science and Technology
under the supervision of Prof. Tadeusz Andruniów.

Keywords: retinal; rhodopsin; retinal protonated Schiff base; one-photon absorption; two-photon absorption; coupled-cluster theory; time-dependent density functional theory; QM/MM; polarizable embedding; nonlinear optical properties

Wrocław 2026

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Prof. Tadeusz Andruniów, for his guidance, scientific insight, encouragement, and continued support throughout this doctoral work.

I am also grateful to Jógvan Magnus Haugaard Olsen of the Department of Chemistry at the Technical University of Denmark for his generous assistance with Dalton and for the many valuable technical discussions that supported the computational work.

I gratefully acknowledge Wrocław University of Science and Technology for financial support through an increased doctoral scholarship awarded for the 2023/2024 academic year.



All calculations were performed using facilities of Wrocław Centre for Networking and Supercomputing.

Glossary of abbreviations and symbols

Abbreviations in alphabetical order.

1PA	one-photon absorption
2PA	two-photon absorption
2PE	two-photon excitation
2SM	two-state model
3PA	three-photon absorption
a.a.	amino acid
ARM	Automatic Rhodopsin Model
B1LYP	Becke one-parameter Lee–Yang–Parr hybrid functional
B3LYP	Becke three-parameter Lee–Yang–Parr hybrid functional
BHandHLYP	Becke half-and-half Lee–Yang–Parr hybrid functional
BLA	bond-length alternation
BLYP	Becke exchange with Lee–Yang–Parr correlation
BO	Born–Oppenheimer
BR	bacteriorhodopsin
CAM-B3LYP	Coulomb-attenuating method B3LYP functional
CASPT2	complete active space second-order perturbation theory
CASPT3	complete active space third-order perturbation theory
CASSCF	complete active space self-consistent field
CatCh	calcium-translocating channelrhodopsin
CC	coupled cluster
CC2	second-order approximate coupled cluster
CC3	third-order approximate coupled cluster
CCS	coupled cluster with single excitations
CCSD	coupled cluster with single and double excitations
CCSD(T)	CCSD with a perturbative treatment of connected triple excitations
CCSDT	coupled cluster with single, double and triple excitations
ChR-I	channelrhodopsin-1
ChR-II	channelrhodopsin-2
CI	configuration interaction
cAMP	cyclic adenosine monophosphate
cGMP	cyclic guanosine monophosphate
CP2K	open-source quantum chemistry and molecular dynamics software package
DFT	density functional theory
EE	electrostatic embedding
eNpHR3.0	enhanced Natronomonas pharaonis halorhodopsin 3.0
ES	excited state
EXX	exact exchange
FLIM	fluorescence lifetime imaging microscopy
FP	fluorescent protein
FPE	frozen polarizable embedding
FQY	fluorescence quantum yield
GGA	generalized gradient approximation
GPCR	G protein-coupled receptor
GS	ground state

Gt	transducin G protein
HBN	hydrogen-bond network
HF	Hartree–Fock
HR	halorhodopsin
IR	infrared
KS	Kohn–Sham
LCAO-MO	linear combination of atomic orbitals into molecular orbitals
LDA	local density approximation
LRCH	long-range-corrected hybrid
M11	Minnesota 2011 range-separated hybrid meta-GGA functional
MAE	mean absolute error
MAPE	mean absolute percentage error
MBPT	many-body perturbation theory
MD	molecular dynamics
ME	mechanical embedding
MM	molecular mechanics
MN15	Minnesota 2015 global-hybrid meta-NGA functional
MO	molecular orbital
MP2	second-order Møller–Plesset perturbation theory
MR-CISD	multireference configuration interaction with single and double excitations
MS-CASPT2	multistate complete active space second-order perturbation theory
NE	no embedding
NeoR	neorhodopsin
NEVPT2	second-order n-electron valence state perturbation theory
NIR	near-infrared
NLO	nonlinear optical
NpHR	Natronomonas pharaonis halorhodopsin
PBQC	Poisson–Boltzmann quantum chemical method
PDB	Protein Data Bank
PE	polarizable embedding
PSB	protonated Schiff base
QM	quantum mechanics
QM/MM	quantum mechanics/molecular mechanics
QMC	quantum Monte Carlo
RGR	retinal G protein-coupled receptor
RI	resolution of the identity
RI-CC2	resolution-of-the-identity second-order approximate coupled cluster
RMSD	root-mean-square deviation
rPSB	retinal protonated Schiff base
SCF	self-consistent field
SHG	second-harmonic generation
SOS	sum-over-states
SR-I	sensory rhodopsin I
SR-II	sensory rhodopsin II
TD-DFT	time-dependent density functional theory
TPA	two-photon absorption
TR	thermophilic rhodopsin
UV	ultraviolet
WCSS	Wrocław Centre for Networking and Supercomputing

WWII	Second World War
WT	wild type
XCF	exchange-correlation functional
XMCQDPT2	extended multiconfigurational quasi-degenerate second-order perturbation theory
ZINDO	Zerner's intermediate neglect of differential overlap

Symbols and units

a.u.	atomic units
Å	ångström
D	debye
ΔE	vertical excitation energy
$\Delta\mu_{0n}$	change in permanent dipole moment between the ground state and the nth excited state
δ^{2PA}	rotationally averaged two-photon transition strength
eV	electronvolt
f	oscillator strength
fs	femtosecond
GM	Göppert–Mayer unit of two-photon absorption cross-section
$\lambda_{\max}$	wavelength of maximum absorption or emission
μ_0	ground-state permanent dipole moment
μ_n	permanent dipole moment of the nth excited state
μ_{0n}	transition dipole moment between the ground state and the nth excited state
nm	nanometre
ns	nanosecond
ps	picosecond
S_0	electronic ground state
S_n	nth singlet excited state
σ^{2PA}	two-photon absorption cross-section
T_1/T_2	single- and double-excitation amplitudes or excitation character
ω	angular frequency

Contents

Acknowledgements	3
Glossary of abbreviations and symbols.....	4
1. 1 Introduction	9
1.2.1: Type-I and type-II Retinals and Rhodopsins structure	15
1.2.2: Retinal structure affects absorption spectra	16
1.3.1 Protein environment impact on retinal geometry	19
1.3.2 Protein environment impact on the spectral properties	21
1.3.3 Electrostatics and polarizations	22
1.3.4 Mutual polarization schemes	26
1.4.1: Two-Photon Absorption has many benefits over One-Photon techniques	28
1.4.2: Experimental measurements of 2PA	29
1.4.3: Theoretical calculations of 2PA	30
1.4.4: The QM/MM model.....	38
2. Objectives.....	42
3. Methodology	43
3.1: Schrödinger's Time-Independent Equation and the Hartree-Fock method.....	43
3.2: Perturbation Theory	46
3.3: Coupled Cluster	47
3.4: Density Functional Theory (DFT)	48
3.5.1: Multiscale Modeling with QM/MM	50
3.5.2: Partitioning the system.....	51
3.5.3: QM/MM is an interacting system	52
3.5.4: Modeling of 1PA and TPA signals.....	54
3.5.5: Converting Two-Photon Transition Moments to Cross-Sections	57
4. Results and Discussion (gas phase)	58
4.1: Climbing the CC Hierarchy and TD-DFT Benchmarking of Retinal Models.....	58
4.1.1: The reduced sized models and the benchmarking strategy	59
4.1.2: Computational methods and benchmarking	60
4.1.3.1 Results and Discussion: CC hierarchy	62
4.1.3.2: CC hierarchy in higher excited-states of PSB3	69
4.1.4: Impact of different XCF	73
4.1.5 Higher Excited-states in TD-DFT	83
4.2 Tuning δ^{2PA} in 11-cis rPSB via De/Methylation Patterns: Benchmarks in RI-CC2 and TD-DFT Validation	85
4.2.1 Evidence of 2PA Tuning in rPSB.....	85
4.2.2 Methodology.....	86

4.2.3 De/methylation Impact on Spectral Properties at the RI-CC2 level for the S ₁ state	88
4.2.4 Performance of TD-DFT Functionals in the S ₁ Excited-state	95
4.2.5 Performance of TD-DFT Functionals in the Higher Excited-states	100
4.2.6.1 2-State Model as a Good Qualitative Predictor of QR	102
4.2.6.2 Retinal Protonated Schiff Base with and without β -ionone Ring.....	103
5. Results and Discussion (Rhodopsin Protein)	104
5.1 QM region size convergence for 1PA/2PA in Rhodopsin	104
5.1.1 The Rhodopsin model	105
5.1.2 Molecular Dynamics Calculations Details.....	108
5.1.3 Electronic Properties and Spectral calculations	109
5.2.1 The M1 Ensemble and its Averages in Polarizable Embedding	110
5.2.2 The M2 Ensemble and its Averages in Polarizable Embedding	117
5.2.3 Single-Frame QM Region Expansion and Spectral Convergence	122
6. Summary and Conclusions	124
7. Supporting Information.....	129
Citation:	135
Scientific Output.....	147

1.1 Introduction

Since prehistoric times, organisms as small as bacteria and larger than trees utilize light at various wavelengths for their own benefits. In vertebrates color vision is processed by Cone opsins in the cone cells, while scotopic vision is processed by Rod opsins (Rhodopsin) in the rod cells [1, 2]. The original Rhodopsin protein discovered by Heinrich Müller in 1851 [3] may be one of the first steps for scientists to uncover the exact mechanism to which these various organisms are able to exploit the power of light. It wasn't until over 20 years later however, when Franz Boll and Willhelm Kühne published a series of paper demonstrating the effect of *bleaching* of the Rhodopsin protein when under illuminated conditions, and its regeneration in the dark [4–10]. Fast forward to 1955, George Wald, along with contributions from Ragnar Granit and Halden Keffer [11, 12], published the first article elucidating the exact photoreceptor process in Rhodopsin which awarded the three scientists a Nobel Prize in Physiology in 1967 [13]. In the present day, multiple classifications of photoactive proteins have been discovered ranging from Rhodopsin and Rhodopsin-like proteins, fluorescent proteins, photoenzymes, photopumps [1, 14], as well as synthetic photoactive dyes [15].

As noted by Müller during his discovery, the color of rod cells appear red in which he reflects so in its name [2]. The word '*Rhodopsin*' itself is derived from two ancient Greek words; *rhodon* and *opsis*, which translates to rose and sight respectively [1]. It is generally accepted that visible light lies within approx. 400 – 750 nm, although the boundaries of human vision are not sharply defined and may extend slightly into the near-ultraviolet and near-infrared regions under favorable conditions. Similarly, infrared (IR) has also been of great interest particularly during the invention of the radar during WWII, where pilots were suspected to be capable of perceiving intense IR radiation [16]. This phenomenon was studied by Wald and co-workers in 1947. The team intuitively observed that over 800 nm the rod cells are more sensitive than cone cells, although subjects still perceive colors ranging from red to yellow due to the foveal vision, within the peripheral retina only the perception of colorless light was observed [17]. With the advent of modern technologies including modern optical, electrophysiological, biochemical and computational tools, this long standing question was systematically studied in both a physical and theoretical setting in a 2014 paper by Palczewska and co [16]. The team demonstrated the visual perception of IR in 30 human test subjects, all of which perceived the laser as red for wavelengths shorter than 1040 nm, and for wavelengths longer than 1040 nm, perceived wavelength was half of the excitation wavelength [16]. More importantly, the team compared the minimum power required to induce a visual sensation in test subjects for short

pulse (~0.9 ps) versus long pulse (~0.5 ns) laser at 1000 nm wavelength and noted that long pulse requires over 100 times more power, suggesting a nonlinear optical process. The team further proved their findings via direct analysis of intact retinas isolated from mice eyes, as well as biochemical analysis of crystalized and in solution Rhodopsin. Palczewska and co-workers determined that indeed the activation and subsequent isomerization of Rhodopsin is activated by a 2-Photon Absorption (2PA) process, and no involvement of the Second Harmonic Generation (SHG) mechanism was detected, when illuminated with IR light.

The current state of interests in Rhodopsin goes further than simply human or animal perception of IR, various applications for both eukaryotic and prokaryotic Rhodopsins have been adopted in multiple fields [1]. The discovery of the first prokaryotic Rhodopsin, or rather microbial Rhodopsin, called bacteriorhodopsin (BR) in 1971 by Osterhelt and Stoeckenius [1, 18] as well as halorhodopsin (HR) in 1977 by Matsuno-Yagi and Mukohata [19] further opened up the possibilities for Rhodopsin applications due to its outward proton or inward chlorine pump. Since then, even more similar photoactive proteins with various functions have been found in both Eubacteria and certain lower Eukaryotes which have also been included in the category of microbial Rhodopsin, like the sensory Rhodopsins (SR-I, SR-II) that control bacterial phototaxis [20, 21] and Channelrhodopsins (ChR-I, ChR-II) in green algae that acts as a light-gated cation channel [22, 23]. This is of course in contrast to higher order Eukaryotes who's Rhodopsins are G-Protein Coupled Receptors (GPCR) that transduce light stimulus to the G-protein signaling cascade [24]. This variety in both its intrinsic nature, due to its origin, as well as its function ushered in a large collection of modern tools utilizing Rhodopsins, most notably within the optogenetics [1, 14, 18] field, but also in optical biopsy [25], light-driven pumps and channels [18], and even as fluorescent tags themselves [26, 27].

The term optogenetics, popularized in 2010 by Deisseroth, is described as “the combination of genetic and optical methods to achieve gain or loss of function or well-defined events in specific cells of living tissue” [28]. The goal for neurologists and neurobiologists that study activity and function of the brain has been to control the activation of neurons, the discovery of various types of Rhodopsins and the ability to easily express them *in vivo* and *in vitro*, paved way to the rise of optogenetics as a field [1, 14, 18]. However, in a broad sense optogenetics encompasses more than just the Rhodopsin and the ability to control neural activations. Deisseroth also expands the term to include any light-responsive controllable tools that can be targeted and deliver effector functions [29]. In essence, optogenetics should have the 3 primary abilities; i) deliver light into tissues under investigation, ii) ability to target specific cells of interests, and iii) obtain compatible readouts and perform analysis like targeted imaging or electrical recording of neural activities [29]. One of the earlier hurdles

identified by researchers in implementing the Rhodopsin proteins in optogenetics are 2 fold. Firstly, is the abundance of the available chromophore within Rhodopsin called retinal. The 11-cis retinal chromophore of animal Rhodopsin is not as readily abundant in animal cells, whereas microbial Rhodopsin's all-trans retinal is [1, 30]. The second reason being the photo-bleaching effect that is part of animal Rhodopsin's photocycle [1, 30]. Upon excitation, 11-cis retinal isomerizes into all-trans in animal Rhodopsin, and all-trans into 13-cis in microbial Rhodopsin [30, 31]. While 13-cis can be regenerated back to all-trans via spontaneous thermal isomerization in microbial Rhodopsin, 11-cis cannot be regenerated in-situ and must be enzymatically replaced [30, 31]. Furthermore, sustained and repetitive activation of animal Rhodopsin's GPCR causes a decline in the signaling response [32]. Thus, the addition of this step in the animal Rhodopsin's photocycle creates another layer of complexity that must be accounted for when crafting chimeric Rhodopsin based optogenetic tools like the 'optoXR'. Regardless, scientists have designed a multitude of optoXRs, based on vertebrates like rat [33, 34] and bovine [35] to name a few. In fact, bovine Rhodopsin is among the most extensively studied opsins [1]. In terms of optoXRs, chimeric bovine Rhodopsin with cytoplasmic loops replaced with Gs and Gq-coupled hamster and human adrenergic receptors, respectively. The chimeras demonstrated *in vitro* to control the release of cyclic AMP-coupled signaling cascade in extracted coronal slice of mice accumbens upon optical stimulation [35]. The feasibility of these optoXRs was possible due to the sufficient levels of retinoid within mammalian brain for the regeneration of 11-cis Retinal [36]. Furthermore, within the same study, the team also demonstrated *in vivo* behavioral control of mice. The optoXR was transgenically expressed in freely moving mice attached to a fibre optic stimulus source, which allowed temporally precise control over signaling in the accumbens neurons. The mice were given optical stimulation in targeted chambers, and after 3 days the mice showed clear preference for said targeted chambers [35]. To combat the deficiency of bleaching in animal Rhodopsin, many bleach resistant Eukaryote Rhodopsin has been studied originating from jumping spiders [1], jellyfish [1], and mosquitos [1]. These Rhodopsins form a bistable product that do not need their retinals enzymatically replaced and may simply revert back to their original dark state [1].

On the other hand, microbial Rhodopsins has enjoyed a wide selection of species with versatile sequences, and functions originating from marine freshwater Archaea, terrestrial Eubacteria and lower Eukaryotes [30]. The discovery of Channelrhodopsins, specifically ChR-II from green algae, had originally sparked the feasibility of creating optogenetic tools. The first demonstrations commonly cited to be that of Boyden *et al.* 2005 [37], the team was able to precisely modulate single spikes as well as spike trains in isolated hippocampus of mice [37]. This control yielded a wide array of applications in neuroscience, and ushered in many variants of Channelrhodopsins [38–42].

Halorhodopsin (HR) is also another variant of microbial Rhodopsin that has received plenty of attention. The *Natronomonas pharaonis* HaloRhodopsin (NpHR), a homologue of HR originating from an alkalophilic bacterium, was first demonstrated for its light-driven, inward facing chloride pumps by Sato *et al.* in 2005 whom also coined its acronym [43, 44]. Due to NpHR having favorable capacity for expression, purification, and stability, most HR investigations were done through the NpHR homologue and is commonly used as a neural silencer [44]. NpHR has been used in conjunction with ChR-II as a photoinhibitor via yellow light, and photoactivator via blue light, respectively [45]. The pair combination, as demonstrated by Han and Boyden in 2007, allowed for hyperpolarization and depolarization of hippocampal neurons of rats at the millisecond timescale [45]. In a more recent study, a modern and popular version of HR, enhanced NpHR 3.0 (eNpHR3.0) has also extended the timescale to that of minutes for long-lasting inhibition of neuronal cells via modulation of light stimulus wavelengths [46], further extending the applications of HR. Since then, tens of species have been demonstrated for use in optogenetics separated into at least 6 main categories based on their function, which includes; proton pumps, chloride ion pumps, sodium ion pumps, photosensors, light-gated ion channels, and enzymorhodopsins [30]. However, one key advantage that animal Rhodopsins may offer over bacterial ones is its increase in light sensitivity, even more so than the ultra-light sensitive ChR-II variant, Calcium Translocating Rhodopsin (CatCh), or the eNpHR3.0 [47].

In light of the expansion of Rhodopsins used in optogenetics, the search for the understanding of its precise molecular mechanisms of these proteins has justifiably also received great attention. To this end, great effort has been put into the functional conversion of microbial Rhodopsins which will help reveal its structural and functional correlations as well as identify key elements of its function [18, 48]. Microbial Rhodopsins have highly conserved domains particularly within its chromophore binding pocket [30] as well as within its transmembrane helices, owing its functional difference mainly to their side-chains, structural water, and ions [48]. Despite the low sequence similarities between microbial Rhodopsins, crystal structures retain high fidelity with a root-mean-square-deviation (RMSD) of approx. 2 Å among selected cases [18, 48]. This fidelity allowed for the functional conversion of microbial Rhodopsin [48]. One of the earliest examples of conversions was performed on BR, a natural light gated outward proton pump mutated at the 85th amino acid sequence from Aspartic acid (Asp85) to Threonine or Serine. This conversion functionally turned BR into HR, an inward chloride pump despite the sequence similarity of only 68% [48]. The success of these types of studies also ushered in an even greater variety of activity profiles for specific applications [18]. For example, considerable work has been put forth into the red-shifting of Rhodopsin variants. With an absorption maxima above 800 nm, the deeper tissue penetration avoids inserting optical fiber

cables into sensitive organs like the brains of test subjects when deploying the light stimulus. For instance, Red light activated Channel Rhodopsin (ReaChR) designed through mutation exhibits an absorption maxima between 590~630 nm [49]. In addition, Retinal analogs like 3,4-dehydroretinal (A2) was modified with an additional carbon double bond within its polyene chain (C10-C11) was inserted into ReaChR and ChrimsonR (ChR mutant from algae [39]) resulting in a shift of 30 and 67 nm respectively, when compared to the original all-trans-Retinal (A1) [50]. Significant efforts has also been poured into Thermophilic Rhodopsins (TR) with additional thermostable mutations [51]. TR with quadruple point mutations displayed retention of function even at ~105.3°C. Conversions in the form of chimeras of animal-microbial Rhodopsins have also been demonstrated [18, 35, 48]. As described earlier, the chimeric bovine Rhodopsin with human and hamster adrenergic G-protein receptors converts regular bovine Rhodopsin into a light-activated control over the reward-driven behavior of mice. However, due to difficulties in repeated activation of animal Rhodopsins, BR as well as GR (*Gloeobacter Rhodopsin*) coupled with the cytoplasmic loops of bovine Rhodopsin were developed and shown to trigger Gt activation over multiple illuminations without any photobleaching [48].

In the context of optogenetics and bioengineering [18, 48], Rhodopsin and its chromophore Retinal has been demonstrated to be a versatile tool thanks to a few key factors. Perhaps the main driving force behind its popularity is the ability for Rhodopsin to be expressed in animal tissue, and even target specific neurons with selective promoters [37]. Secondly, Retinal's photoisomerization is coupled to the subtle changes to the overall protein conformation allowing for a variety of applications across the Rhodopsin family as described earlier. They are also non-toxic hence being naturally expressed in eukaryotes and prokaryotes. Its structural similarities also allow for a versatile design principle for an increased variation of functionality that may cater to highly specialized situations. All these advantages results in Rhodopsin and Retinal being a prime candidate in delivering specific requirements in monitoring and regulating neuronal activity, or signal transducers and proton/ion transport. It seems from the vast and complex diversity of these retinal-proteins an assortment of tools may be produced, some of which may not yet be realized. Therefore, the discovery of NeoRhodopsin (NeoR) in 2020 set an unprecedented further expansion of the Rhodopsin protein family towards the aspect of autofluorescence [26]. NeoR are Rhodopsins found in the fungus *Rhizoclostridium globosum* and is attached to a type III guanylyl cyclases for cyclic GMP upregulation [26, 52]. They exhibit an unusually high fluorescence quantum yield (FQY) of 20% for a Rhodopsin, with an absorption maxima at $\lambda_{\text{max}}=690$ nm and emission maxima at $\lambda_{\text{max}}=707$ nm [26]. Along with the natural fluorescence of NeoR, its naturally highly red-shifted absorbance near-IR also makes it an excellent candidate for a non-invasive deep tissue optogenetic tool, particularly for localization in the brain

[53]. Although as of the time of writing, no NeoR based optogenetic tool has been demonstrated, multiple studies have been published into the investigation of optimizing NeoR, either via retinal analogs or NeoR mutants [26, 53, 54]. For example, retinal analogs, like the 3,4-dehydroretinal (A2 retinal), incorporated into the NeoR-opsin has been demonstrated to already red-shift NeoR to 759 nm, constituting an approximately 69 nm shift [26]. At the same time, it seems that over the past few years, most efforts towards red-shifted color-tuning has been surpassed by naturally further red-shifted variants like the aforementioned ChrimsonR or Ruby-ACR (anion ChannelRhodopsin) [54, 55]. SR-I, in another example, has also been shown to only minimally red-shift post modification, even after 4 decades since its discovery [54]. Regardless, the discovery of naturally near-IR absorbing NeoR offers a wonderful opportunity for design of optogenetic tool in use with red-light. Therefore, the elucidation of both the physical and electronic structure of NeoR either by spectroscopic or theoretical approaches are currently of great interest [27, 54, 56]. It is important to note that, along with Rhodopsin's photo-intermediates, fluorescence in Rhodopsin has been achieved via directed mutagenesis and lowering of pH, where fluorescence quantum yield is linked to the pKa of Rhodopsin's counterion. However, up until the discovery of NeoR, observed fluorescence quantum yields of both microbial and animal Rhodopsin has been few and far between [57].

Outside of neural diagnostics in optogenetics and fluorescence, Rhodopsins have also received attention in diagnostics in optical biopsies. In a 2020 study by Palczewska and co-workers the team demonstrated using Two-Photon Excitation (2PE) in conjunction with Fluorescence Lifetime Imaging Microscopy (FLIM) to perform a non-invasive biopsy to detect early onset of retinal diseases in mice [25]. Current common ophthalmic imaging techniques determine retinal dysfunction based on structurally identifying physical damage within the tissue, therefore detection of disease is not possible at an earlier stage. Relying instead on the intrinsically fluorescent retinal metabolites involved in the visual process, the possibility to identify disease-specific biomarkers early on becomes feasible [25]. The team demonstrated a non-invasive identification of fluorescent diretinoid-pyridiniummethanolamine (A2E), metabolites produced from mutations of the ABCA4 transporter and retinol dehydrogenase in *Rdh8^{-/-}* mice. Additionally, *Rpe65^{-/-}* mice which contains a mutation of a retinoid isomerase that results in an over accumulation of retinyl esters, replicating the human form of Leber congenital amaurosis, were also targeted for 2PE fluorescence lifetime imaging [25]. 2PE offers the benefit of using much longer wavelengths like IR bypassing the need for the lens of the eyes to filter out ultraviolet (UV) light, coupled with modulation of the laser's pulse repetition frequency mitigated the risk for eye tissue alteration during diagnosis. In fact, 2PE or Two-Photon Absorption techniques offer many other advantages, over single photon methods (see section 1.4).

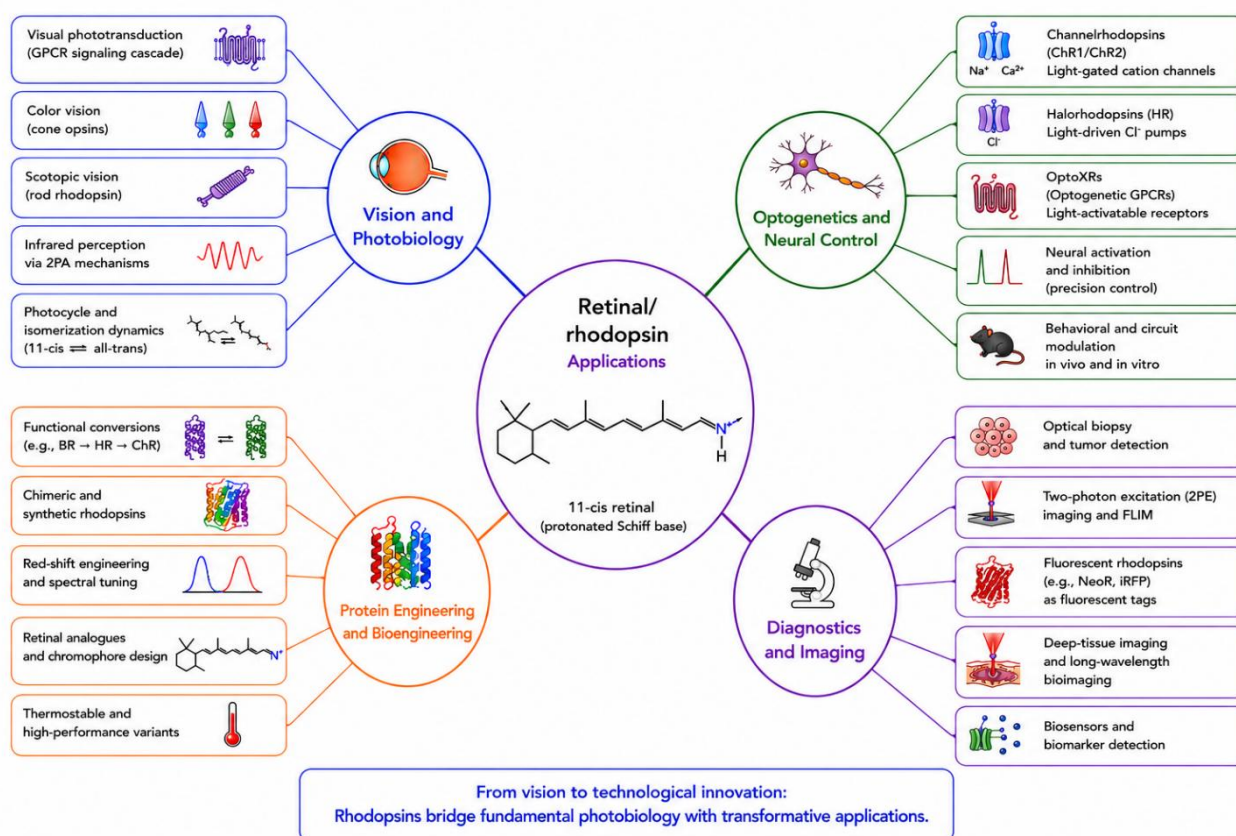


Figure 1. Schematic overview of selected retinal and Rhodopsin within photobiology and their applications discussed throughout this work.

1.2.1: Type-I and type-II Retinals and Rhodopsins structure

Microbial Rhodopsin's versatile functions observed in both its wild-types and mutated variants is rooted in the small differences in their side chains, water configurations, and ions [48]. However, microbial and animal Rhodopsins share virtually no amino acid sequence similarity to each other [58, 59]. Furthermore, the retinal in microbial Rhodopsins possess the all-trans configuration and photoisomerizes into 13-cis, while animal Rhodopsin's retinal possesses the 11-cis to all-trans configuration [1, 18, 30]. Although this differentiation is generally true, in some cases like for retinochromes found in cephalopods, photo-isomerases in honey-bees, and the retinal pigment epithelium-retinal G-protein coupled receptor (RGR), the photoisomerization process begins from the all-trans configuration and isomerize into the 11-cis form upon photo activation [60–62]. For these cases however, phototransduction is not triggered due to the proteins being photoisomerases involved in the complementary regeneration of 11-cis retinal [61–63]¹. Despite these differences, the structure of all these retinylidene share the same 7-transmembrane α -helices that is embedded within the cell

¹ Although RGR retains the seven-transmembrane architecture and retinal-binding features characteristic of rhodopsin-like GPCRs, direct G-protein-mediated signal transduction has not been demonstrated. Its established role is instead as a light-dependent retinal photoisomerase that supports visual-chromophore regeneration and may therefore contribute indirectly to sustained phototransduction.

membrane, where the N and C terminus are facing outwards and inwards respectively [30]. Both microbial and animal Rhodopsin's retinal chromophore are attached to the protein main chain via the retinal protonated Schiff base (rPSB) to the ϵ -amino group of a conserved Lysine on the 7th(or G, in microbial Rhodopsins)-helix [63, 64]. General classification refers to microbial Rhodopsins as type-1 and encompasses the eukarya, bacteria, and archaea kingdoms, while animal Rhodopsins are referred to as type-2 and includes both vertebrates and invertebrates [1, 65]. Animal Rhodopsin's (type-2) photocycle are the prototypical GPCR [30, 31]. Distinct conformational changes occur within the Rhodopsin protein resulting in multiple photointermediates that occur at a time-scale ranging from picoseconds to millisecond [24, 30, 31, 66]. Unlike type-1 Rhodopsins, type-2 regeneration of retinal is not directly cyclic and requires hydrolysis of retinal before it is reisomerized back to its 13-cis configuration either enzymatically [30, 31, 66, 67] or photochemically [60, 68].

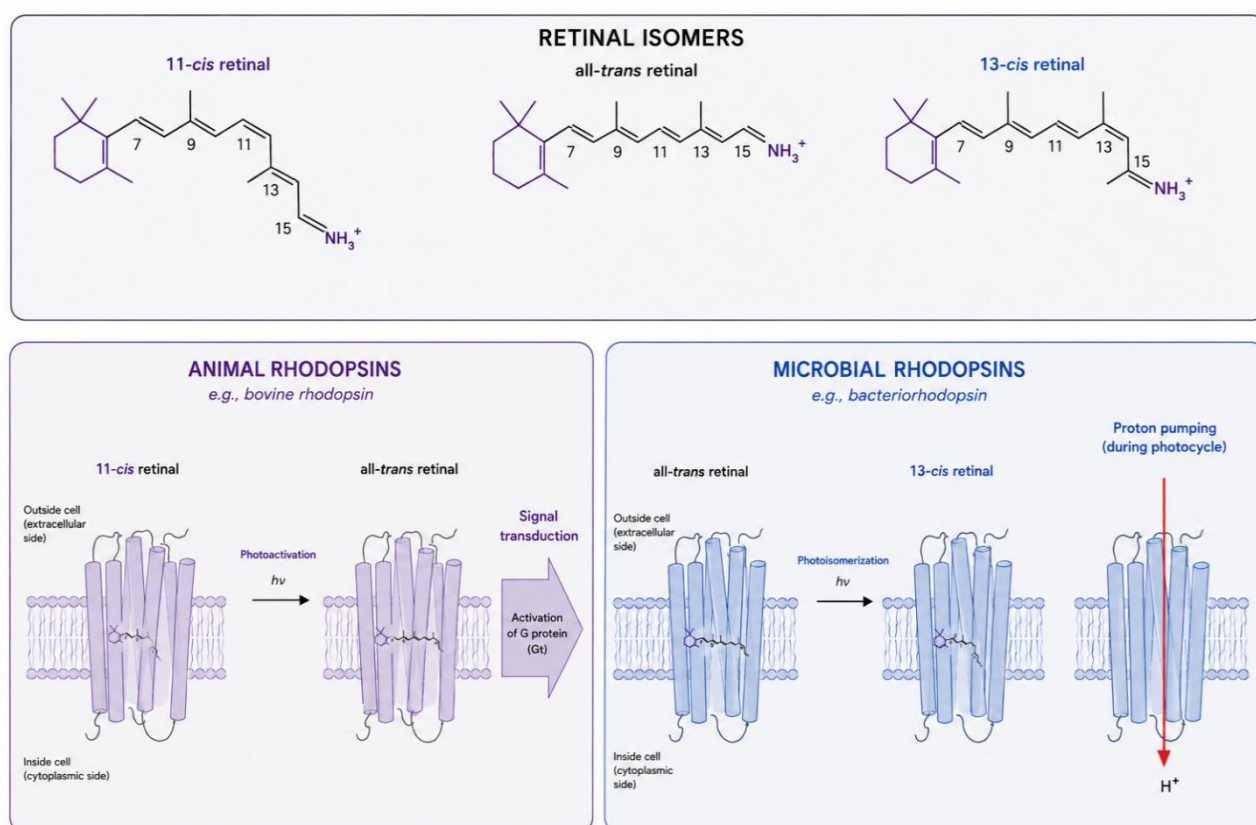


Figure 2. Schematic comparison of the retinal isomers and their photoisomerization pathways within animal and microbial Rhodopsins.

1.2.2: Retinal structure affects absorption spectra

Gas-phase absorption data play a significant role in determining spectral properties devoid of external factors, like charges and dipoles from the environment, relying solely on the molecular structure. Despite the experimental challenges associated with retinal gas-phase studies, as well as its extremely

fast photoisomerization reaction [69], relatively few *in vacuo* investigations have been performed in pursuit of understanding and tuning its spectral properties [69–74]. In 2005 the first gas phase experiment was performed for the all trans retinal by Anderson *et al.* and found the absorption max to be around 610 [69]. In a follow up publication, Anderson and co-workers also recorded the absorption band of 11-cis retinal to be about 610 nm for the $S_0 \rightarrow S_1$ transition, additional retests for the all-trans chromophore resulted in an absorption maxima at 620 nm for the same state [70]. The team also tested the $S_0 \rightarrow S_2$ transition for both chromophores which resulted in an absorption of 385 and 390 nm for the all-trans and 11-cis retinal respectively, making both the absorption bands in S_1 and S_2 states of both chromophores nearly identical [69, 70]. These results mark a significant step in elucidating the geometrical effects of retinal on its spectral properties. Previously all published results compared retinal in a methanol environment with an absorption maxima at 440 nm [74–76] to that of retinal within its native Rhodopsin environment of about 500 nm. This resulted in erroneously ascribing the opsin-shift as bathochromic, as opposed to hypsochromic as the *in vacuo* results would suggest [74]. The results demonstrate a clear electrostatic effect on the tuning of retinal and Rhodopsin's spectral properties, these topics will be further discussed in chapter 1.3 while this sub-chapter will focus only on retinal's geometrical effects.

Due to the homologous nature of the different types of Rhodopsins, it may come as no surprise that their binding pockets are flexible enough to accommodate structurally modified retinal chromophores and produce artificial pigments [77–81]. In many cases, these artificial pigments differ from the native chromophore by only the insertion or deletion of a methyl group. Despite these seemingly minor modifications, methylation patterns can strongly influence photophysical properties, including the excited-state lifetime, quantum yield, and thermodynamic stability, while retaining the characteristic photoactivity associated with retinal's conjugated π -system, 11-cis conformation, and protonated Schiff base [69, 82]. These methyl groups directly interact with the Rhodopsin binding pocket through tight and specific electrostatic interactions of the pocket's side chains. Further discussion on these specific interactions can be found in section 1.3. As noted in the sections above, retinal analogs like the A2 retinal has been experimentally incorporated into various different Rhodopsins. Introducing a modified A2 retinal into the Channel Rhodopsin variants, ReaChR and ChrimsonR, red-shifted their absorbance by as much as 67 nm [50]. The A2 retinal has an additional double bond inserted between the 3rd and 4th carbon along its polyene chain, extending its system of conjugated double bonds for an increased delocalization of π electrons. Additionally, NeoR has also been inserted with the non-modified A2 to also bathochromically shift its absorbance by nearly 70 nm [26]. λ_{max} of the 13-methyl retinal analog, where the methyl group located on the 9th carbon of retinal's polyene chain was removed (see figure 2), was regenerated in opsin by Blatz (1969) and determined to be blue shifted

by 35 nm when compared to native Rhodopsin, to a total of 465 nm [83]. In a flash photolysis experiment, the same 13-methyl Rhodopsin was also shown to have reduced G-protein activation to as little as 8% of what native Rhodopsin is capable of [78, 80]. Other analogues including the 9-methyl Rhodopsin (carbon-13 methyl removed from native retinal) and 9,10-dimethyl (methyl on carbon 9 and 10 on polyene chain only – see figure 2) were measured for their absorbances in CD spectroscopy and visible absorption spectroscopy [84]. For the lowest lying ES visible absorption spectra recorded a λ_{max} of about 498-500 nm for both structures. While CD spectra on the other hand recorded a 490 nm absorbance for 9-methyl with much lower intensity than native Rhodopsin, and the same 490 nm for 9,10-dimethyl but with over twice the intensity as Rhodopsin [84]. Other species of Rhodopsin like ChR have also received much attention in terms of red-tuning its absorption spectra [50, 85–89]. It has been known that the ChR chromophore; 3,4-didehydro-all-trans retinal (A2) shifts all-trans retinal (A1) absorbance by up to 20 nm due to the additional double bond within its β -ionone ring. Additionally, both isolated A1 and A2 demonstrated a 15-20 nm red shift when differing segments of its polyene chain were elongated by an additional C-C double bond [85]. A1 and A2 elongated at the C10-C11 position within the same study were also incorporated into the ChR 1 and 2 chimera, C1C2, opsin. This shifts the methyl group from the native C13 position to the C15 position and demonstrated a distinct red-shift by 25 and 33 nm, respectively, from their non-elongated counterparts [85]. Despite these successful shifts, the relative yield of these pigments were fairly lower than their native counterparts by a reduction in relative absorbance by 0.250 for A1 and 0.032 for A2 [85]. Additionally, the same elongated A2 at the C10-C11 position was also incorporated into ChrimsonR which resulted in a 67 nm red-shift compared to A1-ChrimsonR for a max absorbance of 654 nm [50].

Many quantum ab-initio methods have also been exploited to help better understand both the photoisomerization of retinal and its spectral properties such as excitation energies (ΔE) and oscillator strengths (f), among other properties [87–94]. According to Sugihara *et al.*, excitation energies of several retinal analogues, including 9-methyl, 13-methyl, 9,10-dimethyl, and 9,10,13-trimethyl, were computed using Complete Active Space Self-Consistent Field (CASSCF) wavefunctions combined with second-order perturbative energy corrections through CASPT2 [95]. Of the 2 analogues, namely 9,10-dimethyl and 9,10,13-trimethyl, the inclusion of methyl constituents at the C9 and C10 position causes enough of a sterical strain to destabilize the planarity of the retinal. This was later supported by Walczak *et al.* through the discovery of 2 separate ground-state equilibrium structure for the analogues with 2 or more methyl substituents [92]. The team noted the position of methyl groups along the polyene chain has a profound effect of shifting the absorbance by 5-10 nm depending on its position, similar to the elongation study of the all-trans retinal described earlier. Despite the

chromophore model being shortened at the β -ionone ring in order to reduce computational time, the general trends observed for the absorbance shifts of the analogues relative to native Rhodopsin remain the same for both the computed and experimental values. The computed absorbances are generally 50 – 60 nm blue-shifted from the experimentally observed values, however, 9,10,13-trimethyl demonstrated the largest red shift at 14 nm relative to the native chromophore, while 13-methyl exhibited the largest blue shift at 24 nm [95]. The team also added that a 20° torsion along the C11=C12 double bond caused by the steric interference of methyl groups may also constitute as much as a 10 nm absorbance shift. Therefore, the red-shift from 9,10,13-trimethyl is due to the addition of a methyl group at the C10 position, while the blue-shift in 13-methyl is the combined effect of decrease in torsion along the C11=C12 as well as demethylation of C13. The same torsional angle has also been found by Walczak *et al.* to be proportional to the change in permanent electric dipole moments ($\Delta\mu$) from the ground to excited-states ($|\mu_{ES} - \mu_{GS}|$) [92]. It seems that introducing a 10° and 27° torsion due to the analogues 9,10-dimethyl and 9,10,13-trimethyl, decreases the $\Delta\mu$ by ~0.80 D and ~1.3 D from their planar counterparts, respectively. The decrease in $\Delta\mu$ is due to the charge-transfer character of the excitation, as was experimentally observed by Irving *et al.* [69, 82], and later computed by Olivucci *et al.* [96]. In the ground-state retinal localizes a positive charge at its Schiff base, and upon excitation evenly distributes this positive charge along the conjugated double bonds within the polyene chain towards the β -ionone ring [69, 82, 92, 96]. Anderson and co also demonstrated this charge-transfer effect on retinal's absorption maxima. All-trans retinal with a neutral Schiff base, akin to stabilizing the ground-state Schiff base with a negative counterion, was recorded *in vacuo* with an absorption maxima of about 487 nm, over 120 nm blue shifted from its protonated counterpart [69].

As mentioned in section 1.1, one of the proposed gateways to reach near IR optogenetic tools could include two-photon spectroscopy in retinal/rhodopsin, and indeed a few computational methods have been employed to study their 2PA spectra [16, 88, 89, 97, 98]. However, these publications do not systematically investigate the effect of de/methylation along the polyene chain on the 2PA properties of 11-cis retinal. Thus, the author has targeted these structural modifications to systematically analyze their influence on the 2PA properties of 11-cis retinal. This pertains to objective no. 2 (see Section 2; Objectives).

1.3.1 Protein environment impact on retinal geometry

The retinal chromophore is situated in the center of the 7-transmembrane helices of Rhodopsin where its geometric and electronic properties are closely influenced by the tight and specific binding of the

closest amino acid (a.a.) residues and water molecules [73]. This results in not only the direct control of its photoisomerization quantum yield, but also its absorption spectra shape. This can be directly observed when comparing the isolated 11-cis retinal's λ_{max} of about 610 nm to that of the full Rhodopsin's of 500 nm, creating a ca. 100 nm hypsochromic shift and giving rise to the so called 'opsin-shift'. Spectral tuning can also be observed in human cone cells; the same 11-cis retinal acts as the chromophore for all 3 visible pigments, red (560 nm), green (530 nm), and blue (435 nm) [99]. Furthermore, bacteriorhodopsin (BR) and sensoryrhodopsin-II (SR-II) are also prime examples of how the environment may fine tune the absorption spectra. BR and SR-II share the same all-trans-retinal and also share a high degree of homology, yet their absorbances are fairly different at 568 nm and 497 nm, respectively [100]. Therefore, spectral properties like excitation energies, and even $\delta^{2\text{PA}}$ signals can be tuned through the understanding of the chromophore-protein interactions.

At the same time, the environment may also affect the geometries of the chromophore itself, making the elucidation of the relationship between the chromophore and its absorption spectra more difficult. In fact, not just the highly twisted analogues 9,10-dimethyl and 9,10,13-trimethyl, but Sugihara and co-workers found that all their methylated analogues [95] adhered to a small range of -15° to -26° twist along the C11=C12 double bond once bound to the opsin. Using a combination of retinal analogues and positional constraints on key amino acids, namely Lys296 and Glu113, the team demonstrated the exact contributions from the binding pocket that holds the chromophore in place. It was found that 3 main interactions contributed to the C11=C12 dihedral twist. One from fitting the β -ionone ring to the hydrophobic pocket, the second from covalently bonded Lys296 to retinal, and thirdly the formation of a salt-bridge between N16/H16 of retinal and Glu113. The formation of the salt-bridge and the locking of the β -ionone ring reduces the distance between C6 and N16 of retinal causing the twist along C11=C12, given a methyl substituent at C13 is present. Interestingly, the 9-methyl analogue further reveals that contribution to the twist mainly comes from H16's salt-bridge and Lys296. Constraint on either of these (three backbone atoms of Lys296) causes a twist of about -10° (-11.7° total), while a combination of both increases the strain to -15° . Mutating this Glu113 to alanine breaks the salt bridge between the oxygen of Glutamic acid and H16 resulting in an increased twist. At the same time mutation of Lys296 with alanine results in a more flattened structure of the retinals. However, when both residues have been mutated, the 9-methyl retinal maintains a twisted structure, proving that the binding pocket indeed supports a naturally twisted structure, as seen from the twist due to steric hindrances of methyl groups at the C9, C10, and C13 position. Applying similar scheme of constraints on the heavily twisted retinal, 9,10,13-trimethyl, it becomes clear that the binding pocket limits the amount of torsional strain that can be exerted on C11=C12. Thus the team concluded that the binding pocket has evolved to keep the chromophore in a specific twisted

conformation with great stability, reducing the differences in the analogues geometries, save for the largest difference being the dihedral angle [95]. This stability is of course only broken with the photo-excitation event.

1.3.2 Protein environment impact on the spectral properties

Many calculations have also been dedicated to the exact mechanism of how the Rhodopsin protein environment directly affects the absorption spectra via mutagenesis. Snedkov and co-workers performed computational screening of C1C2 chimeric Channelrhodopsin at the 2nd order approximate coupled cluster singles and doubles excitation (CC2) and the time-dependent density functional theory (TD-DFT) level with a polarizable force field (PE; see 3.5.3 methodology) to describe the electrostatic interactions of the retinal and the protein environment [89]. C1C2, and its mutants had their excitation energies calculated at the CC2 level, while δ^{2PA} properties used TD-DFT with the CAM-B3LYP exchange-correlation functional (xcf) [89]. Mutations along the closest residues acting as counter ions to the Schiff base, such as Asp292 and Glu162, has demonstrated the largest effects on both excitation energies and δ^{2PA} [88, 89]. In fact, both Asp292 and Glu162 displayed a relative increase of over 200% in δ^{2PA} transition strengths when compared to the wild-type [88, 89]. However, Asp292 is speculated to be the primary proton acceptor in ChR2 and may lose its function upon mutation, therefore the primary target seems to be Glu162. Mutation of Glu162 with all possible a.a. were screened for their absorbances and displayed that lysine and arginine substitutes had the greatest red-shift of up to ca. -0.7 eV. At the same time the team also noted that, due to the large size of these side chains, as well as their positive charge, glutamine may be a more realistic substitution while still red shifting by ca. -0.4 eV. In a separate study, the Poisson-Boltzmann/quantum chemical (PBQC) and charge density coupling (CDC) methods have also been used to describe the polarization of the protein environment by the retinal. Specifically, the TD-DFT/BHLYP xcf and CHARMM22 force field were used to determine absorbance shifts between Rhodopsin and its mutants [101]. Colette *et al.* chose 5 a.a. residues within the binding pocket, namely Asp83, Glu122, Ala164, Phe261, and Ala292, for mutational studies. Residues number 122, 164, and 261 were located towards the β -ionone ring, while the remaining residues were located closest to the PSB (protonated Schiff base). Aspartic and glutamic acid of Asp83 and Glu122 were mutated to asparagine and glutamine, resulting in a 2.5 and 22.5 nm blue-shift in PBQC, respectively. On the other hand, alanine and phenylalanine of Ala164 and Phe261 were mutated to serine and tyrosine, also resulting in a red-shift of ~ 1 and 7.5 nm, respectively. And lastly, alanine of Ala292 was mutated to serine for an 8 nm blue-shift [101]. In a 2016 study, Olivucci *et al.* unveiled their new modeling protocol dubbed Automatic Rhodopsin Model (ARM) for the streamlined construction of Rhodopsin

models and its mutants [102]. This new protocol allowed the comparison of absorbances and oscillator strengths between multiple Rhodopsin variants to their mutants in a streamlined matter. Although this report was aimed at benchmarking the ARM protocol against experimentally observed results, the team reported a shift of up to 74 nm, from the computed native Rhodopsin, among their mutants. Experimental and computational studies on modified human cellular retinol binding protein II (hCRBP II) fitted with an all-trans retinal via its covalently bonded PSB were also mutated for absorption spectra shifts [103, 104]. 11 mutant variants were designed and demonstrated an absorbance range of 425 to 644 nm [103, 104]. Even machine learning algorithms have been used to study and predict absorbances in microbial Rhodopsins [105]. However, none of these studies focus on 2PA properties. Only a handful of systematic computational investigations into 11-cis retinal and Rhodopsin's 2PA properties are available, none of which focus on mutational effects [16, 97, 98]. Nevertheless, many of the same electronic properties repeatedly used to rationalize shifts in absorption spectra, such as excitation energies, oscillator strengths, and dipole-based descriptors, also contribute directly to nonlinear optical response. Consequently, strategies used for tuning linear optical properties may also provide useful starting points for understanding and engineering 2PA behavior in retinal and Rhodopsin systems.

1.3.3 Electrostatics and polarizations

As briefly described above, it is rather difficult to disentangle the geometrical relationships of retinal and Rhodopsins from their spectral properties. This is due to the electrostatic forces produced by the protein environment and the factors that influence it. These factors include the long range nature of electrostatic forces, the mutual polarization between the retinal and the environment itself, and the counterions in the form of a.a. residues with positive or negative charges. For example, in the 2004 study by Sugihara *et al.*, the team modeled the protonated and deprotonated retinal with and without its immediate surroundings including a crystallographic water, covalently bonded Lys296, and the a.a. Glu113 and Thr94 [106]. Utilizing the CASSCF/CASPT2 method, the team determined that f values greater than 0.93 is generally associated with the S_1 excitation state, and is accompanied by a strong change in dipole moments both in magnitude and direction. The PSB models with and without an environment both demonstrate their dipole moment's negative charge center shifting, and pointing closer towards the Schiff base, upon excitation to the S_1 state. This suggests that there is a shift of negative charge going towards the Schiff base, similar to what was suggested with Mulliken charges by Olivucci *et al.* [96]. The dipole moment amplitudes have also changed, the isolated PSB experiences a decrease in μ (11.35 reduced to 5.80 D) and conversely the embedded PSB experiences an increase (9.27 increased to 18.52 D), upon excitation to the S_1 state. However, when comparing

both of their ground-state μ to one another, a reduction of ca. 2 D can be seen in the embedded model due to the strong negatively charged Glu113 that localizes the positive charge via a salt bridge to the nitrogen of the PSB, crucially stabilizing the ground-state [106]. This results in the S_{0-1} $\Delta\mu$ values not differing greatly between the two models. In fact, the difference between the two $\Delta\mu$ vectors is only about 0.64 D (obtained from the vector difference of the two $\Delta\mu$ values). Both models therefore exhibit large $\Delta\mu$ values, with the embedded PSB being only slightly larger, further suggesting that the main contribution of the counterions is electrostatic in nature. Additionally, to eliminate the possibility of retinal's geometrical effects on the absorption spectra, the excitation energies were compared between PSB that has been geometry optimized with the environment included but excluded from its spectral properties and PSB which includes both. The models that has only been optimized with the environment and not spectrally calculated with the environment show a blue shift of only 0.1 eV, while models that did include the environment in its spectral calculations showed a significant shift of 1 eV, similar to observations made from the opsin-shift effect [106]. The effects of these counterions also demonstrate an increase in bond length alternation (BLA) within the polyene chain but is somewhat reduced in the C9 to C12 region [106].

Other Rhodopsin mutational studies have also shown shifts in excitation energies decomposed from both electrostatic and electronic polarization (of retinal by the environment) effects [107]. Bovine Rhodopsin and its mutants were computed for their excitation energies using TD-DFT with a mechanical and electrostatic embedding scheme (ME and EE, see 3.5.3 methodology), they represent an embedding scheme that accounts for electrostatic and electrostatic plus polarization effects, respectively. Including the electrostatic effects (ME) to isolated PSB's geometry doesn't shift the absorbance by much; red-shift of ca. 8 nm (527 nm shift to 535 nm). At the same time, also including electrostatic effects to PSB's spectral calculation (that is the electrostatic effects of the protein environment is included), the absorbance is blue-shifted back by ca. 30 nm, smaller than the previously reported opsin-shifts. When polarization effects alone are included in isolated PSB's geometry a fairly large red-shift of 25 nm can be seen, combined with electrostatic effects the total shift is ca. 33 nm (527 nm shift to 560 nm total). However, when including electrostatics and polarization effects on both the geometry and spectral properties, this red-shift drops significantly to ca. 1 nm. In this case, the total net shift due to polarization effects on both the geometry and the spectra is nearly zero. This is due to the change in retinal's geometry from the polarization effects being counterbalanced by the decrease in hydrogen bond (Hbond) distance between the protonated Schiff base of retinal and carboxylic oxygen of Glu113. Thus, the major contribution to wt-Rhodopsin's spectral shift is attributed to the ca. 30 nm blue-shift due to electrostatic effects only, resulting in a net blue-shift of 24 nm to 503 nm [107]. The team notes that this counterbalance effect

helps support the localization of the positive charge on the Schiff base as well. The polarization effects are more pronounced in the mutated models where glutamic acid Glu113 and Glu122 were both mutated into glutamine, reducing the carboxylic oxygen to an amine side chain. In this case, residue 122 experienced a blue-shift due to electrostatic effects by ca 31 nm similar to the wild-type, and experienced another 11 nm blue-shift when also considering polarization effects. Residue 113 experienced a 15 nm blue and 9 nm red-shift for electrostatic and polarization effects, respectively. The team also noted that as much as a 5 nm shift is attributed to the polarization effects on geometry optimization among the two mutants [107]. Tomasello *et al.* further decomposed the counterion Glu113's role in shifting excitation energies using only electrostatic effects [108]. The team gradually increased the complexity of their system model by adding additional a.a. residues in a stepwise manner to the protein environment, going from isolated chromophore till all a.a. residues within 3.5 Å, and finally the full Rhodopsin protein. The inclusion of Glu113 displays a ca. 130 nm (~+15kcal/mol) blue-shift from the isolated retinal. At the same time inclusion of as many as 14 a.a. residues, red-shifts this back by approximately 50% (~7kcal/mol), effectively partially shielding Glu113. The team summarized that almost all of this shielding effect is attributed to a.a. residues within 2.5 to 3.5 Å, where the residues located further away do not add-up in a constructive manner to shield the 113 counterion [108]. Based on the charges of the shielding agents, the team constructed a scheme for tuning the absorbance. In order to destabilize the ground-state and stabilize the excited-state to create a smaller gap between the ground and excited-state energies, ergo creating a red-shift, positively charged residues should be placed near the Schiff base, or negatively charged residues near the β -ionone ring. The opposite is also true for blue-shifting, that is positively charged residues near the β -ionone ring or negative residues near the Schiff base would stabilize the ground while destabilizing the excited-state [108]. Furthermore, negatively charged Glu113's stabilization of the ground-state via its salt bridge has been intimately linked to the reduction of its permanent electric μ , while destabilization of the excited-state is linked to the increase in its μ for the S_1 state as suggested by Sugihara *et al.* above [106].

These findings are in line with Anderson *et al.*'s earlier isolated retinal calculations [69] and further supported by Colette *et al.* in 2018 [101]. In fact, Colette has also proposed the "OH- dipole orientation rule" for highly polar residues located near retinal. The residues 292 and 261, which were mutated to serine and tyrosine, demonstrated a blue and red-shift respectively, as was described above in section 1.3.2. The negative end of the oxygens (hydroxy) within their side chain are coordinated towards the Schiff base and β -ionone ring, leading to these absorbance shifts. These residues also correspond to the same residues investigated by Tomasello that are within 2.5 Å of retinal, and shift their absorbances in the same manner. Collete and team further describes that the orientation of the

hydroxy dipoles play a crucial role in the exact shifts of the excitation energy. For example, mutated serine 292 creates a hydrogen bond network (HBN) with a nearby water molecule, making its hydroxy oxygen within its side chain partially negative and face towards the PSB side of retinal (hence its hydrogen is coordinated by the water molecule), thus stabilizing the ground-state. This mutation to serine on residue 292 shifts the absorbance by ca. 3 nm and the reorientation of the water molecule contributes another ca. 4 nm, resulting in the aforementioned (section 1.3.2) 8 nm blue-shift [101], and thus giving rise to their proposed “dipole orientation rule”. HBN significance on spectral properties has also been highlighted by Olivucci *et al.* in Rhodopsins and its mutants [102]. Mutated Rhodopsin with a deprotonated Glu181 has also demonstrated a large blue-shift of ca. 6 kcal/mol. Hydrogen that is bonded to the carboxylate oxygen in Glu181 usually forms an Hbond with the oxygen of a nearby water. Post deprotonation and the carboxylate oxygen now coordinates with the same water’s hydrogen instead, ergo flipping the orientation of the water molecule [102]. A.a. side chain flexibility has also been identified to affect the HBN and affect absorbances, causing a large variance in recorded energies. The team found that averaging the excitation energies over at least 10 random models (frames) taken from its molecular dynamics (MD) trajectory, is sufficient to reduce the variance from side chain flexibility. Averaging beyond just 10 models only decreases the variance by less than 1 kcal/mol [102]. Furthermore, Thiel and co-workers also highlighted that HBN may contribute as much as 23% of the absorbance shift seen in the difference between the homologous BR and SR-II [100]. Despite the plethora of data available for other spectral properties, schemes directly related to 2PA transition strengths for Rhodopsins is lacking. Looking over to other photoactive proteins like the red fluorescent protein (FP) DsRed could help draw some parallels. In a study by Kongsted *et al.*, DsRed was investigated with TD-DFT and the CAM-B3LYP xcf using an EE scheme [109]. DsRed, like Rhodopsin, also experiences a redistribution of charges upon excitation to the S_1 state and is characterized by a change in dipole moments. Contrary to Rhodopsin however, isolated DsRed experiences a lower ground-state μ_0 (ca. 3.84 a.u.)² that slightly increases upon excitation to the S_1 state (ca. 3.87 a.u.). Upon embedding the chromophore within its protein environment in an EE scheme, a high ground-state μ_0 (ca. 10.98 a.u.) is observed and decreases in magnitude (ca. 9.68 a.u.) upon excitation to the S_1 state. The team notes that there is a preferential increase in the ground-state μ_0 upon embedding [109]. At the same time, Drobizhev *et al.* also observed the quadratic Stark effect in DsRed [110]. The team developed a method to calculate $\Delta\mu$ as well as the electric field strength within the FP from reported 1PA and 2PA values, and determined that ΔE is proportional to the square of permanent μ [110]. This may pose as an issue, similarly to

² Permanent dipole moment magnitudes were obtained from the Cartesian dipole moment components reported by Kongsted et al. (Table 2) [109] according to equation 4.1. Values are reported in atomic units (a.u.), consistent with the original publication.

Rhodopsin, the increase in 2PA intensity is directly correlated to the increase in $\Delta\mu$ as seen from both the two-state model (2SM) and the quadratic response theory (see methodology). This suggests that an increase in $\Delta\mu$ in DsRed due to an increase in ground-state permanent μ would likely also result in an undesirable blue-shift. Therefore Kongsted *et al.* suggests a strategy to increase $\Delta\mu$ by stabilizing the S_1 excited-state while also dampening its μ to get the added benefit of a red-shift [109].

1.3.4 Mutual polarization schemes

The use of previously mentioned methods like ME considers the electrostatic field of the protein environment as effective point charges in place of the individual atomic centers (see methodology). This neglects the higher order electronic effects like electronic polarization and charge-transfer interactions. To account for many of these properties, it is necessary to correctly model the local electric field experienced by an embedded species, like retinal. The local electric field is affected by the interactions involving the chromophore retinal, its environment, and an external field, in the case of 2PA being the incident light. However, the protein environment is not only polarized by the external field but also by its interaction with the perturbed retinal, further contributing to the local field. The following are examples of how including these mutual polarization effects may change the spectral properties, detailed discussion of their methodology and their issues can be found in section 1.4.3. Environmental effects in terms of electrostatic and mutual polarization have been decomposed by Sneskov and co-workers in the same publication as their C1C2 mutational studies on δ^{2PA} and excitation energies [89]. The team calculated ΔE as well as δ^{2PA} in 4 different models: *in vacuo*, *in vacuo* but with the geometry of retinal influenced by the protein environment, in protein with electrostatic effects up to the quadrupole moment, and the same electrostatic effects with additional polarization of both the retinal and environment. The team deduced that firstly, isolated retinal that has its geometry affected by the environment experiences a small BLA increase of 0.05 Å which red-shifts ΔE by -0.05 eV when compared to the ungeometrized one. Secondly, this model also experiences a huge increase in δ^{2PA} signal by over 3 times. On the other hand, comparing isolated retinal to the model with electrostatic effects included, an expected large blue-shift of ca. 1.05 eV can be seen, and is also accompanied by a reduction of δ^{2PA} by 40%. Finally, when looking at all effects included, that is electrostatic and mutual polarization effects included in the interaction between the retinal and its environment, ΔE is only increased by 0.84 eV, red-shifted by 0.24 eV from purely the electrostatic effects, as well as an increase back to 100% relative δ^{2PA} (that is the same 2PA intensity as isolated retinal) [89]. In a separate study, Olivucci and co used the extended multiconfigurational quasi-degenerate 2nd order perturbation theory (XMCQDPT2) method with a mutual polarization embedding (PE) scheme to calculate σ^{2PA} cross-sections via the sum-over-states (SOS; see

methodology) approach for the bovine Rhodopsin [97]. The team generally focuses on the effects of transition dipole moments (μ_{01}) on σ^{2PA} , but also report permanent electric dipole moment values. The team notes that large transition dipole moments originating from an S_1 state can heavily influence an increase in σ^{2PA} values of the 2nd and 3rd ES. For example, the transition dipole moments when going from the $S_1 \rightarrow S_2$ to the $S_1 \rightarrow S_3$ transition increases from 5.0 to 7.5 D respectively, and is also accompanied by an increase in σ^{2PA} from 231 to 771 GM. Additionally, the team also reported fairly large ES permanent dipoles of 15.0, 14.0, and 12.6 D, for the S_1 , S_2 , and S_3 states respectively. Although both the ground-state and change in permanent dipole moments were not reported the decrease in ES permanent μ did not result in a decrease in 2PA signal. Instead, an increase of over 1.5x as seen from S_1 's 2PA signal of 472 GM versus S_3 's 771 GM, likely due to the enhanced transition dipole moment [97]. With limited number of investigations into σ^{2PA} effects of Rhodopsin, the Dsred FP investigated by Kongsted *et al.* is once again used to demonstrate the significance of modeling the correct electric field representing the environment [109]. The team calculated additional embedding schemes other than the electrostatic embedding (EE) one described above, including a mutually polarizable scheme (PE; chromophore region and environment mutual polarization – see methodology), and the frozen polarizable environment (FPE), an extension to EE where the electronic degrees of freedom of the environment has also been relaxed due to the changing electron density of the chromophore region, but only for the ground-state. It was demonstrated that both EE and FPE fail to recover PE's 2PA intensity by approx. half mainly due to the loss of $\Delta\mu$ in both EE and FPE. More specifically, in going from EE, where the chromophore region is polarized by the environment, to having a mutual polarization scheme in the ground state only (FPE), an increase in ground-state μ_0 of over 0.20 a.u. is observed, from 10.98 to 11.21 a.u. as recalculated from the vector components reported in Table 2 of Ref. [109]. Further inclusion of the electronic relaxation of the environment upon excitation dampens the excited-state μ in the PE scheme even more, that is, both EE³ and FPE experience an enhancement of the excited-state μ by ca. 0.20 a.u. when compared to PE. Thus resulting in the PE scheme having the largest δ^{2PA} (18.10 a.u.) and the largest $\Delta\mu$ (1.906 D) among the three schemes [109]. Furthermore, Kongsted and co noted that the angle between $\Delta\mu$ and μ_{01} are also nearly aligned, adding to the enhancement of 2PA signal when using 2SM. Interestingly, FPE still has a slightly lower δ^{2PA} signal (ca. 0.20 a.u.) than EE despite having a slightly larger $\Delta\mu$ of over 0.13 a.u. This could be attributed to the difference in transition dipole moments (μ_{01}) where electrostatic embedding's is slightly larger by 0.14 a.u., similarly to what was observed by Olivucci *et al.* in Rhodopsin [97].

³ Kongsted refers to EE as non-polarizable embedding (NPE)

The cumulation of the previously discussed publications provides an outlook to the main factors that may affect spectral properties and may also extend to δ^{2PA} . Clearly, the charge distribution from the Schiff base to the polyene chain upon excitation is of vital importance in enhancing δ^{2PA} signal. Many of the examples above points to a rational design via enhancing this charge distribution by stabilizing the ground or destabilizing the excited-state in order to induce a large $\Delta\mu$. An issue one may face however, is an inadvertent blue-shift due to the increased energy gap between the ground and excited-state as a consequence of shifting the ground-state to a lower energy via stabilization or higher energy via destabilization of the excited-state. In order to achieve both an increased 2PA intensity and an absorbance red-shift one would need to destabilize the ground-state while simultaneously dampening its μ or stabilizing the excited-state while increasing its μ . Therefore, the author also aims to elucidate the relationship of these other spectral properties like $\Delta\mu$ and ΔE to 2PA strengths within retinal/rhodopsin and pertains to objective no. 3 (see section 2; objectives). Additionally, the author does not focus on mutational effects of the protein environment interacting with retinal in this study. However, the reliability of these effects on Rhodopsins spectral properties depends heavily on the careful treatment of Rhodopsin and retinal's interaction. Therefore, investigation into their treatment pertains to objective no. 5.

1.4.1: Two-Photon Absorption has many benefits over One-Photon techniques

The benefits brought by, and the applications in which Rhodopsins are utilized in, complement the advantages of 2PA, further highlighting its advantages over its single-photon counterpart. 2PA, as the name suggests is the process by which 2 photons of light is nearly simultaneously ($\sim 10^{-18}$ - 10^{-16} s) absorbed by a light-activated molecule, such as a chromophore [111, 112]. In the case of Rhodopsin, the chromophore retinal releases this absorbed energy in the form of isomerization. 2PA and other multiphoton processes are often referred to as nonlinear optical processes due to its quadratic dependence on light intensity [111]. Light intensity being a measure of photons in a unit area per unit of time means that 2PA requires light intensity squared at a certain focal point compared to that of a single photon process. This in turn allows 2PA's activation to be most probable at this focal point while greatly dropping off at further distances [111, 112]. In fact, 2PE probability within the focal plane is 10^7 times more likely than that of the same experimental setup but with an unfocused laser beam [111]. This intensity-squared dependence is what gives rise to the localized nature of 2PA, and the advantages that comes along with it, such as the so called 'optical-sectioning' capabilities for intrinsic 3-dimensional control of the focal point. This reduction in excitation of molecules outside of the focal point enables 2PA to be used in highly delicate cells or tissues due to localizing

photobleaching and photodamaging effects to only the focal point [112]. Additionally, the depth to which 2PA may be used in samples is also increased. At lower depths of heterogeneous biological materials scattering of the incident light is increased. Yet, due to the intensity-squared dependence of 2PA processes, scattering does indeed increase with depth but photons scattered to out of focus plane are extremely unlikely to generate background signals like unwanted excitations [112]. Longer wavelengths, such as those within the NIR or IR region, also compounds to 2PA's reduction of light scattering for deeper tissue penetration. The use of two photons allows the laser power to be reduced by half, essentially doubling the wavelength, which not only reduces phototoxic effects but also reduces scattering. This places 2PA applications on Rhodopsins comfortably within the +700 nm range.

1.4.2: Experimental measurements of 2PA

Despite these aforementioned perks of 2PA techniques over one-photon absorption (1PA) ones, the road to achieving 2PA has not been without its challenges. In both theoretical calculations and experimental measurements over the past two decades, 2PA techniques has been encumbered by persistent difficulties. 2-photon absorption spectra measurements can be particularly challenging especially for the fluorescent proteins (FP). Whereas the 1PA spectra and extinction coefficients of fluorescent proteins are comparatively well established, reported 2PA cross-sections vary substantially between studies [113]. For the enhanced Green Fluorescent Protein (eGFP), reported σ^{TPA} for the lowest lying excited-state ($S_0 \rightarrow S_1$ transition) vary between 1.5 GM [114], 20 GM [115], 30 GM [116], 40 GM [117], 60 GM [118], 180 GM [111] and 300 GM [119] depending on the source. As was suggested by Drobizhev *et al*, the root of these discrepancies stem from a few experimental procedural errors [113]. Firstly is the type of measurements used, 2PA measurement can be broadly categorized into direct and indirect methods [120]. Direct 2PA methods observe the decrease of the intensity of the incident beam itself, and include methods like nonlinear transmission techniques (NLT) and Z-scan [113, 120]. Meanwhile indirect methods measure the various possible decay path of the molecule itself as it returns from the excited to the ground-state, or directly probes the excited-state itself [120], and is commonly associated with the 2-photon excited fluorescence (2PEF) method [113]. Although using direct methods may seem to be the most straightforward path, the requirements of high intensity laser light, as well as high concentration of proteins, introduces not only further uncertainty over the concentration of proteins in the ground versus excited-state but also spurious secondary nonlinear optical processes like thermal lensing, stimulated emission and scattering, excited-state absorption, among others [113, 120, 121]. Indirect methods like 2PEF however, require less intense incident laser light, lower concentration of proteins [113], as well as elimination of other

undesired nonlinear optical processes that may accompany 2PA [121]. The other primary reason for 2PA discrepancy can be traced to the determination of the concentration of mature proteins [113]. Conventional protein concentration is commonly done by bicinchoninic acid assay (BCA assay) or alkaline denaturation. These techniques introduce errors due to also considering misfolded and non-mature proteins leading to an underestimated extinction coefficient under Beer-Lambert's law [113, 122]. Instead, Drozhizev *et al.* demonstrated using an all optical approach with the Strickler-Berg equation to relate fluorescence radiative lifetime and quantum yield to the extinction coefficient where neither the lifetime nor the quantum yield rely on the total protein concentration [113, 123]. The different issues presented above describe how experimental measurements may fall short in providing a conclusive and meaningful insight to the mechanisms governing 2PA properties. Therefore, computational methods provides an alternative tool that may synergize and complement experimental results for a more complete picture.

1.4.3: Theoretical calculations of 2PA

Traditional photobiology or photochemistry utilize different analytical techniques to uncover the mysteries behind various reactions. When considering the interaction of light with matter, spectroscopic type of techniques, like the 2PA spectroscopy, are the preferred choice. These powerful techniques allow the disentanglement of the relationship between the chromophore and its flexible and complex macromolecular structure, i.e. its protein environment. However, these techniques generally provide a global statistical image of the situation rather than a direct mapping of molecular to electronic description of the reaction. Computational methods provides a useful complement to help interpret these results but is also dependent on the important trade-off between accuracy and efficiency (time spent running calculations, computational resources, etc.). 1PA properties (f), including change in vertical excitation energies (ΔE), and 2PA properties (δ^{2PA}) of Rhodopsins and isolated retinals have often been described using multireference (MR) methods, which generally provide the most reliable results for states exhibiting substantial multireference character. Various methods based on the multireference 2nd order perturbation theory like the complete active space (CASPT2), extended multiconfigurational quasi-degenerate (XMCQDPT2), and N-electron valence state (NEVPT2), have been used to calculate vertical excitation energies [90–94, 124–126], oscillator strengths [97, 125, 126], and even 2PA [16, 97] properties from the ground to the lowest excited-state for reduced [90–94] and full sized structures [97, 125–127] of the retinal protonated Schiff base (rPSB). These methods were used in conjunction with multiconfigurational approaches like complete active space self consistent field (CASSCF) to get the full dynamic and static correlation of electrons. At the same time, conventional CASSCF treats all possible electronic configurations within the active

space, which often leads to system size issues as the number of determinants in the wave function increases exponentially with the system size. Modifications to CASSCF with the density matrix renormalization group (DMRG) algorithm was able extend the available active space beyond the usual limitation of approximately 16 electrons and 16 orbitals. Although the DMRG algorithm has only been demonstrated in a small number of cases in retinal [127], similar conjugated double bonded systems like carotenoids are well suited for this method [128]. Additional methods based off quantum monte carlo (QMC), 2nd and 3rd order variants of coupled cluster (CCSD, CCSD(dt), CC3), and multireference configuration interaction with singles and doubles excitation (MR-CISD) have also been tested on the smaller rPSB models like PSB3 (3 indicates the number of conjugated double bonds within the reduced size model, see figure 2 and figure 4) [90, 91, 124]. These high accuracy calculations provide some of the best reference data available but at the cost of huge computational demands. Although perturbative methods like PT2 have been demonstrated to compute 2PA in systems as large as 50 or more atoms [97], other methods like the 3rd order coupled cluster methods like CC3 are even less computationally feasible and pushes the size down even further to the 20s of atoms, run over a 1000 hours, depending on the software used [129]. Regardless, the size limitation itself is enough to prohibit the applicability of these methods in realistic biological systems without extensive tinkering within the methodology.

One such possible tinkering is the use of lower accuracy (more approximated) but higher efficiency methods capable of calculating linear and nonlinear optical properties. Methods such as the 2nd order approximate coupled cluster singles and doubles excitation (CC2) and the time-dependent density functional theory (TD-DFT) have both gained increasing popularity for this very purpose [18, 92, 94, 129–138]. The use of these robust methods is accompanied by the verification or validation of their accuracy and the choice of exchange correlation functions (XCF) inherent in the DFT method (see Methodology and below). Many chromophores including ones from fluorescent proteins as well as Rhodopsins have had both their electronic properties and geometries extensively benchmarked at the CC2 level against higher level methods [94, 129–138]. Walczak and co-workers compared the ΔE values in isolated retinal *in vacuo* computed at the CC2 level of theory against CASPT2 [92]. The results demonstrated good agreement to the reference CASPT2 with geometries optimized at the Møller-Plesset second-order perturbation (MP2) level of theory, where the mean absolute error (MAE) of all retinal models was 0.02 eV. Valsson and Fillipi also demonstrated CASPT2 results being in line with experimental absorption performed in gas phase for the full 11-cis retinal (retinal model terminated with a single methyl group). Walczak and co-workers's results also corroborated the findings of Aquino *et al.* where CC2 results using the PSB4 (see figure 4) are able to recover MR-CISD with pople's size extensivity correction (+P) results quite well, with deviations as small as 0.02

eV [136]. On the other hand, Aquino *et al.* also noted that CC2 generally overestimate doubly excited-states in all PSB models [136]. CC2 and XMCQDPT2 using protonated and unprotonated cis and trans-retinal models have also been compared by Wolanski and co-workers against experimental measurements [137]. The results showed CC2 overestimating protonated forms of retinal by 0.13-0.17 eV, and under estimating deprotonated forms by 0.12-0.21 eV for the lowest lying excited-state, while the second lowest excited-state is over estimated by 0.13-0.16 eV. Interestingly, CC2 generally performs better than CCSD in retrieving ΔE values of Multi State (MS) -CASPT2 [135]. CC2 deviates no more than 0.06 eV in reduced PSB models, e.g. PSB3 – PSB5, while CCSD deviates by nearly 0.20 eV. [135]. Furthermore, in a benchmark study of 28 organic chromophores encompassing 152 singlet and 71 triplet (223 total) states by Thiel and co-workers, CC2 also performed statistically slightly better than CCSD when compared to either of the reference, MS-CASPT2 and CC3 with the TZVP basis set [139]. The benchmark study showed that CC3 provided the best overall agreement with the reference data for singlet excited states, followed by CC2 and CCSD. Average deviations from MS-CASPT2 were approximately 0.2 eV for CC3, 0.3 eV for CC2, and 0.5 eV for CCSD [139]. Although CC3 remained the most accurate method, the comparatively small deviation of CC2, combined with its substantially lower computational cost, supports its widespread use as a practical alternative for larger chromophore systems [139]. A similar trend was observed for the 71 triplet states dominated by single excitations. CC3 exhibited the lowest mean absolute deviation (MAD) relative to CASPT2 (0.08 eV), followed by CCSD (0.16 eV) and CC2 (0.19 eV) [139]. Despite the superior accuracy of CC3, the performance of CC2 remained competitive when considering its substantially lower computational cost [139].

Deviations of CC2 and CCSD results are a little less pronounced in terms of oscillator strengths, f . In Thiel's benchmark study of various chromophore classes, oscillator strengths in CC2 and CCSD with the TZVP b.s. has an r-value of 0.995, showing no systematic deviation between the two methods. CC3 oscillator strengths were not computed within this study, instead when compared to the MS-CASPT2, f in the majority of states is higher than the CC methods [139]. Additionally, the group notes that although oscillator strengths are largely dependent on basis sets (b.s.), when comparing f available in literature, derived from the same CASPT2 method but with differing b.s., results are generally in the same range with an r-value of 0.9102 [139]. These results are further supported by Walczak *et al.*, where CC2 results deviate from the reference SS (single state) -CASPT2 anywhere from as low as 0.01 to 0.16 for various retinal models reduced at the β -ionone ring [92]. In spite of this, the team noted that the change in dipole moments from ground to excited-states ($\Delta\mu$) are severely underestimated by as much as half, going from 9.53 D in CASPT2 to 5.40D in CC2 [92]. To support SS-CASPT2's (S-IPEA shift) reliability as reference for CC2, their results have also been

validated against higher level methods, like CASPT3 as well as MR-CISD+P, in retinal's PSB3 model [93]. In terms of f , CASPT2 performs quite well when compared to MR-CISD+P, being underestimated by only 0.007. This value, however, is increased when compared to CASPT3, to 0.032. Similar to f , ΔE of CASPT2 is also underestimated when compared to reference CASPT3 by 0.16 eV, and it is even more underestimated when compared to MR-CISD+P by almost twice at 0.32 eV [93]. $\Delta\mu$ on the other hand was not computed at the CASPT3 level, but when compared to MR-CISD+P (albeit with differing b.s.), CASPT2 underestimates it by approx. a tenth of a Debye, depending on the method used to resolve the geometry [93].

In spite of the decent amount of benchmark calculations with respect to excitation energies and oscillator strengths, quadratic response based 2PA reference calculations are seldom performed except for a limited number of publications [97, 131, 138, 140, 141]. As mentioned above, the constraint in high-level methods like those within Coupled Cluster are limited by the exponential increase in computational time as one climbs the CC hierarchy. CC2, CCSD, and CC3 scales at a rate of N^5 , N^6 , and N^7 , respectively, where N is the number of basis functions and thus limiting the higher level calculations to the 10s of atoms when computing 2PA. Nonetheless, Paterson and co-workers benchmarked CC2 with the Resolution of Identity approximation (RI-CC2), as well as CCSD, against CC3 for small molecules including formaldehyde and diacetylene for their lowest excited-state, as well as multiple states in water. CC2 with Dunning's b.s. in diacetylene for the $1\Pi_g$ state overestimates CC3 by as low as 8%. At the same time CC2 with double augmented Dunning's b.s. consistently overestimates formaldehyde's $1A_2$ state by over twice as much as CC3 regardless of increasing b.s. size (up to quadruple zetas). Water had performed the worse out of the three in regards to the lowest excited-state ($1B_1$), where CC2 overestimates CC3 by almost 4 times. At the same time, higher excited-states can increase in discrepancy even more either due to resonance enhancement effects or contributions from triples excitations. In fact, triples excitations in CC methods can contribute as much as 15-20% to 2PA strengths, depending on the state [138]. Furthermore, the team also demonstrated a strong dependence of CC2 on the atomic basis set, highlighting the importance of including at least a single diffuse function, as well as multiple zeta functions (2 or more) to describe the valence orbital [138]. In an article by Zalesny and co-workers a set of 30 organic molecules of various sizes ranging up to 6-non Hydrogens within its structure, encompassing 82 singlet states total, were benchmarked for their 2PA transition strengths by CC2 and quadratic response/Equation-of-Motion (EOM) CCSD against the CC3 method [141]. The team showed that CCSD performed markedly better than CC2 for all singly-excited-states where the mean absolute percentage error (MAPE) ranges between 15-20% for the two CCSD variants and rises to over 40% in CC2. Considering only Rydberg states, the MAPE for CCSD drop to 10-12% while CC2 remain in the 40%

range, however, excellent correlation R^2 value is demonstrated in all 3 methods and are in excess of 0.95. Valence states on the other hand show a general increase in MAPE and decrease in R^2 for all 3 methods, demonstrating a higher difficulty in describing valence states over Rydberg. Interestingly, CCSD more than doubles its MAPE from Rydberg states to 25-37%, while CC2 only marginally increases it to ca. 54%. Regardless, R^2 for valence states still maintain excellent correlation of 0.95 or above for all 3 methods [141]. CC2 has also been measured for its reliability in computing states with a doubly excited character, marked by its %T1 assessment being lower than 80% (at least 20% contributions from double excitation). Generally CC2 and CCSD fail to properly describe states with doubly excited character due to its intrinsic nature (see methodology), this benchmark is no exception. Similar to what was noted by Aquino *et al.* in PSB [136], Zalesny and co demonstrated CCSD and CC2 with quadratic response theory had an Absolute Percentage Error (APE) of ca. 100% and over 200%, respectively in both butadiene ($1A_g$) and cyclopentadiene ($1A_1$). The lowest error was found in benzene ($1E_{2g}$) and recorded an APE of 15 and 28% for CCSD and CC2 respectively. In summary, while the performance of CC2 on spectral properties is highly dependent on the compound, excitation character, and state investigated, CC2 generally provides a reasonable compromise between accuracy and computational cost. Although CCSD typically provides more accurate δ^{2PA} values, CC2 often recovers excitation energies well and has demonstrated good agreement with multireference reference methods in retinal systems. Combined with its substantially lower computational cost, this makes CC2 attractive for studying larger chromophores and identifying qualitative trends in optical properties.

In another attempt at mitigating computational cost, TD-DFT has also been fitted with the quadratic response formalism for 2PA calculations. The benefits of the highly efficient DFT has recently been implemented in large scale qualitative estimations of 2PA strengths (up to thousands of atoms) [141], however its results are dependent on its choice of xcf. The choice of xcf adds an additional layer of difficulty into consideration when using DFT methods, but also benefits from the availability of many different functionals, whose performance may vary depending on the system and property under investigation. Therefore, a significant step when using TD-DFT is to benchmark the functionals in consideration against higher level methods. Laurent and Jacquemin discusses in a review the strategies in selecting the most accurate and adequate xcf through multiple benchmark studies for computing optical properties in both organic and inorganic molecules [142]. The duo notes that a singular universal xcf doesn't exist, but rather a handful of factors that can influence DFT performance, that is: molecular species and property under investigation (structure), transition character of excited-state, source of reference, and the type of xcf. The choice of xcf boils down to increasing the complexity of the approximation for electron exchange and correlation energy by

climbing up a metaphorical ladder of 5 total steps, as was proposed by Perdew and Schmidt [143]. The foundations of DFT developed by the Hohenberg-Kohn theorem are traditionally computationally implemented through non-interacting Kohn-Sham orbitals (see methodology), where the lowest rung on the ladder is represented by the simplest functional: the local-density approximation (LDA). LDA along with the next two rungs on the ladder, generalized gradient approximation (GGA) and meta-GGA can be considered as pure xcf, xcf that do not contain exact exchange energy (EXX) contributions from Hartree-Fock (HF). The 2 next rungs up are the hybrid xcfs and the non-local or also referred to as long-range corrected hybrid (LCRH) functionals. Hybrid xcf include a fixed percentage of EXX while LRCH have an increasing percentage of EXX with increasing electron distance. LRCH has been noted to significantly improve charge-transfer character of excited-states [143], but also adds an additional attenuation parameter ω which controls the threshold distance of electron interaction.

Thiel's set of 28 molecules benchmark was used as a basis for investigating many xcf, Laurent and Jacquemin compiled over 40 different TD-DFT xcf and their mean signed/absolute errors (MSE/MAE) compared against Thiel's theoretical best efforts (TBE) in the TZVP b.s. They noted that in going from pure xcf to increasingly larger contributions from EXX, in excess of 50%, as well as LRCH with ω in excess of 0.3 a.u., the xcfs go from underestimating (pure) to overestimating (hybrids/LR) Thiel's excitation energies for the singlet state. The two adds that these effects would not be limited to Thiel's set of molecules hence, increasing EXX in hybrid functionals would provide a more localized description of the excited-state. Furthermore, the two points out to Grimme's functional, B2PLYP, as having the smallest MAE of 0.22 eV, and coming in close are the hybrid functionals with 20-30% EXX like PBE0 (0.24 eV) and B3LYP (0.27 eV), while range-separate functionals like the CAM-B3LYP performed marginally worse (0.31 eV) [143]. In terms of triplets state, results are systematically underrated regardless of the choice of xcf, and are generally larger than for the singlet states. Out of the 63 triplet states tested, the two noted that M05/M06-2X, BMK, M11-L, and DFT/MR-CI performed the best with MAE slightly below 0.30 eV. Peach and Tozer in a separate study determined that applying the Tamm-Dancoff Approximation (TDA) of TD-DFT can help alleviate these problematic cases of triplet states [144]. On the other hand, Bredas *et al.* also tested DFT-Tight Binding (DFT-TB) Thiel's 63 triplet states and determined an MAE of only 0.07 eV, perhaps the best performance for excited triplet states [145]. Rhodopsin and its modified rPSB has also been the subject of several DFT benchmark studies [92, 129, 130, 136]. Walczak *et al.* also compared their aforementioned [92] retinal analogues using 3 xcf; BLYP (pure), B3LYP (hybrid), and CAM-B3LYP (LRCH) against CASPT2 excitation energies. The team determined that at the MP2 level geometry, the max error can be quite pronounced and ranges from 0.07 eV in BLYP to

0.31 eV for CAM-B3LYP. The largest errors come from the twisted structures (i.e. twisted 9,10-dimethyl and 9,10,13-trimethyl) and is compounded by the difficulty in describing the charge-transfer character of the S_1 state by the functionals themselves. Comparing the MAE of these 3 xcf an interesting pattern of lower EXX being more in line with CASPT2 results, BLYP has an MAE of 0.04 eV, B3LYP is slightly larger at 0.16 eV, while the range separated CAM-B3LYP is at 0.27 eV. The team concluded that BLYP performed the best in recovering CASPT2 energies due to cancellation of errors from an underestimated BLA and tendency for TD-DFT methods to overestimate excitation energies in charge-transfer states of PSB5 models [92]. Additionally, the team finds that when using CASPT2 excitation energies with varying ground-state geometry methods, B3LYP showed the most reliable geometry equilibrium with a small max error of 0.05 eV when compared against MP2 models. CAM-B3LYP comes in as a close second with the same max error and MAE but slightly higher minimum error and overestimated BLA. And in the last place, BLYP with the least accurate equilibrium structure causing an overestimation of as high as 0.12 eV [92].

Oscillator strengths computed at the TD-DFT level has been benchmarked in multiple studies. In small molecules like N_2 , CO, formaldehyde, and more, Tawada *et al.* tested pure, hybrid, and LRCH functionals against experimental and symmetry adapted cluster-CI (SAC-CI) results and found three LRCH functionals to be most successful [146]. In slightly larger molecules like benzene, phenol, aniline, and fluorobenzene tested by Miura *et al.* compared 7 different functionals against CC and experimental results. The team concluded that xcf with more EXX produced more realistic f and recommended the BHLYP functional [147]. Thiel's group also benchmarked 28 molecules over multiple excited-states and noted that B3LYP outperforms BLYP [148, 149]. In an extensive study by Caricato *et al.* 28 xcfs were tested on 11 small organic molecules encompassing 30 valence and 39 Rydberg states against EOM-CCSD. The team noted that within carbonyls, alkenes, and azabenzenes subgroups, on average CAM-B3LYP performed the best [150]. Walczak *et al.* also corroborated many of the findings above with PSB analogues at the S_1 state. The team noted that their reference CASPT2 demonstrated large f that is closely followed by DFT, starting with CAM-B3LYP being the largest, followed by B3LYP, and the smallest being BLYP in both f and BLA. The team also noted the twisted structures (i.e. 9,10-dimethyl, 9,10,13-trimethyl) have smaller f than their planar counterpart [92].

Permanent electric dipole moments on the other hand is rather difficult to be properly captured by TD-DFT methods. 7 diatomic and tetra-atomic molecules encompassing 10 excited-state dipole moments were compared against experimental values by Furche and Ahlrichs using three pure (including BLYP) and two hybrid (including B3LYP) functionals. The pure hybrids demonstrated

MAD of 0.11-0.12 D, while the hybrid functionals performed slightly worse at ca. 0.17 D [151]. Thiel's benchmark study of 28 molecules compared against CASPT2 demonstrated an MAD for B3LYP and BHLYP of 0.59 and 0.61 D, while also noting that 12 and 13% of all cases respectively have deviations larger than 1 D [148]. Thus alluding to xcf with less EXX being more accurate than ones with larger contributions for the selected scenarios. On the other hand, pure and hybrid functionals were also tested against EOM-CCSD by King in furan and pyrrole, where the hybrid functional (B97-1) performed better in excited-state dipoles with a deviation of 0.5 and 1.1D respectively [152]. Furthermore, for the lowest ES in larger molecules like N-Phenylpiperidone-malondinitrile, Rothlisberger and co-workers reported a small deviation of 0.1 and 0.2 D in the hybrid PBE0 functional when compared against CC2 and experimental values, respectively. These deviations of PBE0 is much less pronounced compared to its pure counterpart, PBE where deviations rise to ca. 2.5 D [153], contrary to what was seen in the smaller molecules earlier. In addition to this, Laurent and Jacquemin also notes in their review study that TD-DFT tends to overestimate ES dipole moments [142]. Walczak *et al.* also notes that TD-DFT also has trouble capturing ES with charge transfer character like those in their PSB5 analogue set. More specifically, the team demonstrated that hybrid functionals like CAM-B3LYP outperforms pure functionals like BLYP but is still underestimated by 3-6 D when compared to the CASPT2 reference for $\Delta\mu$ [92].

2PA values in DFT has also been benchmarked in a handful of studies. Most likely the first benchmark study in 2PA for TD-DFT methods was performed by Masunov and Tretiak in 2004, where the duo used the hybrid B3LYP xcf on 15 different polyconjugated molecules, including stilbenes and bis(styryl) benzene derivatives, compared against experimental values [154]. Similarly to what was noted above by Walczak *et al.* [92], Masunov and Tretiak demonstrated B3LYP underestimating the BLA within the tested compounds geometry which also results in an underestimation of its transition energies creating a red-shift in both 1PA and 2PA. However, the team notes that MAEs are quite low for 2PA and 1PA's transition energy at 4.6% and 6.9%, respectively. In terms of 2PA cross-sections itself, the two noted that TD-DFT and experimental results agree relatively well, particularly in terms of following experimental trends with an exception of a single outlier due to doubly excitations in unsubstituted stilbene. However, calculated values may diverge by as much as an order of magnitude to as little as a few percentage with no systematic patterns in their deviations [154]. Paterson and co-workers also tested pure (BLYP), hybrid (B3LYP), and range separated (CAM-B3LYP) xcf in their 2PA benchmark [138]. Across all functionals, the significance of including a single diffuse function in augmented basis sets is demonstrated by an increase of an order of magnitude towards convergence to CC3. The difference between the functionals are also large. In formaldehyde's 1A2 state, BLYP overestimates CC3 by nearly 5 times. In going to B3LYP, this deviation is decreased significantly

but is still overestimated by double. CAM-B3LYP on the other hand converges to CC3 with only a 15% error margin, by far the best result [138]. Similarly, diacetylene's $1\Pi_g$ state also converges in the same manner as formaldehyde. BLYP produces δ^{2PA} that is twice as large as the reference CC3, B3LYP reduces this deviation slightly, while CAM-B3LYP reduces it to 10%. Additionally, similar to lower tier CC methods mentioned above, the improving convergence to CC3 by increasing b.s. size did not improve much in going from double augmented double zeta (d-aug-cc-pVDZ) to quintuple zeta (d-aug-cc-pV5Z) [138]. In a separate benchmark study, Freise *et al.* compared pure (BP86), hybrid (B3LYP), and range separated (CAM-B3LYP) functionals against CC2 on 10 different molecules like push-pull stilbenes. The team found that CAM-B3LYP performed the best and had no significant qualitative difference to CC2 [155]. Furthermore, Laurent and Jacquemin theorize that the success of LRCH in the previously mentioned articles can be attributed to the strong charge-transfer character of the excited-states investigated [142]. CAM-B3LYP has also been compared to CC3 for water, hydrogen fluoride, ethylene, and octatetracene by Knippenberg *et al.* The team demonstrated that CAM-B3LYP deviated from CC3 by only 15%, but unsurprisingly could not model doubly excited-states [156]. In the more recent study by Zalesny *et al.* two hybrids, and 3 range separated xcf were employed against CC3 in quadratic response [141]. Across 76 singly excited-states, BHLYP and CAM-B3LYP performed the best with an MAE of 15.8 and 14.6 a.u., being on average underestimated by -2.9 and -3.9 respectively. However, the spread of errors is quite significant in their respective SDEs of 30.1 and 32.2 a.u. When looking at only Rydberg states, BHLYP and CAM-B3LYP perform similarly with a mean absolute percentage error (MAPE) of 26.2 and 25.2% respectively. And similar to what was found in CC methods above, valence states are considerably more difficult to describe, here the MAPE for all xcf types increase by double to ca. 50-70% [141]. To the best of the author's knowledge, there currently are no systematic study of TD-DFT against higher level methods on Rhodopsin, retinal, or reduced retinal models in terms of 2PA. Therefore the author aims to elucidate the best performing functional for TD-DFT against higher level methods like CC3 which pertains to objective no. 1.

1.4.4: The QM/MM model

Even with the efficiency of DFT, modeling the entirety of the Rhodopsin protein would still not be feasible, nor necessary. Besides from modulating the level of theory used in calculating these electronic properties, another approach to tackling the computational efficiency of modeling larger biological systems is the use of multi-scale systems like the hybrid quantum mechanical/molecular mechanics (QM/MM) method. This method will be further explored in the methodology section. In this section, the author introduces the concept of QM/MM and explore the current state of knowledge

of this model on Rhodopsin 2PA signals. The QM/MM scheme first introduced by Martin Karplus, Arieh Warshel, and Michael Levitt, and allowed theoreticians to model large scale systems by partitioning the investigated system into multiple parts, thus awarding the trio with the 2013 Nobel prize in chemistry [157]. QM/MM relies on partitioning the investigated system into two main parts; a QM subsystem where the phenomenon of interest would occur which requires higher level quantum mechanical description, in the case of Rhodopsin this would be the photoisomerization of retinal. The QM region may be expanded from simply the retinal model to a larger area including surrounding amino acid side chains and crystallographic waters. The QM is embedded within the MM system, consisting of the rest of the system which can include the rest of the protein chain, waters, and ions. The MM system electrostatics is represented in a simpler scheme, commonly, point charges in the place of atomic centers using classical Newtonian physics and a force field (see methodology). The interaction between the QM and MM subsystems must also be carefully represented. These simple point charges can be expanded to higher order multipoles like dipoles, quadrupoles, etc. in which dipole polarizabilities may be derived in order to induce the QM electron density with the MM charges, this effectively creates the Electrostatic Embedding (EE) scheme. A further expansion to this scheme would be a mutually polarizing scheme where the MM subsystem may also be induced by the QM, this is known as the Polarizable Embedding (PE) scheme.

The results herein are focused on absorption rather than emission or fluorescence, yet discrepancies like that of 2PEF has also been observed in theoretical results of wt-Bovine Rhodopsins, similar to those presented in the previous section (1.4.2). This can of course be understood as the difference in standardization of the methodology in investigating 2PA properties of Rhodopsin. In fact when looking at two separate publications by Olivucci *et al.* 2018 [97] and Palczewska *et al.* 2014 [16], one can also see a large discrepancy between these reported values. Olivucci *et al.* demonstrated QM/MM where the QM portion is modeled using multistate multiconfiguration XMCQDPT2 with orbital configurations produced by CASSCF (XMCQDPT2//CASSCF), an MM modeled by the AMBER force field, and a mutually polarizable embedding (PE) scheme to calculate σ^{2PA} cross-sections via the sum-over-states approach. On the other hand, Palczewska *et al.* modeled their Rhodopsin's QM region using DFT with the CAM-B3LYP xcf, while the MM region was modeled by the general Amber force field (GAFF) with an EE scheme where 2PA was calculated via the quadratic response approach. Furthermore, the two studies also have differing QM region sizes, where Olivucci *et al.*'s QM had a total of 54 atoms, while Palczewska's consists of approx. 62 atoms. The differences in these parameters results in the two lowest excited-states between the two publications to have drastically differing results: S_1 computed with a 472 GM at ca. 950 nm peak against 1 GM at ca. 1100 nm, and S_2 with 231 GM at ca 678 nm against 2 GM at ca. 1020 nm, for Olivucci and

Palczewska's respectively [16, 97]. This is not unexpected as seen from the examples above of how different levels of theory (QM), electrostatic interactions, and geometry (in terms of wild-type and mutational geometries) can affect spectral properties. In terms of QM/MM however, another aspect that one must consider is also the QM region size. Rhodopsin one-photon properties, including excitation energies, has been investigated in terms of QM region size in a handful of publications [89, 157, 158]. Firstly, Snedkov's computational screening of Channelrhodopsin (CR) mutants C1C2 have had their BLA, excitation energies, and relative δ^{2PA} strengths compared across 5 total combined QM size ranges in terms of geometry optimization and spectral calculations [89]. The team used PE with the RI-CC2 and TD-DFT/CAM-B3LYP method to calculate excitation energies and 2PA strengths (TD-DFT only). Geometries of C1C2 were optimized at the B3LYP in the QM region, while the MM used the Optimized Potential for Liquid Simulations (OPLS) 2005 force field. With the same QM size in spectral properties, going from no QM geometry optimization (that is, simple MD preparation of assigning missing structures, protonation states and optimizing hydrogen positions from a crystal structure, see further in methodology), to QM/MM geometry optimization where the QM size is the lone retinal saturated with a methyl group, BLA decreases by 0.06 Å, excitation energy increases by 0.12 eV, while relative δ^{2PA} drops by 50%. The increase of QM geometry optimization to the largest size that includes the a.a. Lys296, Asp292, Glu162, Trp163, and a singular crystallographic water did not change its properties by much, BLA increased by a mere 0.01 Å, excitation energies by only 0.03 eV, and increased relative 2PA by 10%. When considering the full QM size in geometries and varying the spectral QM size from lone retinal to also including the covalently bonded Lys296, excitation energies increase by 0.06 eV while relative 2PA strengths decrease by a fair 20%. Expanding the spectral QM size to all 4 a.a. and water molecule also doesn't affect excitation energies by much, only dampening it by 0.02 eV, while relative 2PA is increased back by 10%. Fujimoto and co also investigated the effect of increasing QM up to a range of 3 sizes for Rhodopsin, bacteriorhodopsin (BR), and sensory Rhodopsin (SRII) in the EE scheme using SAC-CI [157]. The team notes an increase of ca. 0.40 eV across all 3 Rhodopsin types, when going from having just the bare retinal chromophore within the QM region to the mid-size one which includes covalently bonded lysine, singular counterion a.a. and water molecule. Furthermore, when expanding the QM region even further to include an additional two a.a. and 3 water molecules in BR and SRII, a difference of 0.00 to 0.01 eV was observed, making the mid-size QM region sufficient in terms of ΔE [157]. In the same study oscillator strengths and transition dipole moments (μ_{01}) were also disentangled, the team notes the decrease in μ_{01} but an increase in f when going from bare retinal in the QM region to the mid-size QM. They concluded that the increase in f is likely due to increase in ΔE . Change in dipole moments were also calculated for the 3 Rhodopsins, but only at the pure QM (no MM), and bare retinal QM size, displaying a large negative $\Delta\mu$, increasing in negativity when

bound within the opsin, similar to the findings of Sugihara [106] described in 1.3.3. Moreover, mid-size QM region also showed great agreement with experimental values and demonstrated a max deviation of 0.05 eV across all 3 Rhodopsins [157]. Valsson and co-workers had tested QM sizes in Rhodopsin using the semiempirical ZINDO approach for excitation energies, which are later refined at the TD-DFT/CAM-B3LYP level [158]. ZINDO allows calculation of excitation energies at higher efficiencies but with the trade-off being more approximate, although ZINDO has been reasonably benchmarked using retinal [159] and Rhodopsin intermediates [160], it also exhibits a red-shift when compared to higher level methods [161]. In any case, Valsson and co-workers used ZINDO with the EE scheme to compute average excitation energies from 1500-2000 frames taken from the last 10 ps of their QM/MM MD trajectory in up to 13 different cluster sizes, starting from the bare chromophore with an additional a.a. added for every step size, till the largest QM size being all atoms within 5 Å of retinal (720 atoms total). Going from the averaged bare chromophore, M1 cluster, to the average of one additional a.a. (Glu113), M2 cluster, ΔE blue-shifted from 2.18 eV to 2.45 eV, similarly to Fujimoto above. Expanding the QM to include two additional a.a. (M3 and M4) gradually increases ΔE to ca. 2.53, further expansion actually begins to red-shift the energies before finally arriving at a stable convergence at M10. In fact, going from M10 (250 atoms, 9 additional a.a. total and 7 waters) to all atoms within 5 Å of retinal (the largest size), only a small increase from 2.40 eV to 2.41 eV can be observed. Refinement of ZINDO excitation energies by TD-DFT/CAM-B3LYP averaged over 11 individually selected frames revealed that including the EE scheme to the bare retinal chromophore QM size can drastically help converge ΔE towards a more detailed QM system, mitigating the loss of a more accurate QM description. That is, adding EE to the bare retinal QM (M1) increases ΔE from 2.39 to 2.65 eV, while adding EE to the M2 size lowers ΔE from the pure QM of 2.70 eV to 2.66 eV. When further testing CAM-B3LYP in larger cluster sizes, the team concluded that depending on the size chosen for EE scheme, ΔE may shift as much as 0.2-0.3 eV, and recommends a minimum of M10 size for convergence to experimental value [158]. Lastly, the team notes that in order to reliably capture the dynamic, thermal, and environmental variations in the protein-chromophore system, one should avoid using a singular frame (taken from MD that is) to calculate excited-state properties. Hence, one could arrive at the correct answer (convergence to experimental absorbance) due to cancellation of errors, as was also corroborated by others [102, 162].

In terms of 2PA signals, to the best of the author's knowledge, no systematic investigation into the QM size in various embedding schemes has been performed for Rhodopsin, looking over at FPs again here could also lead to some insights. Grabarek and Andruniow had performed QM size investigations into neutral and anionic forms of GFP, as well as citrine FP using the BhandHLYP xcf with the PE, EE, and no embedding (NE, akin to calculations *in vacuo*), as well as additional comparisons to CAM-

B3LYP post publication [162]. The duo investigated increasing the QM size from the bare chromophore to a max radius of 0.5 nm, with an interval of 0.05 nm per step size. The two noted that the brightest 1PA band is qualitatively well described by just the bare chromophore in terms of EE scheme, but does red-shift by up to 0.15 eV when increased to max QM size. On the other hand, f intensities in EE scheme converges to a steady value much slower with increased QM size, even after including ca. 250 atoms (0.30-0.35 nm radius) f still slowly decreases with increasing QM size. 2PA on the other hand at just the bare chromophore QM size had neither quantitatively nor qualitatively correct spectra for excitations above 4.0 eV. In short, to converge to a steady value and have the same qualitative description in both 1/2PA spectra as with larger systems above 300 atoms, a QM of at least ca. 220 atoms is required in EE. Finally, the PE scheme is able to reduce the QM size to ca. 80-100 atoms and maintain a converge 1/2PA spectra. CAM-B3LYP was compared to BhandHLYP results post publication, the duo in their supporting information (SI) noted that CAM-B3LYP generally red-shifts BhandHLYP functional by 0.1-0.2 eV at absorption peaks. Yet, the two functionals retain the same quantitative trends in ΔE and 1/2PA values when increasing QM size, but differences are more visible in higher excited-states [162]. Looking back to the results of Olivucci [97] and Palczewska [16], one raises the question of how large Rhodopsin's QM size should be and which embedding scheme should be employed in order to obtain converged 2PA properties. While Olivucci et al. employed a higher-level QM method, the convergence of the QM region itself was not systematically investigated. This therefore leads into objective no. 4.

2. Objectives

Throughout the Introduction chapter, several of the primary objectives of this work were introduced. Broadly, these objectives can be associated with either methodological aspects, focusing on establishing reliable computational protocols, or practical aspects concerning the physical mechanisms governing spectral tuning and nonlinear optical response in retinal and Rhodopsin systems. The main objectives of this work are summarized as follows:

1: Evaluate the performance of increasing the electron correlation by increasing the coupled cluster hierarchy from CC2, CCSD, to CC3 on ΔE , f , μ , and 2PA signal for reduced sized protonated Schiff base retinal models *in vacuo*. Furthermore, analyze various xcf in TD-DFT, with emphasis on CAM-B3LYP and BHandHLYP, against CC3 to benchmark their performance for use in larger scale systems. In addition, two types of basis sets were also evaluated with this set of methods, namely Pople's 6-31+G* and Dunning's aug-cc-pVDZ.

2: Investigate a systematic way to tune 2PA signal in Rhodopsin's 11-cis rPSB based on modulating the methyl groups pattern along its polyene chain which produces a total of 7 models calculated at the RI-CC2 level *in vacuo*. Furthermore, in light of the results from objective no. 1, two additional xcf: M11 and MN15, for a total of 4 xcf are also benchmarked against RI-CC2 to determine both its quantitative accuracy and qualitative similarities in trends. Other spectral properties like ΔE , f , and μ which may be in/directly correlated to δ^{2PA} signal are also analyzed based on the effects from structural modulations which leads into objective no. 3.

3: Disentangle the correlation between different spectral properties like ΔE , permanent electric dipole moments, and transition dipole moments and how they affect the δ^{2PA} and σ^{2PA} signals based on quadratic response and the two-state model.

4: Determine the best system setup in investigating TD-DFT 2PA spectra in the Rhodopsin QM/MM system. The objective here is to determine in a systematic way the most efficient yet accurate QM subsystem size by increasing its size and composition in a stepwise manner and comparing their corresponding 2PA spectra with those of reference methods and experimental data.

3. Methodology

3.1: Schrödinger's Time-Independent Equation and the Hartree-Fock method

Quantum mechanics provides chemists, and physicists alike, the theoretical basis for understanding molecular interactions at an atomic scale. Computational chemists are able to apply this basis to complex molecular systems in order to predict desired chemically relevant properties and phenomenon. The behavior of any molecular system for state i can be fully described by the time-independent Schrödinger equation:

$$\hat{H}\psi_{(i)} = E_{\{i\}}\psi_{(i)} \quad (3.1)$$

Where $\hat{H}$ is the Hamiltonian operator which represents both kinetic and potential energy of the system, E is the total energy of state i , and Ψ is the total wavefunction for both electronic and nuclear coordinates. For the purposes of this methodology section, the state focused will be the ground-state

($i=0$), unless otherwise stated. In terms of this thesis, the foremost relevant molecular properties are related to its electronic structure, that is the state and motion of the electrons that are involved in Rhodopsin/retinal's excitation. Therefore, it is commonly assumed that the motion of electrons and nuclei can be decoupled through the Born-Oppenheimer (BO) approximation due to their large mass disparity, allowing one to solve the electronic Schrödinger equation independently of the nuclear motion. On the other hand, it is commonly known that there is no exact analytical solution for multi-electron systems to the Schrödinger equation. Therefore, to apply the Schrödinger equation to realistic applications, numerical and approximation methods are used instead. One of the earliest approximations is in the form of the Hartree-Fock (HF) method. The many electron HF wavefunction ($\Psi_{(0),HF}$) is approximated by a set of single electron functions known as a spinorbitals, $\Phi_i(j)$, arranged as a Slater determinant and represents a single configuration of the electrons. Each spinorbital consists of a spatial component, $\phi_i(j)$, and a spin component, $\alpha/\beta(j)$, for the j -th electron. This arrangement automatically satisfies Pauli's exclusion principle hence when two rows in the determinant are interchanged, the determinant's sign also changes.

$$\Psi_{HF}^{(0)} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \Phi_1(1) & \Phi_2(1) & \cdots & \Phi_N(1) \\ \Phi_1(2) & \Phi_2(2) & \cdots & \Phi_N(2) \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_1(N) & \Phi_2(N) & \cdots & \Phi_N(N) \end{vmatrix} \quad (3.2)$$

This original wavefunction constructed from the single Slater determinant is often times referred to as the trial wavefunction (Ψ_{trial}) in the HF method. Plugging in the Slater determinants back into the time-independent Schrödinger equation one would therefore integrate the expectation value of $\hat{H}$ to solve for the total energy with respect to the trial wavefunction. The energy produced from Ψ_{trial} will always be equal to or greater than the true ground-state E_0 under the variational principle.

$$E(\Psi_{\text{trial}}) = \langle \Psi_{\text{trial}} | \hat{H} | \Psi_{\text{trial}} \rangle \geq E_{\text{actual}} \quad (3.3)$$

To ensure orthonormality of orbitals and therefore uphold Pauli's exclusion principle, Lagrangian multipliers are employed which leads to the construction of integro-differential Fock equations which are laborious to perform hence they are not solvable analytically for systems larger than a single hydrogen. Instead, the Roothan-Hall approach transforms unknown molecular orbitals $\Phi_i(j)$ into known atomic orbitals χ_μ by using the Linear Combination of Atomic Orbitals into Molecular Orbitals (LCAO-MO) producing basis functions.

$$\phi_i = \sum_k c_{k,i} \chi_k \quad (3.4)$$

In modern computational methods, multiple basis functions can be linearly combined to produce a basis set for more accurate orbital descriptions. The basis functions/sets can then produce the electron density matrix which is plugged back into the Fock equation to finally produce the Roothan-Hall equations and compute the new MOs. These new MOs can then produce a new electron density matrix and evaluate its HF energy. This cycle is repeated in an iterative manner until the orbitals are optimized, and subsequent change in HF energies become stable or reached a certain threshold. This process is commonly referred to as the Self-Consistent Field (SCF) approach.

However, the HF approach is incomplete. Its formulation lacks coulombic/dynamic correlations, due to its one-electron approximation, the electron-electron repulsions are not accounted for directly but rather in a mean field. Secondly, since single Slater determinants only represent one specific occupation of spin orbitals (configuration), this may be insufficient in describing nearly/degenerate states, accounting for static correlations. Multiconfigurational and multireference wavefunctions are able to fill in both static and dynamic correlations, respectively. These are also known as post HF methods, for example, the Configuration Interaction (CI) method which can be expanded to include both static and dynamic correlation when using the ‘full’ number of configurations of the determinants. The CI method relies on a linear combination of determinants including the ground and excited (unoccupied virtual) states, where occupied orbitals are replaced by virtual orbitals (effectively promoting electrons to the virtual orbitals) to produce new configurations. These new configurations can include the simultaneous replacement of two (double), three (triple), and up to n -excited orbitals, giving rise to the CI hierarchy. When considering the full-CI excitations up to n -electrons (all electrons within the system) are included, while CISD only includes singles and doubles excitations. Furthermore, the increase in basis set size will also increase the number of orbitals considered and therefore also the determinants considered, where using a complete basis set would make the method the ‘complete-CI’. On the other hand full-CI is computationally unfeasible for any system larger than a few electrons, only lower level methods like CISD is possible for realistic systems. Variations of the SCF methods like the multiconfigurational SCF (MCSCF) or CASSCF have also been developed to be more computationally feasible by only considering a selected number of determinants (MCSCF) or perform full-CI but only within the selected active spatial orbitals (CASSCF). The reduction in complexity allows higher efficiency but with a trade-off of being only able to capture static correlation.

3.2: Perturbation Theory

In order to capture the remaining dynamic correlation, in an approximate way without the use of full-CI, the Many Body Perturbation Theory (MBPT) becomes an enticing choice. One of the more popular flavors is the Møller-Plesset perturbation theory (MP), which relies on adding correlation energies via a perturbation. In the case of MP, its perturbed $\hat{H}$ operator is composed of the non-perturbed $\hat{H}_{\text{(HF)}}$ plus $\lambda\hat{V}$, where $\hat{V}$ represents the perturbation operator and a scalar parameter λ (usually set to 1, simply to denote MP expansion order). The perturbation operator is defined as the difference between the full electron-electron coulomb interaction and HF potential, thus it represents exactly the missing electron correlation in HF. Using Taylor expansion the wavefunction and energy may be expressed in higher orders, starting from the zeroth order term being the unperturbed reference HF wavefunction and energy ($E_{(0)}\Psi_{(0)}$).

$$\Psi(\lambda) = \Psi_{(0)} + \lambda\Psi_{(1)} + \lambda^2\Psi_{(2)} + \dots \quad (3.5)$$

$$E = E_{(0)} + \lambda E_{(1)} + \lambda^2 E_{(2)} + \dots \quad (3.6)$$

The ability for MP to easily expand to higher order terms allows for the development of the MP hierarchy, by including higher order terms and therefore more subtle electron correlation effects, the results should converge towards the true total energy. As usual, this comes at a dramatic increase in computational cost however, the increase from MP2 to MP3 and MP4 may not actually guarantee increased accuracy [163]. Furthermore, although MP2 incorporates doubly excited determinants, it does so perturbatively by adding energy corrections instead of considering a linear combination of determinants within the wavefunction. This means MP is still only single reference and single configuration, although it can capture as much as 80-90% of dynamic correlations [163], extensions to the perturbation theory to encompass these missing correlations have been developed as the CASPT (Complete Active Space Perturbation Theory) methods. Once a multiconfigurational wavefunction has been constructed, such as those performed in MCSCF and CASSCF, a similar process as MP may be employed. Perturbation corrections to the energy, for example those involved in double excitations in CASPT2 (2nd order Perturbation Theory), are added based off the previously constructed multiconfigurational wavefunction. This results in the CASSCF/CASPT2 combo great in capturing both static and dynamic correlations in an approximate way.

3.3: Coupled Cluster

The Coupled Cluster (CC) method is another approach at tackling dynamic correlations. CC is able to treat correlations between pairs of electrons, and even up to large clusters (3 or more) of electrons, similarly to CI. However, unlike CI, CC relies on an exponential ansatz, as opposed to a linear combination one. The CC wavefunction starts similarly to previous methods, where $\Psi_{(0)}$ is taken from Ψ_{HF} , specifically here the HF spinorbitals (Φ_0) is used as the reference determinant.

$$|\Psi_{\text{CC}}\rangle = e^T |\Phi_0\rangle \quad T = T_1 + T_2 + T_3 + \dots + T_N \quad (3.7)$$

The cluster operator, T , can be expanded to include single (T_1), double (T_2), and up to N^{th} simultaneous excitations to virtual orbitals for an N -electron system. Of course if the full N^{th} CC excitations were used this would be equivalent to the full-CI. Plugging in the cluster operator up to singles and doubles (CCSD) into the exponential operator and expanding it via a Taylor series yields;

$$e^T |\Phi_0\rangle = \left(1 + T + \frac{1}{2!}T^2 + \frac{1}{3!}T^3 + \dots\right) |\Phi_0\rangle \quad \text{where } T = T_1 + T_2 \text{ for CCSD} \quad (3.8)$$

As seen from the equation above, it is clear under polynomial expansion that second-order and higher terms would experience squared (or higher order exponents) terms like T_1^2 and T_2^2 which represents disconnected (implicit) double and quadruple excitations constructed from two single and two double excitations respectively. At the same time cross-terms like T_1T_2 and T_2T_1 would also produce disconnected triple excitations constructed from the product of a single and a double excitation. Implicit excitations may be included to infinity but realistically only up till the N^{th} rank in an N -electron system, higher excitations beyond the number of total electrons in a system would simply vanish to zero when acting on the reference determinant Φ_0 . Regardless, this lends itself to CC's inherent size extensivity, recovering additional correlation energies that are not available in equivalent methods like CISD which allows accurate calculations even for much larger sized systems. Additionally, as noted above in eqn. 3.8, CC may be expanded to include explicit higher rank excitations like CCSDT (for explicit triples) up to the N^{th} order but realistic calculations truncate this to CCSD(T) to save on computational cost. To this end, approximations which act as intermediate levels, for example between explicit singles (CCS) and doubles (CCSD), like CC2 have also been developed. CC2 relies on computing singles (T_1) amplitudes exactly in an iterative manner while doubles (T_2) are solved approximately as a perturbation, similar to MP2. CC3 extends CCSD through an approximate iterative treatment of connected triple excitations, providing an intermediate level between CCSD and full CCSDT at a lower computational cost.

3.4: Density Functional Theory (DFT)

An alternative quantum mechanical method developed with high efficiency in mind was first presented by Hohenberg and Kohn in their 1964 theorem [164]. The first part of the theorem states that the ground-state electron density $\rho_0(\mathbf{r})$ uniquely determines the external potential acting on an N-electron system. Therefore, all ground-state properties can be expressed as a functional of $\rho_0(\mathbf{r})$. This reduces $\Psi_{(0)\text{HF}}$ from a function of $4N$ variables, corresponding to 3 spatial coordinates and one spin coordinate for each of the N electrons, to simply a function of 3 spatial coordinates ($\mathbf{r}$).

$$\rho_0(\mathbf{r}) = N \int d\sigma_1 dx_2 \cdots dx_N |\Psi(x_1, x_2, \dots, x_N)|^2 \quad (3.9)$$

The function above describes the probability density of finding any one indistinguishable electron at the 3 fixed spatial coordinates ($\mathbf{r}$), that is after having integrated over its own spin (σ_1), as well as all other electrons' spin and spatial coordinates (collectively x_i). Once the probability density is obtained, the second part of the theorem comes into play, it states that there exists the variational principle for the energy as a functional of $\rho_0(\mathbf{r})$, similarly to that of Ψ_{HF} . That is, minimizing $E[\rho(\mathbf{r})]$ with respect to all physically valid electron densities will yield the true ground-state energy, where $\rho_0(\mathbf{r})$ would be the electron density of a fully interacting, correlated system. Despite the exact nature of DFT in principle, in practice the true $E[\rho_0(\mathbf{r})]$ is not known, as the explicit functional forms of the kinetic and electron-electron interaction energies in terms of $\rho(\mathbf{r})$ are unknown. To this end, Kohn and Sham (KS) proposed replacing the fully interacting N-electron system with a fictitious non-interacting system. The KS-DFT energy is expressed as the sum of non-interacting kinetic energy functional ($T_s[\rho]$), the external potential ($V_{\text{ext}}(\mathbf{r})$), which treats electrons as a stationary cloud, classical coulomb repulsion functional ($E_H[\rho]$), describing the classical electron-electron interaction of the electron density, and finally the exchange correlation energy functional $E_{\text{xc}}[\rho]$, which treats all many-body effects outside of the non-interacting system. In fact, it is the E_{xc} that makes DFT so efficient. By isolating the most complex component of the total energy, this small but significant contribution representing the quantum effects not captured by the previous terms can be treated approximately. $E_{\text{xc}}[\rho]$ specifically is able to correct for three main deficiencies, i) the difference between the true Kohn-Sham kinetic energies, ii) exchange energy arising from the antisymmetry of the wavefunction, and iii) dynamic electron-electron correlation energy.

$$E[\rho] = T_s[\rho] + \int \rho(r) V_{ext}(r) dr + E_H[\rho] + E_{xc}[\rho] \quad (3.10)$$

In practice, this energy functional is only evaluated at the end of a KS-DFT loop, once the electron density has converged (that is, minimal change from e.g. p_{0r} and p_{1r}). The loop, similar to that of SCF, begins with an initial electron density (p_{0r}) which is used to construct the effective potential (V_{eff}) with only 3 constituents, $V_{ext}(r)$, $E_H[p]$, and $E_{xc}[p]$. V_{eff} then enters into the KS equations to produce new orbitals, which in turn can produce the new densities (p_{1r}). The loop is repeated till convergence has been met.

$$V_{eff}^{(0)}(r) = V_{ext}(r) + V_H[\rho_0(r)] + V_{xc}[\rho_0(r)] \quad (3.11)$$

As seen in the equation above, the quality of the effective potential V_{eff} , and consequently the convergence of the electron density $p_1(r)$, depend on the form of the exchange-correlation potential V_{xc} . Since $V_{xc}[p]$ is defined as the functional derivative of the exchange-correlation energy $E_{xc}[p]$, the accuracy of the resulting potential ($V_{xc}[p]$) depends directly on how well the underlying functional ($E_{xc}[p]$) is constructed and parameterized. These functionals are usually parameterized with reference data obtained from higher level methods with more accurate descriptions of the wavefunction, like CI, CC, as well as quantum monte carlo methods, and even with experimental measurements. This gives rise to DFT's semi-empirical nature.

The development and classification of these functionals to fit specific physical and mathematical demands is organized according to the metaphorical ladder presented by Perdew and Schmidt. At the lowest level we have the LDA method, it approximates the exchange-correlation energy in DFT by assuming that each point in the electron density, $p(r)$, behaves like part of a uniform electron gas. Although it performs well in extended periodic systems, due to its nature it lacks an inhomogeneous description of electrons at each local point leading to only an approximate treatment of inhomogeneity in the global description of the system. This makes the description of many properties involving non-local correlation and exchange, such as those found in non-covalent long-range interactions like dispersion and charge transfer, poorly represented in LDA. To this end, the next rung on the ladder, GGA, incorporates density gradients ($\nabla p(r)$) to model the missing inhomogeneity of the electron density at each local point. This helps improve things like total potential energies and molecular geometries, but still lacks long-range descriptions. Notable variants includes the PBE and BLYP functionals. The next step, meta-GGA functionals, includes an additional kinetic energy density term that helps improve the description of orbital localizations and distinguish electron shell structures.

Regardless, functionals on the lower tier of this metaphorical ladder relies on local or semi-local approximations of electron treatment which does not consider non-local exchange correlation effects. By including exact HF exchange energy into these functionals, true non-local static exchange character can be introduced, creating the next rung on the ladder, hybrid functionals. Hybrids greatly improve exchange description corrections which may affect properties significant properties like excitation energies and electron delocalization. However, depending on the exact variant, only a certain percentage of exact exchange is included, for example the popular B3LYP consists of 20% HF exchange, and BHandHLYP includes 50%. Meaning, the aforementioned functionals may still underestimate non-local behavior, particularly at large separations of donor and acceptor orbitals which may result in incomplete descriptions of charge-transfer states and Rydberg states. The next evolution on the same rung of the ladder is of course the range-separated variants, LRCH. The upgrade here is to simply divide short-range exchange to that of DFT and long-range exchange to being 100% HF exact exchange. This keeps computational efficiency high where possible while still modelling the long-range exchange potential decay as accurately as possible with HF. Notable LRCH variants include the CAM-B3LYP functional.

3.5.1: Multiscale Modeling with QM/MM

As noted in 1.4.4, the need for multiscale modeling is rooted in increasing computational efficiency due to the necessity in using QM-methods to describe changes in the electronic structure involved in properties like excitation energies, charge-transfer, multi-photon absorption, and dipole moments. Depending on the exact QM-method, computational cost can rise exponentially, for example at N^{4-7} in HF up to CC3. DFT has been demonstrated to compute ground-state properties for a system up to few thousand atoms [165]. However, this only constitutes a small to medium size protein, much less the entire Rhodopsin protein in an aqueous environment, and not to mention more complexed calculations like excited-state properties or geometry optimization. As noted in the Introduction, calculations or measurements done in specific environments can have a drastic effect on absorption and geometric properties. Therefore, to tackle the issue of protein in solution continuum solvation models like the polarizable variant (PCM) has been developed. This allows one to avoid modelling every single atom in the solvent, which can reach up to 10s or 100s of thousands of atoms, to a single continuous polarizable dielectric medium. This highly increases the computational efficiency but would still require one to model the entire protein in QM, which for Rhodopsin is unlikely. The popular solution to this is the hybrid QM/MM method, introduced in 1.4.4. As noted above, QM/MM is able to benefit from advanced QM methods while maintaining relatively low computational cost by partitioning the entire system into two parts. A QM region governed by advanced QM methods

which likely will contain the phenomenon of interest, and an MM region governed by semi-empirical force fields using Newtonian equations of motion.

3.5.2: Partitioning the system

As with any partitioning, actual regions must be physically separated from each other, this of course is not physically accurate when considering a protein structure of covalently bonded segments. To this end, careful and specialized treatments of this border region needs to be accounted for. To avoid any artefacts, it is generally advised when selecting the partition region to cut weakly polarized C-C covalent bonds, a particularly common area being the C α . Once cut, to avoid any unphysical dangling bonds, these cut areas are saturated with a Linker atom. The atoms that are directly bonded to the Linker atom at the QM region are generally denoted as Q1 atoms, while the one directly bonded at the MM region are denoted as M1 atoms. Directly attached to both the Q1 and M1 atoms are the (up to 3 possible) Q2 and M2 atoms, respectively. These Q1 and Q2 and M1 and M2 atoms make up the frontier atoms. Using the link atom scheme, the Linker atom itself would be placed between the Q1 and M1 atoms, with a distance of approximately 0.5 Å from the M1 frontier atom.

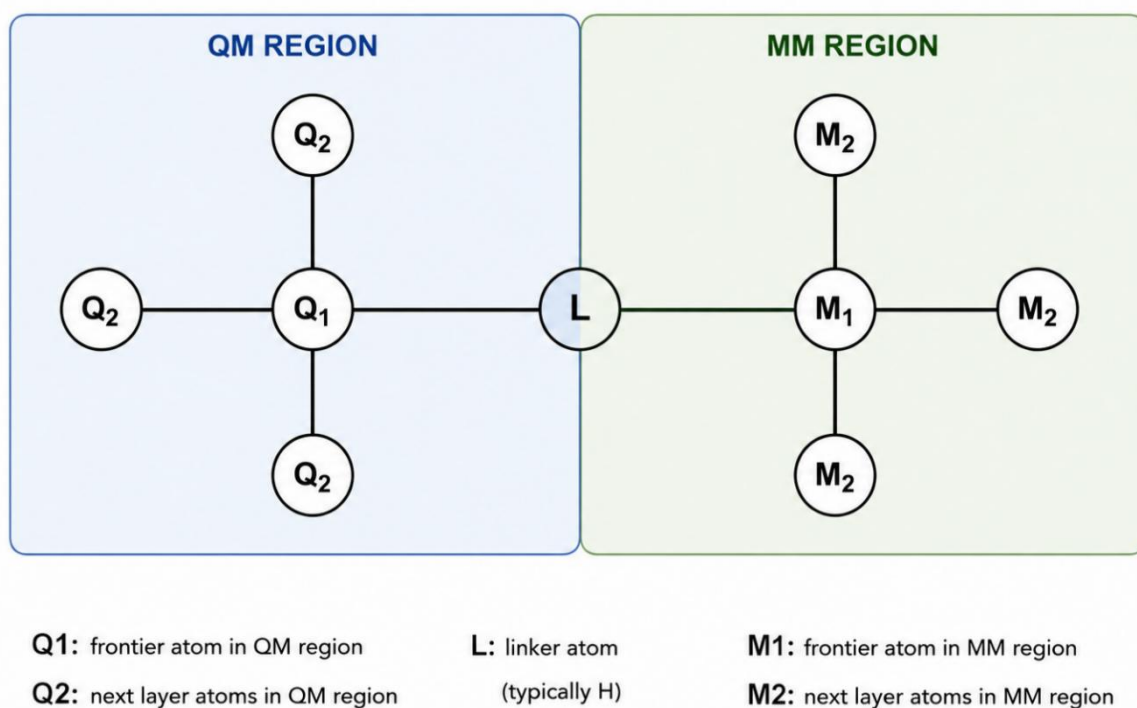


Figure 3. Schematic representation of QM/MM partitioning showing frontier atom definitions and linker atom (L) placement at the partition boundary.

The Linker atom here are often represented as hydrogen and is directly incorporated into the QM region. This may pose as an issue when considering the electrostatic interactions between the two regions which may easily overpolarize the Linker atom due to the proximity of the M1 and M2 atoms. On the other hand, a key strength in the QM/MM method is its ability to choose the level of complexity on how the regions may influence each other, allowing one to dampen or enhance electrostatic polarization effects as needed.

3.5.3: QM/MM is an interacting system

When considering the total energy of a QM/MM system, the coupling between the two regions adds an additional term for the interaction energy.

$$E_{total} = E_{QM} + E_{MM} + E_{int} \quad (3.12)$$

The QM/MM interaction energy generally consists of electrostatic and van der Waals (vdW) contributions, and may additionally include bonded terms when covalent bonds cross the QM/MM boundary. The electrostatic term is therefore of great interest in terms of polarizabilities due to it being one of the strongest interactions, long-range nature, and have a strong impact on the electronic structure itself. As mentioned in 1.4.4., 3 (or 4, if considering ‘no embedding’) types of embedding schemes have been popularized for describing this coupling, including the mechanical (ME), electrostatic (EE), and polarizable (PE) embedding schemes.

ME is considered to be the simplest QM/MM embedding scheme; it models electrostatic point charges in place of atomic positions in both the QM and MM regions. As a result, electrostatic interactions between the QM and MM regions are computed entirely within the MM force field framework, rather than being derived from the QM Hamiltonian. Furthermore, careful consideration is required when assigning point charges to QM atoms to ensure compatibility with the MM force field. This poses an inherent limitation: describing the QM electron density using fixed point charges is insufficient in capturing its dynamic nature, particularly during processes like electronic excitation or charge transfer. Additionally, since only point charges are considered, with no higher-order terms such as those in a multipolar expansion, mutual polarization between the QM and MM regions cannot be captured. Thus, ME effectively couples QM and MM regions only at a mechanical level, namely through bonded interactions and geometry. The next step in improving the interaction between the regions is to incorporate the electrostatic potential of the MM region into the QM hamiltonian. By computing each region using their respective methodologies, the MM electrostatic potential can be

explicitly included as an external field acting on the QM electron density. This allows the QM electronic wavefunction to be solved in the presence of MM charges, enabling the QM region to be polarized by the static electrostatics of the MM environment. Despite these improvements, mixing parameterized point charges in the MM region with true electron densities in the QM region may not be ideal in situations involving strong polarization, charge transfer, or cases where the MM charge model poorly represents the underlying physical behavior. Nevertheless, this implementation is among the most widely used embedding scheme due to its ease of access. The extension of the EE scheme results in the PE scheme where mutual polarization can occur. PE introduces atomic polarizability tensors to the MM region, which allow induced dipole moments to form at each MM site in response to the electric field generated by the QM electron density. Importantly, these induced dipoles are not fixed quantities. The induced dipoles depend on the QM electron density, while the QM electron density itself is influenced by the induced dipoles. Therefore, the induced dipoles and QM wavefunction must be updated self-consistently throughout the QM optimization procedure. The induced dipole at MM site i is given as:

$$\mu_i^{ind} = \alpha_i \cdot F_i^{tot} \quad (3.13)$$

where α_i is the polarizability tensor and F_i^{tot} is the total electric field at site i . The total electrostatic energy in PE can thus be written as the sum of the electrostatic interaction of QM and MM permanent multipoles (equivalent to the electrostatic energy of interaction in EE) and the energy due to induced dipoles (E_{ind}) in the MM responding to the QM. The total E_{ind} is thus given as:

$$E_{ind}^{total} = -\frac{1}{2} \sum_i \mu_i^{ind} \cdot F_i^{tot} \quad (3.14)$$

Where F_i^{tot} is the total electric field at site i , composed of contributions from the QM nuclei, QM electron density, and static multipoles located in the MM region. Induced dipoles are computed iteratively and self-consistently in each QM wavefunction optimization cycle, leading to a better polarization description at the cost of higher computational cost. At the same time, due to this dynamic mutual adaptation between the QM electron density and induced MM dipoles, changes to the electron density during electronic excitation may also be captured which improves the description of excitation and spectral properties.

For practical reasons, the remaining vdW forces are typically treated at the MM level where reliable force field parameters are readily available, adapted from similar existing atom types, or newly

parameterized using quantum mechanical calculations or empirical data. In most cases, vdW forces are modeled with a Lennard-Jones potential, which can introduce artefacts known as the electron spill-out (ESO) problem. This artefact occurs due to the electron density of the QM region not experiencing repulsive exchange (as well as dispersion) effects from the MM region. This allows the electron density to extend unphysically into the surrounding environment and can result in artificial low-lying excited-states due to overstabilization of excited-state electron density [166–168]. ESO is especially pronounced in the PE scheme and when using diffuse basis functions, and while removing these diffuse basis functions may mitigate this issue, more reliable techniques based on accounting for the exchange interaction at the QM level have been achieved [166–168]. A highly efficient and readily available method proposed by Hättig *et al.* involves using the all-electron Effective Core Potential (ECP) in place of MM sites alongside other typical polarizable force field components like multipoles and polarizabilities [166]. The ECPs create a Pauli repulsion barrier that prevents ESO. Although ESO may also lead to overpolarization of the QM region due to the absence of short-range repulsive interactions, the inclusion of the linker atom within the QM region, positioned near strong electrostatic fields from the MM, is another frequent source of additional overpolarization. To this end, it is common practice to set multipoles and polarizabilities to zero within a certain range from the QM atoms. In order to maintain a consistent integer total charge, the removed point charges, as well as higher-order multipoles and polarizabilities, in the MM are redistributed to the nearby sites.

3.5.4: Modeling of 1PA and TPA signals

Spectroscopic processes such as light absorption inherently involve changes in the electronic structure induced by electromagnetic fields over an extremely short time scales. Upon photoexcitation, the system evolves from the ground-state Ψ_0 to an excited-state Ψ_j . In spectroscopy, the process is captured as a time-dependent pulse and would therefore also require a time-dependent formalism of quantum mechanics, enter in the time-dependent Schrödinger equation (TDSE):

$$\hat{H}(t)\Psi(r, t) = i\hbar \frac{\partial}{\partial t} \Psi(r, t) \quad (3.15)$$

Where,

$$\hat{H}(t) = \hat{H}_0 + \hat{V}(t) \quad (3.16)$$

Here, $\hat{V}(t)$ is a time-dependent perturbation of an external electromagnetic field like light, acting on the unperturbed Hamiltonian, $\hat{H}_0$. The time-dependent $\Psi(r, t)$ can be expressed as a linear combination of time-independent stationary eigenstates Ψ_i of the time-independent $\hat{H}_0$, which allows one to track the system's evolution through different electronic states;

$$\Psi(r, t) = \sum_i c_i(t) \Psi_i(r) \quad (3.17)$$

Here, $c_i(t)$ represents the time-dependent amplitudes, coefficients of each stationary Ψ_i 's contribution to the overall wavefunction $\Psi(r, t)$ at time t . However, this method is computationally unfeasible since it requires computing all excited-state wavefunction Ψ_j explicitly, along with the ground-state Ψ_0 . To circumvent this direct time-dependent perturbation formalism, a widely adopted framework known as the response theory evaluates how the system's properties change in response to the perturbation $\hat{V}(t)$ without needing to compute every Ψ_i . The response theory is motivated by the Ehrenfest theorem, which provides a framework for describing the time evolution of quantum mechanical expectation values of observables, such as dipole moments or polarizabilities, analogously to classical equations of motion. By expanding an observable property's expectation value (like the induced dipole moment) as a Taylor series in the external electric field strength, this yields linear, quadratic, and cubic responses corresponding to first, second, and third order derivatives, respectively. Typically, it is more efficient, and more directly connected to experimental observables, to compute response functions in the frequency domain rather than in the time domain. Most spectroscopic techniques probe how molecules respond to electromagnetic fields at specific frequencies (or wavelengths), making a frequency-resolved approach both natural and computationally advantageous. In the context of the TDSE, the perturbation $\hat{V}(t)$ is typically expressed as the product of the dipole moment operator and an external electric field:

$$\hat{V}(t) = -\hat{\mu} \cdot E(t) \quad (3.18)$$

$E(t)$ is represented in the form;

$$E(t) = E_0 e^{-i\omega t} \quad (3.19)$$

Where E_0 is the electric field strength at $t=0$, and $e^{-i\omega t}$ represents a complex exponential form of an oscillating field, whose real part corresponds to the cosine wave under Euler's formula.

In practice, this means computing the system's response to an oscillating field at each frequency ω of interest. Response functions are expressed as sums over poles, frequencies corresponding to electronic excitation energies, with residues at these poles carrying the physical information about transition strengths and spectral intensities, which provides a direct bridge between theory and experiment. Each term in the expansion of induced dipole moments corresponds to a distinct physical process characterized by its perturbative order: the residue of the linear response function captures

first-order effects such as absorption spectra and oscillator strengths; quadratic response residues capture second-order phenomena like two-photon absorption (δ^{2PA}); and cubic response residues describe third-order processes such as third harmonic generation. Additionally, double residues of the quadratic response function give access to transition moments between excited-states.

An alternative to the quadratic response formalism is the Sum-Over-States (SOS) approach. Instead of solving response equations directly, the SOS approach constructs the two-photon transition moment as a sum over the complete set of eigenstates of the time-independent Schrödinger equation, including both ground and excited-states. In principle this allows multi-photon absorption processes to be described as sequential virtual one-photon transitions through intermediate excited-states. Therefore, the 2PA process is modeled as the linear combination of these transitions, and for linearly polarized light of and two photons of equal energy with parallel polarization, rotationally averaged two-photon transition moments (δ^{2PA}) is given as;

$$\delta^{2PA} = \frac{1}{15} \sum_{ab} (2S_{ab}S_{ab}^* + S_{aa}S_{bb}^*) \quad (3.20)$$

Here, S_{ab} are cartesian components of the 2PA transition moments;

$$S_{ab} = \sum_{n \neq 0} \left[\frac{\langle S_0 | \hat{\mu}_a | S_n \rangle \langle S_n | \hat{\mu}_b | S_f \rangle}{\omega_n - \omega} \right] + \left[\frac{\langle S_0 | \hat{\mu}_b | S_n \rangle \langle S_n | \hat{\mu}_a | S_f \rangle}{\omega_n - \omega} \right] \quad (3.21)$$

Here, equation 3.21 sums over all the excited intermediate (virtual, S_n) states that lead up to the final excited-state of interest (S_f , where S_0 is the ground-state), $\hat{\mu}_a$ and $\hat{\mu}_b$ are components of the dipole moment operator along cartesian directions (a,b = x, y, and z). ω_n is the excitation energy and ω is the photon energy, half of the energy required to reach the final state. This formulation is advantageous due to only using one-photon transition moments, qualities readily available from standard excited-state quantum calculations, as opposed to computing full two-photon transition tensors via quadratic response which contains SOS implicitly. However, because the SOS requires the full sum of intermediate states, the computation cost would quickly rise with the number of excited-states involved and would therefore require a truncation of the expansion. As a result, the accuracy of the resulting two-photon transition moment becomes sensitive to the number and character of the excited-states involved. A common truncation in 2PA is the use of the two-state model (2SM), which provides a simplified version of the SOS formalism, as seen in equation 3.21

above. By retaining only the ground and final excited states, 2SM relates the 2PA strength directly to differences in permanent ($\Delta\mu_{0f}$) dipole moments between ground and excited-state, transition (μ_{0f}) dipole moments, and the angle between these two vectors (θ);

$$\delta^{2PA} = \frac{16}{15} \left(\frac{|\mu_{0f}| |\Delta\mu_{0f}|}{\omega_f} \right)^2 (2\cos^2\theta + 1) \quad (3.22)$$

3.5.5: Converting Two-Photon Transition Moments to Cross-Sections

The obtained quadratic response residue corresponding to δ^{2PA} is a microscopic quantity not experimentally observed. To this end, conversion to macroscopic σ^{2PA} in Goppert-Mayer (GM) is necessary through the equation;

$$\sigma^{2PA} = \frac{N\pi^3 \alpha a_0^5 \omega^2}{c} \cdot \delta^{2PA} g(2\omega, \omega_n, \Gamma) \quad (3.23)$$

Here N is an integer number for comparison to experimental values, α is the fine structure constant, a_0 represents the Bohr radius, ω is the photon energy, c corresponds to the speed of light, and finally $g(2\omega, \omega_n, \Gamma)$ is the line shape function that describes the spectral broadening of an excitation and depends on twice the photon energy, the excitation energy ω_n , and Γ is the empirical damping parameter. In theoretical 2PA calculations, it is standard to assume degenerate two-photon absorption, where both photons possess equal energy (ω), each corresponding to exactly half the excitation energy required to reach the final electronic state (as noted in 3.5.2). This simplification is computationally advantageous and aligns with the assumptions used in most quantum chemical response theory frameworks. However, it is important to note that in experimental implementations, this degeneracy condition is not always fulfilled. Non-degenerate two-photon absorption, where the two incident photons have different energies summing to the excitation energy, is more commonly encountered in practice due to instrumental constraints or specific experimental designs. Consequently, care must be taken when comparing calculated cross-sections with experimental values, particularly in interpreting the line shape function $g(2\omega, \omega_n, \Gamma)$, which inherently assumes photon degeneracy in most theoretical treatments. In a publication by Beerepoot *et al.* [131], the authors highlight how the different choices of these parameters can greatly influence the final σ^{2PA} and therefore will cause issues when comparison to literature. To this end, the values used in the conversion within this thesis and some descriptors for the choices are described. A Lorentzian (or Gaussian in noted figures) lineshape function is chosen for $g(2\omega, \omega_n, \Gamma)$, where ω is the photon energy, ω_0 is the excitation energy, and Γ is chosen as Half Width at Half Maximum (HWHM):

$$L(2\omega) = \frac{1}{\pi} \frac{\Gamma}{(2\omega - \omega_0)^2 + \Gamma^2} \quad (3.24)$$

The lineshape function is normalized to one and ω is set to $\omega_0/2$ resulting in:

$$L(\omega_0) = \frac{1}{\pi\Gamma} \quad (3.25)$$

This maximum can be inputted back into equation. 3.23 along with the other selected parameters to produce:

$$\sigma^{2PA} = \frac{4\pi^2 \alpha a_0^5 \omega^2}{c\Gamma} \cdot \delta^{2PA} \quad (3.26)$$

This formula for conversion is used throughout the rest of this thesis.

4. Results and Discussion (gas phase)

4.1: Climbing the CC Hierarchy and TD-DFT Benchmarking of Retinal Models

This chapter is dedicated to the author's own results of 1PA and 2PA spectra of the 11-cis retinal protonated Schiff base (rPSB), its analogs that differ from the wild-type retinal via methylation patterns on its polyene backbone, and reduced size rPSB models *in vacuo*. The results have been published in two separate papers [129, 130].

In sub-chapter 4.1, multiple reduced size rPSB models were calculated at the CC2, CCSD, and CC3 method for their ΔE , f , μ , and 2PA signal. Furthermore, multiple xcfs were also benchmarked against CC3 results to determine the most rational choice for a reliable xcf used in the next stage of calculations that would involve the whole Rhodopsin protein. This pertains to objective 1. The choice of TD-DFT in computing Rhodopsin 2PA signals is due to computational limits, where even CC2 would be too computationally intractable. The xcfs were chosen based on the percentage exact HF energy, as well as their long-range separated contributions.

In this sub-chapter 4.2, the relationship between retinal's structure, more specifically its methylation pattern along its polyene chain, and its 2PA signal were elucidated with the RI-CC2 method. This

sub-chapter aims to determine how different methyl groups within the retinal structure can affect ΔE , f , μ , and consequently the σ^{2PA} cross-section. Furthermore, the angle between $\Delta\mu$ and one-photon transition dipole moments were also used under 2SM hence it can serve as a correlation predictor of 2PA. Sub-chapter 4.2 is also a direct continuation of the findings from 4.1, where the most reliable xcfs were found to be CAM-B3LYP and BHandHLYP. These xcfs along with the MN15 and M11 functionals are also tested on the various retinal analogs against RI-CC2.

4.1.1: The reduced sized models and the benchmarking strategy

As noted in the introduction, the CC methodology increases computational time exponentially as the system size increases. The models used in this study were therefore limited to only the retinal chromophore *in vacuo*, even so, the full retinal structure itself with approx. 50 atoms would still be unfeasible. To make these calculations actually feasible, reduced sized protonated Schiff base (PSB) models were employed. PSB reduces the retinal chromophore by removing the β -ionone ring, leaving only the polyene backbone chain and its methyl constituents. PSB models may be reduced further into smaller variants, where their number denotes the retained conjugated double bonds in the polyene chain, more specifically PSB3 and PSB4 were used here. Among them are also de/methylated variants, PSB3 is reduced to have only 3 double bonds in its polyene chain, as its name suggests, and includes a methylated variant (m-PSB3). PSB4 also follow suite, retaining 4 conjugated double bonds and a single methyl group, as seen in figure 4. Similarly to PSB3, PSB4 also includes a de-methylated variant (dm-PSB4). These smaller models strike a balance between chemical relevance and computational feasibility, allowing the application of more advanced methods like the CC approximate singles, doubles, and triples (CC3). By climbing the CC hierarchy through CC2 (approximate singles and doubles), to CCSD, and then CC3, one can assess their accuracy and establish reliable error margins, insights that can be extended to larger, more complete systems. Despite numerous extensive investigations into the 1PA properties of PSB3 and PSB4 models [133, 135, 169–172], one may still wonder about the effects of removing the β -ionone ring on 2PA properties. This question will be further discussed in the results of 4.2, mainly because the inclusion of the β -ionone ring in higher CC methods like CC3 is currently likely unfeasible. Another factor being this study was the first instance that such large models like PSB3 and PSB4 are being investigated for their 2PA properties. And, it was found that the computational resources required to run PSB3, at approx. 20+ atoms, with the aug-cc-pVDZ basis set was already north of 1700 hours – *let alone* the PSB5, even without the β -ionone ring, at the CC3 level. The PSB3 model in this study were adapted from Grabarek *et al.* 2016 [93], where its ground-state geometry was optimized at the CASPT2 level, along with the Atomic Natural Orbital basis set (ANO-L-VDZP) level with 6 electrons

in 6 π -type molecular orbitals (6,6). PSB3 was also given a methyl group (m-PSB3, see figure 4 below) and later geometry optimized at the CASPT2(6,6)/ANO-L-VDZP level of theory. PSB4 was originally adapted from Szymczak *et al.* [173], where the team used CASSCF with 8 electrons in 8 π -type orbitals averaged over two states. The structure was later refined in the Openmolcas 21.06 software [174, 175], at the CASPT2(8,8)/ANO-L-VDZP level of theory to bring the geometry up to the same level as PSB3. Demethylated (Dm)-PSB4 was constructed from the optimized PSB4 by removing its methyl group.

Within this study, the CC3 method was the highest level on the CC ladder employed. As noted earlier it was shown that including triples excitation, in a perturbative way such as those in CC3, could contribute a significant amount to 2PA strengths, as much as 20%. The goal of this sub-chapter is to investigate the impact of increasing the magnitude of electron correlation from CC2 to CCSD, and then CC3 on 2PA strengths. Secondly, to validate the popular xcfs including BLYP, B1LYP, B3LYP, BHandHLYP, and CAM-B3LYP against CC3 results. And finally, the impact due to Dunning's basis set versus that of Pople variants, as well as the effects of including diffuse functions.

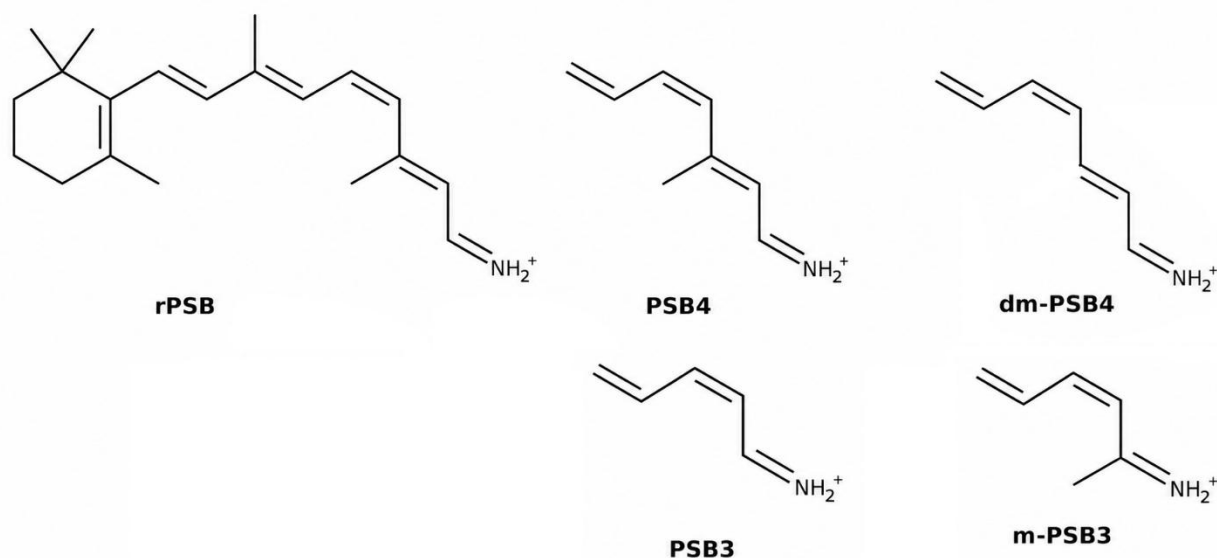


Figure 4. Schematic representation of the various retinal protonated Schiff base (rPSB) used in this chapter. Including the full reduced retinal protonated Schiff base (rPSB), PSB4 and PSB3 together with their methylated (m-PSB3) and demethylated (dm-PSB4) variants.

4.1.2: Computational methods and benchmarking

Methodology: Basis sets

In this study, Dunning's correlation-consistent polarized valence double-zeta (cc-pVDZ) and its augmented counterpart including diffuse functions (aug-cc-pVDZ) were employed. In a similar fashion, we also considered Pople-style split-valence polarized basis sets, namely 6-31G* and its

diffuse-augmented version, 6-31+G*. The effect of diffuse functions in Pople's style b.s. were tested among all structures with the exception of methylated PSB3 (m-PSB3), while the larger Dunning's aug-cc-pVDZ b.s. was only feasible in PSB3 within the CC3 ansatz.

Methodology: TD-DFT

Among the TD-DFT xcf considered in this study are the CAM-B3LYP (65% short-range and 19% long-range) [176], BHandHLYP (50%) [177], B1LYP (25%) [178], B3LYP (20%) [179], and BLYP (0%) [180, 181] functionals. In brackets are the exact HF-exchange energy contributions to the DFT functional. All TD-DFT calculations were performed in the Dalton2020 program [182, 183], where no core orbitals were frozen and deletion of MOs due to linear dependencies are kept at the default eigenvalue threshold of 10^{-6} . 1PA oscillator strengths throughout this study and for the rest of the thesis are provided in length representation, and 2PA cross-sections are obtained using linearly polarized photons with parallel orientation of equal energy, as mentioned in 3.5.3. Both 1PA and 2PA are computed with the linear and quadratic response formalism, respectively, unless otherwise stated. The change in dipole moments are calculated according to:

$$\Delta\mu = \sqrt{\sum_a (\mu_{a,ES} - \mu_{a,GS})^2} \quad \text{where } a = x, y, z \quad (4.1)$$

Methodology: CCx

The spectral results from CC2, CCSD, and CC3 methods were all also computed using the same Dalton2020 suite. Frozen core embedding was applied to all CC methods, the number of core orbitals frozen varies with system size, where PSB3 had 6 frozen cores, m-PSB3 had 7, demethylated PSB4 (dm-PSB4) had 8, and PSB4 (with methyl group) had 9 frozen cores. To further test the effect of frozen core orbitals in PSB models, non-frozen core orbital calculations were repeated for CC2 and CCSD which demonstrated a small discrepancy of ~0.1 to 0.2% depending on the property of interest, method, and basis set used. Furthermore, to validate CC calculations and determine the amount of double excited character, %T2 diagnostics were analyzed. For the S_1 state all CC calculations provided a $T2 < 9\%$ (with the exception of CC3 where $T2 < 13\%$), within the recommended thresholds of $T2 < 15\%$ [184]. Similarly to TD-DFT, linear and quadratic response was employed and $\Delta\mu$ computed in the same manner.

4.1.3.1 Results and Discussion: CC hierarchy

model	method	ΔE	δ^{2PA}	σ^{2PA}	f	$\mu_{01}(\text{au})$	μ_0	μ_1	$\Delta\mu$
PSB3	CC3(Dun)	4.06	2098	12.69	0.614	2.483	4.333	-	-
	CCSD(Dun)	4.20	2065	13.33	0.688	2.586	4.324	0.714	4.618
	CC2(Dun)	4.08	2016	12.29	0.697	2.641	4.204	0.824	4.503
PSB3	CC3	4.10	2456	15.15	0.639	2.521	4.458	-	-
	CCSD	4.23	2425	15.92	0.714	2.624	4.435	0.831	4.906
	CC2	4.12	2201	13.70	0.722	2.675	4.267	0.891	4.642
Meth-PSB3	CC3	4.25	2568	17.01	0.664	2.526	6.786	-	-
	CCSD	4.39	2567	18.11	0.734	2.613	6.794	4.188	5.210
	CC2	4.22	2475	16.17	0.745	2.684	6.628	4.275	4.947
Dm-PSB4	CC3	3.33	12234	49.81	1.018	3.531	6.660	-	-
	CCSD	3.46	11065	48.40	1.127	3.649	6.652	1.036	5.945
	CC2	3.36	9611	39.65	1.142	3.727	6.225	1.118	5.466
PSB4	CC3	3.34	11877	48.52	1.048	3.580	6.568	-	-
	CCSD	3.47	10679	47.01	1.158	3.692	6.578	0.908	5.809
	CC2	3.33	10796	43.84	1.175	3.795	6.148	0.713	5.631

Table 1. S_1 state electronic properties of PSB3, m-PSB3, dm-PSB4, and PSB4 calculated using CC2, CCSD, and CC3.

Rows with (Dun) employ the aug-cc-pVDZ basis set, while the remaining calculations use 6-31+G*. ΔE is reported in eV, δ^{2PA} in a.u., σ^{2PA} in GM, μ_{01} in a.u., and μ_0 , μ_1 , and $\Delta\mu$ in Debye. The CC3 excited-state permanent dipole moments, and consequently $\Delta\mu$ values, were unrecoverable.

Starting with the PSB3 model, 1PA transition energies, microscopic 2PA strengths and its macroscopic conversion to cross-sections, transition dipole moments, permanent electric dipole moments and their differences (between ground and n-excited-state) for the lowest singlet S_1 state are presented in table 1 for all CC methods. Since PSB3 is the only model studied here within CC3 that utilizes aug-cc-pVDZ b.s., during the discussion of PSB3 in CC methods, all values will reference this b.s. unless otherwise stated. In PSB3 the S_1 state is of π - π^* and charge-transfer character given its large $\Delta\mu$. This state is also dominated by a single-reference wave function where the contributions from doubly excitations are below 8% and 9% for CC2 and CCSD, respectively, while CC3 was below 13% for both aug-cc-pVDZ and 6-31+G* b.s. The excitation energy of CC3 is 4.064 eV, compared to MS-CASPT2/ANO-L-VQZP of 3.96 eV [93], makes the deviation a small overestimation of ~ 0.10 eV. While CCSD is slightly more overestimated at 4.197 eV, CC2 follows CC3 more closely at 4.078 eV, likely due to cancellation of errors. Interestingly other spectral properties like oscillator strengths do not follow this pattern. CC3 demonstrates an f value of 0.614, while CC2 gives the highest value at 0.697 marking a 13.5% increase, followed closely by CCSD at 0.688, a 12% increase. This, to no surprise, also leads to transition dipole moments (μ_{01}) following the same pattern as f . CC3 demonstrates the lowest μ_{01} at 2.483 a.u., followed by CCSD with 2.586 a.u. and CC2 at 2.641 a.u., marking a 4% and 6% deviation respectively. Looking over at 2PA strengths, it becomes clear as well here that inclusion of full doubles excitation (as opposed to CC2's perturbative doubles) can also diminish deviations from CC2 at 4%, down to 1.5% of CCSD.

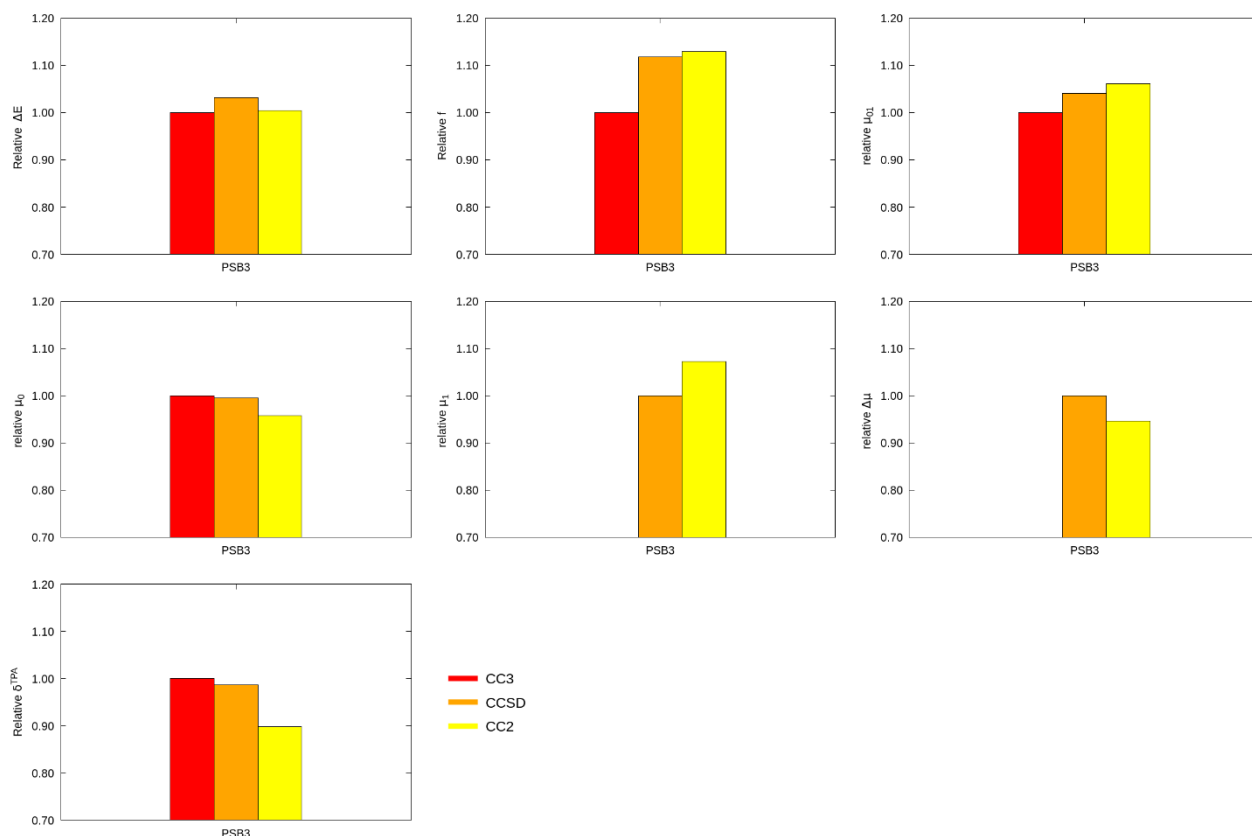


Figure 5. Histogram of relative difference of CC2 and CCSD in the aug-cc-pVDZ basis set of all spectral properties in the S_1 state for the PSB3 model. All values have been normalized to CC3/aug-cc-pVDZ with the exception of excited-state dipoles (μ_1), that were normalized to CCSD instead.

When considering the macroscopic converted σ^{2PA} , the absolute deviations from CC2 and CCSD are closer to each other but also flipped. Here, CC2 has the slight advantage due to higher excitation energy agreement with CC3, leaving macroscopic σ^{2PA} of CC2 at 3.1% underestimated, while CCSD is 5% overestimated. The charge-transfer character of the S_1 state is confirmed via a $\Delta\mu$ of 4.5 to 4.9 D, depending on the specific CC method and b.s. used. CC3 ground-state dipole moment was recorded at 4.333 D, with CCSD in tow at 4.324 D, marking a deviation of merely 0.2%. On the other hand CC2 is slightly more deviated at 4.204 D, a 3% underestimation. Unfortunately, CC3's S_1 (nor its higher states) excited-state permanent electric dipole moment was not recoverable and therefore only CCSD and CC2 data are available. Since CCSD had performed markedly better than CC2 in recovering CC3 ground-state dipole moments, as well as δ^{2PA} strengths, the author here compares CC2 to that of CCSD instead. In both cases of CC2 and CCSD, S_1 demonstrated a large decrease in dipole moment amplitudes at 0.824 D and 0.714 D respectively, leading to the large $\Delta\mu$ as was noted in the introduction [69, 82, 92, 96]. The CC2 μ_1 discrepancy accounts for approximately a 14% overestimation from that of CCSD, at the same time, CCSD's S_{0-1} $\Delta\mu$ is 4.618 D, merely 2.5% larger than CC2's at 4.503 D. This decrease in discrepancy in $\Delta\mu$ is due to the oppositely signed errors when

subtracting the S_0 and S_1 vectors to produce their differences (see 4.1.2 $\Delta\mu$ equation) in CC2. This fortunate cancellation of errors likely also contribute to CC2's good performance in recovering CC3 δ^{2PA} values. To sum up, it seems that for most 1PA spectral properties like oscillator strengths, transition and excited-state permanent dipole moments CC2 is the most overestimated, more so than CCSD. While 2PA properties and their associated $\Delta\mu$ are generally underestimated by only a few percent.

Moving to the 6-31+G* b.s., trends remain true to aug-cc-pVDZ with the exception of δ^{2PA} and $\Delta\mu$ for the S_1 state in CC2 having larger deviations. In Pople's style b.s. the δ^{2PA} jumps from ~ 2200 a.u. in CC2 to ~ 2450 in CC3, an underestimation of over 10%, approximately 2.5 times larger than Dunning's style b.s. ground-state μ_0 , although its deviation jumped up by double of that of aug-cc-pVDZ, the total percentage deviation is merely 0.5% and 4% for CCSD and CC2 when compared to CC3, respectively. On the other hand, the μ_1 for Pople's CC2 deviates from CCSD only half as much as its Dunning's counterpart, at 7%. These two states combine to make the deviation of $\Delta\mu$ only 5.4%, yet still twice the error margin of Dunning's likely leading to the larger deviation in 2PA as well. CCSD on the other hand gives a δ^{2PA} value of 2425 a.u., only $\sim 1.2\%$ below the CC3 value of 2456 a.u. and slightly closer than in aug-cc-pVDZ. Despite the relatively good agreement within the CC hierarchy in PSB3, δ^{2PA} remains sensitive to the choice of b.s. Switching from aug-cc-pVDZ to 6-31+G* increases δ^{2PA} by approximately 17% in both CC3 and CCSD, but by less than 10% in CC2. This weaker basis-set response at the CC2 level explains why its underestimation relative to CC3 increases from approximately 4% in aug-cc-pVDZ to over 10% in 6-31+G*. With that said, there is a high computational cost when jumping from Pople's to Dunning's b.s. as well as climbing the CC hierarchy. To perform quadratic response in CC2 for the aug-cc-pVDZ b.s. using 64-Bit version of Dalton at approx. 20 GB of memory it takes approximately 50 hours. CCSD brings the time up to 71 hours, and CC3 jumps up to a whopping 1771 hours. Switching to the 6-31+G* b.s. can reduce computational times by 6-8 times, depending on the method. The 35 fold increase in computational time from CC2 to CC3 therefore leads to an increase of only approximately 4% in the calculated δ^{2PA} for PSB3/aug-cc-pVDZ.

M-PSB3 was computed using the 6-31+G* b.s., and its results are displayed in table 1. All trends in m-PSB3 remains the same from its demethylated counterpart with the exception of the permanent electric dipole moments. Excitation energy in CC2 remains to have a better agreement to CC3, at a mere 0.7% underestimation, compared to CCSD's overestimation at 3.3%. Oscillator strengths and μ_{01} also follows PSB3 trends, fidelity to CC3 increases with increasing electron correlation with the largest deviations occurring in oscillator strengths at 10-12% overestimated. In δ^{2PA} , the trend indeed

remains the same but with greater amplitudes. CC2 error rate is much smaller than its non-methylated structure at 3.6%, on the other hand CCSD's already great performance in PSB3 is even better here where the error is practically negligible at 0.04%, essentially off from CC3 by only 1 a.u. Ground-state dipole moments is the only property that breaks the trends established in PSB3, but only by slightly. Despite having similarly low deviations to its PSB3 counterpart, whereby both CC2 and CCSD deviates from CC3 only by 2.3% and 0.1% respectively, here CCSD actually overestimates CC3 rather than underestimate it. However, S_1 dipole moments in CC2 is only ca. 2.1% overestimated from CCSD's, far less than PSB3's for the same b.s. at 7%. Thus, it is not surprising that the resulting $\Delta\mu$ of CC2 also shows low deviation of ca. 5.1% from CCSD's. Yet when looking back at PSB3, one may notice that CC2's $\Delta\mu$ deviates from CCSD relatively similarly to what is observed here (5.4% in PSB3, vs 5.1% in m-PSB3) – and still the δ^{2PA} of both have starkly different deviations to their respective CC3 references. This likely stems from the nature of the quadratic response formalism not being governed by $\Delta\mu$, but is instead more sensitive to the accuracy of the individual permanent dipole components and transition moments (full quadratic equation in 3.5.4-5), which in the case of μ_1 for m-PSB3 indeed exhibits better agreement with CCSD than that of PSB3. In summary, the individual dipole moments themselves are more accurate in m-PSB3 than that of 'regular' PSB3 (in 6-31+G* b.s.), leading to improved δ^{2PA} prediction by CC2 through physically consistent polarization of these dipole moments rather than error compensation.

Extending the polyene backbone chain by an extra carbon-carbon double bond from the PSB3 model to the dm-PSB4 has a noticeable effect on all spectral properties. Interestingly, all properties increase by a relatively wide margin from its PSB3 counterpart with the exception of excitation energies that actually decreases instead. Regardless, the trends established in PSB3 remains the same here. Analogously to the PSB3 models, dm-PSB4's lowest singlet S_1 is a charge-transfer π - π^* state characterized by its large $\Delta\mu$ (5.466-5.945 D). Both CC2 and CCSD display a T2 percentage contribution < 9%, where CC3 brings this value slightly up to < 14%. Similarly to the previous two structures, CC2 outperforms CCSD by deviating from CC3 excitation energies by merely 0.8%, while CCSD brings this up to nearly 4%. Other properties also follow suite, oscillator strengths are overestimated at the CC2 level by 12% and decrease to ca. 11% at CCSD. Transition dipole moments are less deviated than its oscillator strengths counterpart at 5.5 and 3% for CC2 and CCSD, respectively. Similarly to PSB3, but unlike m-PSB3, dm-PSB4 here underestimates ground-state μ_0 in both CC2 and CCSD at 6.5% and 0.1% respectively. This is of course in contrast to earlier in m-PSB3, where CCSD actually overestimates CC3 instead. Nevertheless, the high accuracy of CCSD to CC3's ground-state μ_0 is seen in all structures analyzed so far. Interestingly, μ_1 of CC2 overestimates CCSD by approximately 8%, while its $\Delta\mu$ underestimates it by the same 8% margin.

Therefore, it comes as no surprise that δ^{2PA} for CCSD would also be more accurate than that of CC2's and indeed this is the case. CCSD deviates from CC3 by ca. 10% while CC2 brings this value up to a whopping ca. 21.5%, making CC2's dm-PSB4 the least accurate δ^{2PA} so far. This, coupled with the low deviation in transition dipole moments and excitation energies, makes one believe that the increase in the underestimation of δ^{2PA} likely stems from the enhanced polarizability and electronic delocalization associated with the increase of π -conjugated double bonds in dm-PSB4. The perturbative treatment of double excitations in CC2 is too approximate to adequately describe excited-state polarization in this system, resulting in an incomplete quadratic response. Additionally, despite the similar trend of increasing the electron correlation in CC methods bringing about higher fidelity towards CC3, absolute differences are starkly different from PSB3 as noted earlier. Excitation energy decreases across the board by ca. 0.8 eV, accounting for approximately 20% when compared to either m-/PSB3. Oscillator strengths increase by ca. 0.4, over a 40% increase. Transition dipole moments also increased by approximately 1 a.u., while permanent dipole moments for the ground-state jumped up by approximately 2.0-2.2 D, practically a 50% increase across all CC methods. S_1 permanent dipole moments also increased but by a lesser margin at approximately 25%, this results in the $\Delta\mu$ being increased by 18-21%, or approximately 0.8-1.0 D. The absolute increases across all the aforementioned spectral properties cumulate to also an extremely large increase in δ^{2PA} . 2PA strengths jump from around 2200-2500 a.u. in either PSB3 models to dm-PSB4 sitting comfortably at 10-12000 a.u. range.

The addition of a methyl group to dm-PSB4 generates the 'regular' PSB4. Like the previous larger structures (that is, larger than PSB3) PSB4 is computed only at the 6-31+G* b.s. where trends take after its de-methylated counterpart with the exception of δ^{2PA} , and less noticeably S_1 excited-state dipole moments. Just like all the other models studied here, excitation energies in CC2 demonstrates great agreement with CC3, with an underestimation of only 0.3% or the equivalent of 0.01 eV. PSB4's energy however, similarly to the only other methylated structure m-PSB3, are the only two models who's energy is underestimated in CC2. CCSD on the other hand brings the deviation up to 3.8% overestimated. Oscillator strengths and transition dipole moments also follow the trends established in previous models where CCSD improves upon CC2 by a few percentages, in the case of PSB4 this is equivalent to an increase in fidelity from 12 down to 10% overestimation in oscillator strengths and 6 to 3% in μ_{01} . Most notable in PSB4 is the δ^{2PA} . Unlike previously where CCSD would consistently outperform CC2 by a few percentages up to almost 12% (aug-cc-pVDZ b.s. inclusive), here in PSB4 CC2 actually outperforms CCSD by approximately 1%. CCSD in PSB4 underestimates CC3 by approximately 10%, while CC2 improves upon this slightly at 9% as seen in figure. 6. Looking over to permanent electric dipole moments helps shed light on this matter. In PSB4 the

ground-state μ_0 like in all previous models demonstrated a better agreement with CC3 in CCSD than that of CC2, more specifically CCSD is deviated by 0.15% while CC2 brings this value up to approximately 6.4%. However, the direction of the deviation doesn't follow that of dm-PSB4 or PSB3 where both CCSD and CC2 underestimates the CC3 reference. Instead, here like in m-PSB3, only CCSD is overestimated while CC2 remains underestimated. The S_1 excited-state dipole moments on the other hand paint a rather interesting picture. CC2 underestimates CCSD's value by over 21%, making it the least accurate S_1 dipole moment in this study by far. Nonetheless, the difference in $\Delta\mu$ of CC2 to $\Delta\mu$ of CCSD is actually only approximately 3%. Similarly to PSB3/aug-cc-pVDZ, here due to the cancellation of oppositely signed errors, CC2's $\Delta\mu$ has a relatively good agreement with CCSD. It seems that for PSB4 three main factors contribute to the higher δ^{2PA} . Firstly, δ^{2PA} is proportional to the square of transition dipole moments (see eqn. XX), the increase in μ_{01} could push 2PA strengths higher than its CCSD counterpart. Secondly is the lower (and better agreement with CC3) excitation energy, which enters into the quadratic response as the denominator, also pushing 2PA strengths slightly higher with its decrease. And lastly, although quadratic response is sensitive to the individual dipole moments of each state, the cancellation of errors that leads to a more accurate $\Delta\mu$ reflects a net dipole shift that more closely approximates the true charge redistribution upon excitation. This results in a fortuitously accurate depiction of excited-state polarization, and indirectly enhancing the δ^{2PA} .

In the previous PSB3 models, when going from a non-methylated structure to a methylated one, a general increase in all spectral properties can be seen (with the exception of transition dipole moments in CCSD). In PSB4 models however, this is not necessarily the case. For 1PA properties like excitation energies, oscillator strengths, and transition dipole moments, there indeed observed a slight increase. Excitation energies increase consistently, with a slightly more profound impact in PSB3 than in PSB4. In going from PSB3 to m-PSB3 around a 0.10 eV to 0.16 eV bump is observed, while PSB4 pairs bring this down to 0.01 to 0.03 eV only. Oscillator strengths and transition dipole moments also follow similar trends to what was seen with the PSB3 models. In PSB4, oscillator strengths increase slightly by approximately 3%, while μ_{01} brings this value down to 1.3-1.8%, slightly higher than that of the PSB3 pair. However, unlike previously in PSB3 and m-PSB3, 2PA associated properties in the PSB4 pair experiences a decrease with a notable exception in CC2. Ground-state dipole moments demonstrates a <3% decrease across all methods, while S_1 dipole moments experience a large 12 and 36% decrease in CCSD and CC2 methods respectively. When it comes to $\Delta\mu$, CCSD decreases by approximately 2% but conversely CC2 actually increases by 3% instead, in going from dm-PSB4 to PSB4. And thus, δ^{2PA} also increases by approximately 12% in

CC2, as opposed to CC3 and CCSD's decrease of 3-4%. Altogether, these positive shifts help identify how CC2 improved its fidelity to CC3 in terms of δ^{2PA} in the PSB4 model relative to CCSD.

To summarize this sub-chapter, across all PSB models CC2 reproduces CC3 excitation energies with greatest fidelity surpassing CCSD. In PSB3 and dm-PSB4, CC2 consistently provides the closest excitation energies to CC3, typically deviating by less than 1%, whereas CCSD tends to slightly overestimate by 3–4%. This accuracy, however, is offset by a systematic overestimation of 1PA intensities such as oscillator strengths and transition dipole moments, where CC2 overestimates CC3 by ca. 10 – 13%, exceeding CCSD at 8 –10%. For 2PA properties, CC2 performs comparably well in PSB3, its δ^{2PA} and $\Delta\mu$ values differ by only a few percent from CC3 due to fortuitous error cancellation between the ground and excited-state dipoles, but deteriorates as system size increases. For the larger dm-PSB4 system, CC2's limited perturbative doubles manifest as a less accurate δ^{2PA} signal, highlighting the method's sensitivity to increased delocalization due to the extension of the conjugate double bond system. Notably, upon methylation to PSB4 this trend pivots. CC2 slightly surpasses CCSD in δ^{2PA} (by ~1%), consistent with its lower excitation-energy errors and balanced $\Delta\mu$. Taken together, these results position CC2 as a practical and predictive baseline for two-photon response. It delivers consistently accurate excitation energies, captures structure–property trends across model modifications, and provides quantitatively competitive δ^{2PA} at a fraction of the cost.

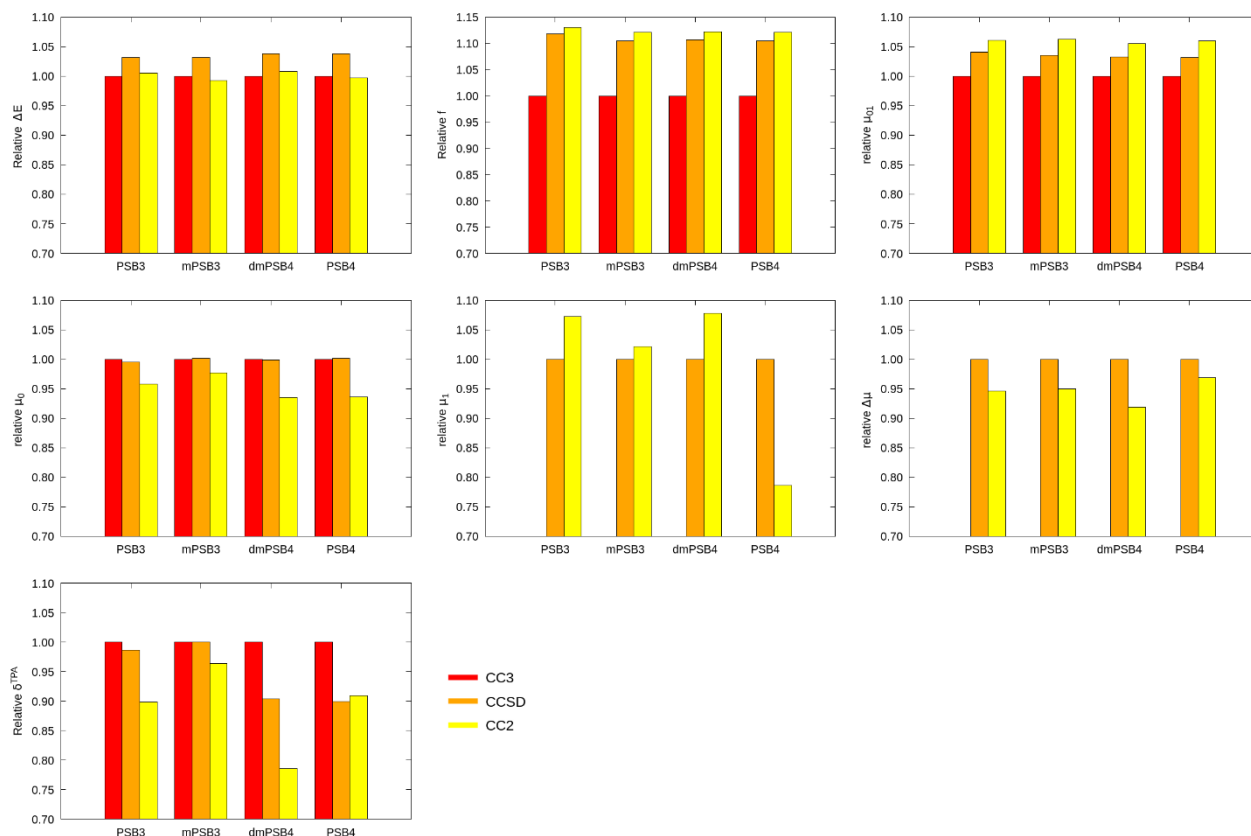


Figure 6. Histogram of relative difference of CCx/6-31+G* of all spectral properties in the S_1 state for all available models. All values have been normalized to CC3/aug-cc-pVDZ with the exception of excited-state dipoles (μ_1), that were normalized to CCSD instead.

4.1.3.2: CC hierarchy in higher excited-states of PSB3

In higher excited-states, CC3 results were not recoverable in the aug-cc-pVDZ b.s. and therefore this section will be mainly focused on referencing the 6-31+G* b.s. up to the S_4 state unless stated otherwise. CC2 and CCSD results on the other hand were computed up to the S_6 state, regardless, the number of 2PA bright states (ergo, absorption peaks) remain at around 3 for all CC methods as seen in Fig. 7 and table 2. The T1 percentages are also listed in table 2, along with the spectral properties for each method. Percentage of single and double excitation character, T1/T2, in all CC methods, with the exception of CC3's second lowest excited-state, demonstrates decent to great performance, hovering at $T2 < 14\%$. CC3/6-31+G*'s S_2 state demonstrated a T2 approximately at $\sim 30.5\%$, well below the recommended threshold of most publications at $< 15\text{--}20\%$ [141, 184]. Looking over table 2, one may quickly notice two major points of contention with the data. Firstly is the presence of negative δ^{2PA} values found in CC2 and CCSD within both basis sets in the S_4 state.

6-31+G*	states	λ	δ^{2PA}	σ^{2PA}	f	μ	$\Delta\mu$	%T1	Hr:Min
CC2	ground	-	-	-	-	4.267	-	-	-
	1	602	2201	13.70	0.722	0.891	4.642	92.76	6:44
	2	389	3345	49.77	0.154	1.816	2.468	91.45	11:53
	3	375	2	0.03	0.000	0.575	3.986	94.28	19:27
	4	355	-5266	-	0.218	3.804	0.577	93.97	29:03
	5	349	6	0.10	0.001	6.959	10.955	95.03	23:29
CCSD	6	332	206	4.23	0.003	3.458	6.351	95.32	-
	ground	-	-	-	-	4.435	-	-	-
	1	586	2425	15.92	0.714	0.832	4.906	91.66	11:00
	2	387	1931	29.16	0.119	2.732	1.718	87.5	10:58
	3	372	2	0.03	0.001	0.893	3.614	93.37	30:13
	4	351	-1176	-	0.180	3.001	1.516	92.51	26:58
CC3	5	346	6	0.12	0.001	6.514	10.607	94.69	36:24
	6	329	121	2.50	0.004	3.494	6.323	94.9	-
	ground	-	-	-	-	4.458	-	-	-
	1	605	2456	15.15	0.639	-	-	87.78	248:17
	2	440	121	1.41	0.124	-	-	69.48	389:07
	3	382	2	0.02	0.000	-	-	90.06	616:35
aug-cc-pVDZ	4	366	2113	35.51	0.135	-	-	86.25	762:15
	states	λ	δ^{2PA}	σ^{2PA}	f	μ	$\Delta\mu$	%T1	Hr:Min
	CC2	ground	-	-	-	4.204	-	-	-
	1	608	2016	12.29	0.697	0.824	4.503	92.64	50:43
	2	393	3774	55.01	0.163	1.960	2.267	91.29	88:03
	3	380	2	0.03	0.000	0.569	3.885	94.05	127:22
CCSD	4	367	-1691	-	0.211	3.530	0.771	93.94	165:09
	5	351	7	0.12	0.002	6.930	10.938	94.79	201:29
	6	336	22	0.43	0.001	0.631	3.783	94.21	-
	ground	-	-	-	-	4.324	-	-	-
	1	590	2067	13.34	0.688	0.714	4.618	91.76	71:39
	2	389	2266	33.78	0.130	2.684	1.658	88.09	146:55
CC3	3	376	1	0.02	0.001	0.945	3.435	93.2	206:06
	4	363	-375	-	0.179	2.902	1.507	92.99	274:14
	5	346	7	0.13	0.002	6.344	10.394	94.63	423:55
	6	332	9	0.18	0.001	0.907	3.476	93.74	-
	ground	-	-	-	-	4.333	-	-	-
	1	611	2098	12.69	0.614	-	-	87.89	1771:39

Table 2. All available electronic properties for higher excited-states $S_1 - S_n$ for the PSB3 model in CC3/CCSD/CC2 methods in 6-31+G* (top) and aug-cc-pVDZ (bottom) basis sets. λ is reported in nm, δ^{2PA} in a.u., σ^{2PA} in GM, μ , and $\Delta\mu$ in Debye. T1% and computational time are also present.

These negative δ^{2PA} values stem directly from the differently signed left and right eigenvectors of both CCSD and CC2's quadratic response formalism when contracted with the dipole operator $\hat{\mu}$. Secondly, and perhaps less noticeable unless observing figure 7, is the difference in state ordering between CC3 and the other two CC methods. At higher excited-states in CCSD and CC2, a smaller absorption peak near 330 nm appears, corresponding to the computed S_6 state. In CC3, however, the analogous state appears as computed S_2 and is located instead at around 440 nm. At the same time, when referring back to Table 2, this state shows a high degree of doubly excited character ($T2 > 30\%$). CC3, while more accurate overall, may still struggle with states dominated by double-excitation character because its treatment of higher excitations remains approximate. Nonetheless, the relatively high $T2$ contribution indicates that CC2 and CCSD are expected to describe this state less reliably. CC3 includes an approximate iterative treatment of connected triple excitations, whereas CCSD and CC2 do not include connected triples, with the doubles contribution also treated approximately in CC2. Consequently, states with substantial double-excitation character may be described less reliably by CCSD and especially CC2. Therefore, for PSB3, CC3 provides an indication of how this state may appear when higher-order correlation is included. The available excitation-energy and intensity patterns suggest that the weak state appearing as computed S_2 in CC3 may instead appear as computed S_6 in CC2 and CCSD. The ordering of higher excited-states differs between CC3 and the lower-level CC methods, particularly for states with substantial double-excitation character. Throughout the remainder of this section, all states are discussed according to their original computed-state labels as reported in Table 2.

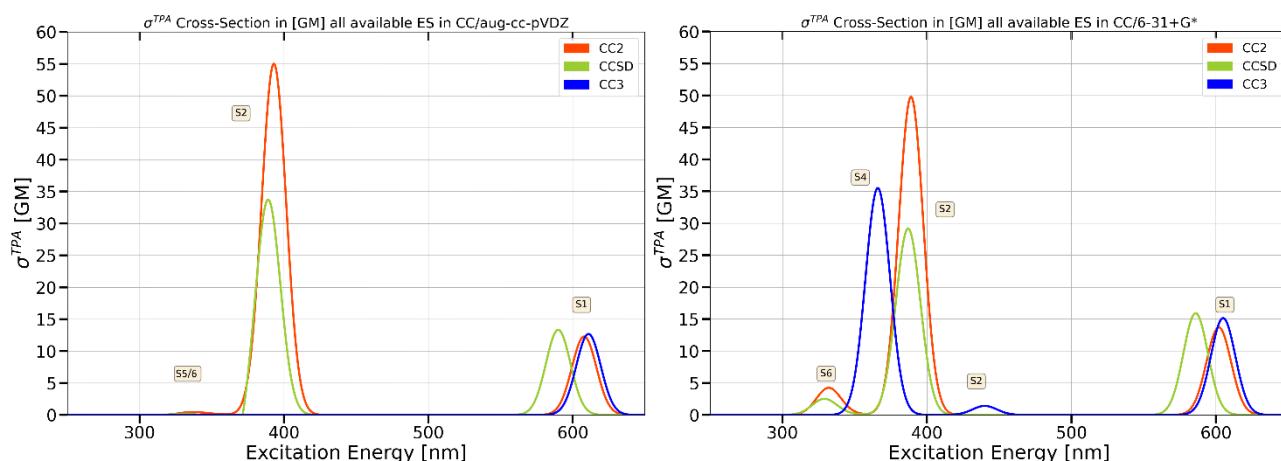


Figure 7. Simulated σ^{2PA} spectra of PSB3 for all available excited states calculated using CC2, CCSD, and CC3 with (left) aug-cc-pVDZ and (right) 6-31+G*. The annotations indicate the computed-state assignments of the corresponding spectral features.

For the purposes of this sub-chapter, excitation energies will also be referred to in nanometers (nm), as opposed to electronvolts (eV) as was done previously. This is to allow easier comparison within the absorption spectra, both among the CC methods themselves and with other references discussed

further below, which are traditionally reported in nm. Under this shaky correspondence, the computed S_6 states of CC2 and CCSD/6-31+G* are blue-shifted relative to CC3 S_2 by approximately 110 nm, or around 25%. Conversely, the computed S_2 states of CC2 and CCSD are red-shifted relative to CC3 S_4 by approximately 20 nm, or around 6%. The comparable σ^{2PA} amplitudes of the first pair, shown in Figure 7, are consistent with this proposed change in state ordering, although they do not establish the assignment on their own. Despite the large excitation-energy difference between CC2/CCSD S_6 and CC3 S_2 , the corresponding 2PA peaks remain in the same general ballpark of 1.5 – 4 GM. Nonetheless, due to the state being relatively dim, such deviations still constitute a 77 and 200% overestimation by CCSD and CC2 respectively. This large deviation is again likely due to the large contributions from doubly excitation character of this state. CC3's S_4 paints a rather better picture, here where singly excitation is dominant (CC3 $T_2 < 13.7\%$) CCSD (S_2) underestimates CC3 by 18% or approximately 6GM, while CC2 (S_2) overestimates it by a whopping 40% or approximately 14GM. Despite the higher singly excitation dominance in this state, both CC2 and CCSD deviates more aggressively from CC3 than in S_1 despite T_2 percentages being within 1-2% of each other. It seems that lower CC methods within the hierarchy become less reliable the higher the calculated excited-states are. Interestingly, this doesn't seem to be the case for 1PA properties like oscillator strengths and transition dipole moments. When comparing the oscillator strengths in table 2 in CC3's S_2 state (0.124) to that of CCSD and CC2's S_6 (0.003-0.004) one quickly notices that 1PA oscillators in CCSD and CC2 are actually dark-states. Interestingly, oscillator strengths appear to follow the original (S_2) computed-state ordering reported in Table 2 more closely than the excitation energies and 2PA intensities. Unsurprisingly this is due to 1PA oscillator strength properties being dominated by singly excited configurations, regardless, the deviations in S_2 are rather high for CC2 at 24% overestimated, while CCSD surprisingly performed even better than in its S_1 at a mere 4% underestimation. Computed S_4 state on the other hand, increases both deviations to 61 and 33.5% overestimated for CC2 and CCSD respectively. Dipole moments in these higher excited states are difficult to compare because the corresponding excited-state dipole moments were not recoverable at the CC3 level. Differences in the ordering and character of the higher computed states between CC2, CCSD, and CC3 also make a reliable one-to-one comparison uncertain. Therefore, the author omits the discussion of this property within the aspect of the CCx hierarchy.

Although CC3/aug-cc-pVDZ results are available for S_1 , they were not recoverable for the higher excited states discussed here. CCSD and CC2 can still provide insight into the absolute differences between the two basis sets. Below is a brief discussion on the effects of going from a more complete basis set like aug-cc-pVDZ to the smaller 6-31+G* (or vice-versa). As seen in Fig. 7 above, the most notable differences are found in the computed S_2 and S_6 states. In both basis sets, CC2 and CCSD

place the bright higher-state feature at S_2 and the weaker feature at S_6 . This differs from CC3/6-31+G*, where S_2 is weak and the bright higher-state feature appears at S_4 . The S_6 response is also substantially weaker in aug-cc-pVDZ than in 6-31+G*, as seen in Table 2. Examining the same Table 2 further, the S_5 excited-state permanent dipole moment obtained with CC2/aug-cc-pVDZ is 6.930 D, very close to the CC2/6-31+G* value of 6.959 D, corresponding to a difference of approximately 0.4%. The same can also be said for the CCSD pair where the deviation is only approximately 2.7%. In fact, all permanent electric dipole moments from the ground up to the computed S_5 state don't deviate from each other by more than 16.5%, with the max belonging to the S_1 state of CCSD. However, the computed S_6 permanent dipole moment is substantially larger in the 6-31+G* basis set, increasing by 448% and 285% relative to aug-cc-pVDZ for CC2 and CCSD, respectively. Together with the larger S_6 σ^{2PA} values in 6-31+G*, this indicates that S_6 is considerably more sensitive to the choice of basis set than the lower states. The available data are not sufficient to determine whether this reflects a change in state character, mixing with a nearby state, or displacement of the corresponding state to a higher root in aug-cc-pVDZ, since states above S_6 were not computed. Unlike in the S_1 state where the σ^{2PA} intensity increases when going from the larger Dunning's basis set to the Pople one, the computed S_2 state is the opposite. Both CC2 and CCSD experiences a decrease in σ^{2PA} intensity by 9.5 and 13.6% respectively when switching to the smaller Pople basis set. Similarly, the δ^{2PA} also decreases by a similar amount at about 11.4 and 14.8%. Yet, the $\Delta\mu$ in Pople's basis set is increased from Dunning's by 9% and 4% in CC2 and CCSD respectively, for this computed S_2 state. Looking over to the oscillator strengths (and therefore transition dipole moments) could help explain the slight difference, where Pople's oscillator strengths are generally smaller by 5.5 and 8.5% in CC2 and CCSD, respectively. Meanwhile, excitation energies follow the same increase as was observed in the S_1 state (in terms of electronvolts). Most excited-states differ by approximately 1% or less between the two basis sets, with the largest deviation consistently occurring for the computed S_4 state at roughly 3% when going from aug-cc-pVDZ to 6-31+G* basis set. Interestingly, the computed S_4 and S_5 states in the 6-31+G* basis set may lead one to believe state-mixing may be at play due to the close excitation energy levels separated by only 6 nm in CC2 and 5 nm in CCSD. These states are also accompanied by the aforementioned large negative δ^{2PA} strengths in the S_4 and an extremely weak δ^{2PA} in the S_5 . However, the same S_4 and S_5 states in aug-cc-pVDZ are more separated at 16 and 17 nm in CC2 and CCSD, respectively. And despite the larger gap between the two states in aug-cc-pVDZ, S_4 retains a large negative δ^{2PA} while S_5 also remains weak. Because the negative S_4 value persists even when S_4 and S_5 are more separated in aug-cc-pVDZ, the author ascribes the negative S_4 response at the CC2/CCSD level to a methodological limitation for this state rather than to simple state mixing or intensity borrowing.

4.1.4: Impact of different XCF

Calculations of PSB3 using TD-DFT with various xcf referenced against CC3 at the 6-31+G* and aug-cc-pVDZ basis sets for the S₁ state are displayed in table. 3 and figure 8 below. Looking at relative differences within the respective basis sets, there is a clear effect on increasing the total amount of exact HF energy within the xcf on all electronic properties across all model structures. Looking firstly over to PSB3/6-31+G*, the first row in figure 8 displays 1PA dominant properties and excitation energies. Here, one can see that from BLYP up till BHandHLYP (purple bars till lime-green bars), the equivalent of an increase in short-range exact HF energy of 0% till 50%, aligns in a stepwise fashion with a general increase in excitation energies from 3.99 eV to 4.33 eV. Interestingly, by increasing short-range energy contributions even further to 65% and also including long-ranges at 19%, CAM-B3LYP drops excitation energy down from BHandHLYP to 4.23 eV. Coincidentally, B3LYP in terms of ΔE performed the best with an overestimation of merely 1%, likely due to a fortuitous cancellation of errors. This trend of increasing HF energy contributions also increasing certain electronic properties is generally tied to 1PA properties, as seen not just in ΔE but also in oscillator strengths and transition dipole moments of figure 8 as well. Unsurprisingly the largest jumps in these generally occur in going from BLYP (no HF energy contributions) to B3LYP (20% HF energy contribution) or between B1LYP (25%) and BHandHLYP (50%), coinciding with the largest jumps in HF energy contributions. For the most case, TD-DFT converges towards CC3 results for 1PA properties with the increase of HF-energy contributions. The exception being, as noted earlier with excitation energies, BHandHLYP, which offers the highest enhancement of all TD-DFT methods. This leads to a consistent overestimation of the CC3 reference within 1PA properties, like oscillator strengths by 9.7%, and μ_{01} by 1.9%, for the S₁ state. Nevertheless this is quite the improvement from CC2 and CCSD's 12 – 13% overestimation in oscillator strengths and 4-6% in μ_{01} . At the same time, CAM-B3LYP overestimates oscillators by merely 3.8% and μ_{01} by less than half a percent, seemingly outperforming all other xcfs and CC methods. Although CC2 still demonstrates the highest fidelity in excitation energies, CAM's ability to match (Pople basis set) or outperform (Dunning basis set) CCSD's energies really demonstrates the capability of this DFT

6-31+G*										
method	ΔE	δ^{2PA}	σ^{2PA}	f	$\mu_{01}(\text{au})$	μ_0	μ_1	$\Delta\mu$	T1	Hr:Min
CC3	4.10	2456	15.15	0.639	2.521	4.458	-	-	87.78	248:17
CCSD	4.23	2425	15.92	0.714	2.624	4.435	0.831	4.906	91.66	11:00
CC2	4.12	2205	13.72	0.722	2.675	4.269	0.891	4.642	92.75	6:44
CAM-B3LYP	4.23	313	2.05	0.663	2.530	3.472	2.115	1.465	-	0:08
BHandHLYP	4.33	260	1.79	0.701	2.569	3.298	2.048	1.364	-	0:05
B1LYP	4.17	197	1.25	0.608	2.440	3.400	1.786	1.702	-	-
B3LYP	4.14	181	1.14	0.588	2.408	3.419	1.718	1.782	-	-
BLYP	3.99	117	0.68	0.489	2.237	3.473	1.435	2.096	-	-
aug-cc-pVDZ										
method	ΔE	δ^{2PA}	σ^{2PA}	f	$\mu_{01}(\text{au})$	μ_0	μ_1	$\Delta\mu$	T1	Hr:Min
CC3	4.06	2098	12.69	0.614	2.483	-	-	4.333	87.89	1771:39
CCSD	4.20	2065	13.33	0.688	2.586	0.714	4.618	4.324	91.76	71:39
CC2	4.08	2019	12.30	0.697	2.641	0.824	4.503	4.206	92.63	50:43
CAM-B3LYP	4.18	323	2.07	0.637	2.492	2.150	1.417	3.463	-	01:40
BHandHLYP	4.29	266	1.79	0.675	2.533	2.075	1.332	3.298	-	00:54
B1LYP	4.12	210	1.30	0.584	2.405	1.841	1.630	3.383	-	-
B3LYP	4.09	195	1.20	0.564	2.372	1.778	1.705	3.398	-	-
BLYP	3.94	132	0.80	0.468	2.202	1.504	1.997	3.439	-	-

Table 3 S_1 electronic and two-photon absorption properties of PSB3 calculated using CC2, CCSD, CC3, and selected TD-DFT functionals with 6-31+G* (upper block) and aug-cc-pVDZ (lower block). Excitation energies are reported in eV, δ^{2PA} in a.u., σ^{2PA} in GM, μ_{01} in a.u., and μ_0 , μ_1 , and $\Delta\mu$ in Debye. T1 contributions and wall times are reported where available.

functional, especially considering it costs nearly two orders of magnitude less computational time than CCSD. Turning over to 2PA associated properties, the cracks of TD-DFT begins to show. Ground-state μ_0 in CC2 and CCSD hovers at about a half to 4% underestimated from CC3, meanwhile TD-DFT methods largely increase this value to the mid 20%, most visible when observing fig. 8 below. In fact, in fig. 8, the ground-state μ_0 histogram exhibits an inverse trend relative to the first row of histograms (concerning 1PA related properties) – the property decreases in a stepwise manner as HF-energy contributions increases, with the same exception as earlier of CAM-B3LYP. Among the xcfs, CAM-B3LYP and BLYP are the least underestimated at approx. 22%, while BHandHLYP remain at the extreme of TD-DFT results being the most underestimated by 26%. As noted above in the CCx section of the analyses, CC3 excited-state and therefore also its $\Delta\mu$ were unrecoverable, and so again, here the reference used will be CCSD. Conversely to the ground-state, the excited-state μ in TD-DFT are grossly overestimated and places the results in the range of 70 to 150% increased from CC3. Interestingly, excited-state μ_1 follows the same trend as in the first row of histograms, where increased HF-energy also increases μ_1 in a stepwise manner. But unlike 1PA properties, the increase in HF energies actually causes xcfs to become increasingly overestimated, with BLYP showing the smallest deviation, still approximately 73% higher than CC3. Notably, this is the first property where CAM-B3LYP lies at the extreme, exhibiting the largest overestimation, with BHandHLYP coming in as a close second. The combination of underestimated ground-state and overestimated excited-state μ yields a large reduction in $\Delta\mu$ in TD-DFT, similarly to the findings of Eriksen *et al.* [185]. As shown in Fig. 8 below, the $\Delta\mu$ histograms display the same inverse trend (relative to 1PA properties) as the ground-state μ_0 but are significantly underestimated relative to CCSD. Similarly to ground-state μ_0 , BLYP with the lowest HF-energy contribution (within this

study), shows the least underestimation at roughly 40% of CCSD. Unlike ground-state μ_0 earlier, CAM-B3LYP here actually performs second worse at 30%, just slightly ahead of BHandHLYP which is at 28% of CCSD's. Regardless of the state of μ and $\Delta\mu$, it seems that δ^{2PA} increases with increasing HF-energy in a stepwise manner – making CAM-B3LYP the least underestimated xcf at 12.7% of CC3 and BHandHLYP a close second at 10.6%, according to figure 8 below. Even with the largest increase in HF-energy contributions, TD-DFT methods are still severely underestimated. This was seemingly foreshadowed in the $\Delta\mu$ properties. As was demonstrated by Beerepoot *et al.* [131] and described in the CC analyses above, the underestimation of dipole moments and their differences can lead to an underestimation of δ^{2PA} . TD-DFT yields $\Delta\mu$ values only 28 – 43% of CCSD; in parallel, its δ^{2PA} collapses to 5 – 13% of CC3. This co-reduction is consistent with the strong $\Delta\mu$ sensitivity of δ^{2PA} , though other factors also contribute. Yet, one may wonder why BLYP, the xcf with the best performing μ and corresponding $\Delta\mu$ performed the worse in recovering CC3 δ^{2PA} strengths. Similar to earlier observations, both the transition dipole μ_{01} and the excitation energies strongly contribute to δ^{2PA} strengths. In quadratic response the two-photon tensor is effectively a weighted sum over intermediate states. Functionals like BLYP, which lack any exact HF exchange, can suffer delocalization error and incorrect long-range behavior that pull charge-transfer excitations down in energy [185, 186], misplacing and crowding the intermediate states in the 2PA manifold near half the final-state energy. With many near-resonant, comparable contributions of mixed sign, these terms cancel in the quadratic-response sum, yielding a small 2PA amplitude. By contrast, CAM-B3LYP's correct long-range exchange relocates and separates these intermediates and gives more reliable transition moments, so a few dominant channels add constructively, producing a larger δ^{2PA} amplitude [176, 187]. Even so, CAM-B3LYP here is unsurprisingly still extremely underestimated, consistent with known TD-DFT limitations for charge-transfer excitations, as was reported in single [187, 188] and double-photon [131] excitation studies by Eriksen *et al.* and others.

Checking the differences between the two available basis sets, all trends in the aug-cc-pVDZ set follow the exact same trends as in 6-31+G* with the exception of the excited-state dipole moments, μ_1 . As shown in Fig. 8, aug-cc-pVDZ increases μ_{11} by ~50% relative to 6-31+G*, which raises the overestimation versus the CCSD reference from ~170–250% to ~210–300%, making aug-cc-pVDZ more overestimated. In contrast, other properties change only marginally or remain similar in fidelity to CC3. δ^{2PA} is also no exception, yielding a modest increase in accuracy of ~1.5–2.5%, consistent with the trends in CC methods. Looking over at absolute increases, the increase in aug-cc-pVDZ's overestimation in μ_1 to the reference CCSD becomes clearer. μ_1 receives a slight total increase of 2 – 5% in DFT methods, accompanied by CCSD's decrease of 14%, exacerbating the overestimation when comparing fidelity. The only other property that also received a total increase when going from

6-31+G* to aug-cc-pVDZ is δ^{2PA} , which increased by 2 to 11%, increasing with lower exact exchange-energy. All other electronic properties in DFT experienced a decrease in value, which is in slight contrast to CC methods where all properties experienced a total decrease. The computational cost rises from ~8 mins in 6-31+G* to ~1 hour and 40 mins in aug-cc-pVDZ, which may still be reasonable to achieve a modest increase in accuracy for δ^{2PA} , but little to no change in oscillator strengths.

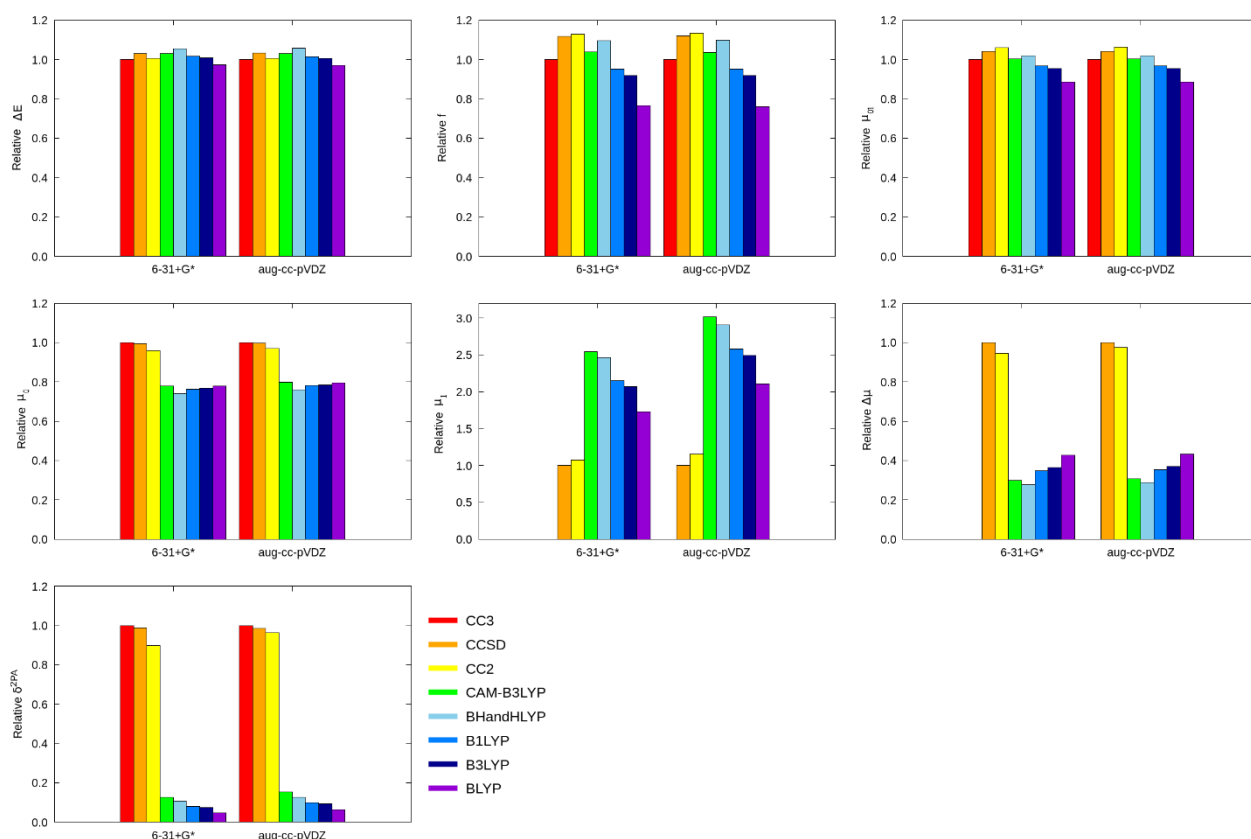


Figure 8. Relative S_1 electronic and two-photon absorption properties of PSB3 calculated using CCx and TD-DFT methods with 6-31+G* (left) and aug-cc-pVDZ (right). Within each basis set, ΔE , f , μ_{01} , μ_0 , and σ^{2PA} are normalized to the corresponding CC3 values, while μ_1 and $\Delta\mu$ are normalized to CCSD instead.

In dm-PSB4, the same trends observed in PSB3 within TD-DFT methods are again present. Increasing HF-energy raises excitation energies, 1PA properties, μ_1 , and δ^{2PA} , with the same exception as before where CAM-B3LYP lowers these values slightly relative to BHandHLYP. $\Delta\mu$ shows the inverse pattern to these properties, just as in PSB3. μ_0 diverges slightly, with all functionals nearly identical, or within 0.5% of one another, except CAM-B3LYP at ~4% higher. These trends are shown in figure 9 below. Absolute total differences follow CC methods. Extending the polyene chain by one carbon double bond red-shifts excitation energies by ~0.70 eV (17%) across all methods. Staying consistent with CC trends, most properties increase with chain extension, except excited-state dipole moments for lower HF-energy functionals (B1LYP, B3LYP, BLYP), which decrease instead.

Single-photon properties like oscillator strengths and μ_{01} increase substantially by ~ 60 and $\sim 40\%$ respectively across all functionals and CC methods. Excited-state μ_1 shows only a modest rise of 5 – 7% for CAM-B3LYP and BHandHLYP, while decreasing by 5 – 15% for the other functionals. This is contrary to CC methods’ increase in μ_1 by $\sim 25\%$. Ground-state μ_0 grows by $\sim 45\text{--}50\%$, scaling with HF-energy contributions, similar to CC’s $\sim 50\%$ rise. Accordingly, $\Delta\mu$ also increases by 80 – 110% with HF-energy in TD-DFT, compared to CC’s more conservative 18 – 21%. Despite the larger $\Delta\mu$ increases in TD-DFT, microscopic δ^{2PA} gains remain smaller, with TD-DFT demonstrating 150 – 270% increases (rising with HF-energy) versus CC’s 340 – 360%.

	ΔE	δ^{2PA}	σ^{2PA}	f	6-31+G*		μ_0	μ_1	$\Delta\mu$	T1	Hr:Min
					$\mu_{01}(\text{au})$						
CC3	3.33	12234	49.81	1.018	3.531	6.660	-	-	-	86.24	1251:38
CCSD	3.46	11065	48.40	1.127	3.649	6.652	1.036	5.945	91.02	42:26	
CC2	3.36	9611	39.65	1.142	3.727	6.225	1.118	5.466	91.90	13:18	
CAM-B3LYP	3.53	1160	5.30	1.044	3.476	5.175	2.270	2.948	-	-	
BHandHLYP	3.61	943	4.50	1.092	3.512	4.966	2.149	2.853	-	-	
B1LYP	3.49	601	2.68	0.966	3.363	4.985	1.713	3.309	-	-	
B3LYP	3.46	533	2.34	0.936	3.322	4.987	1.616	3.408	-	0:57	
BLYP	3.34	292	1.19	0.774	3.076	4.975	1.226	3.788	-	0:47	

Table 4. Electronic properties of the dm-PSB4 model computed at the CCx and TD-DFT level of methodology with the 6-31+G* b.s. for the S_1 state. ΔE is reported in eV, δ^{2PA} in a.u., σ^{2PA} in GM, μ_{01} in a.u., and μ_0 , μ_1 , and $\Delta\mu$ in Debye. T1% and computational time (where applicable) are also present.

As was in PSB3, δ^{2PA} values remain underestimated relative to CC3, and fidelity worsens in dm-PSB4. CC2 deviations grow from 10 to 21.5%, CCSD from 1.3 to 10%. TD-DFT functionals also underestimate more here, with deviations 2 – 3 percentage points larger than in PSB3, shifting from 5 – 13% of CC3 to only 2.4 – 10%, demonstrated in figure 9 below for dm-PSB4. For oscillator strengths, low HF-energy functionals (B1LYP and below) perform slightly worse by a few percentage points than in PSB3, deviating 5 – 24% from CC3, with corresponding μ_{01} deviations of 5 – 13%. BHandHLYP and CAM-B3LYP improve slightly in dm-PSB4, being ~ 1 – 2% less overestimated than in PSB3. Similar patterns are found in μ_{01} , except that all functionals are now underestimated. While BHandHLYP and CAM-B3LYP remain the most accurate, CAM-B3LYP is $\sim 2\%$ less accurate here than in PSB3. In short, functionals with more than 50% HF-energy yield deviations of only 1 – 2% in μ_{01} and 3 – 7% in oscillator strengths, outperforming CC2 and CCSD, like in PSB3. Ground-state dipole μ_0 accuracy decreases modestly (by up to 3%) when moving to dm-PSB4, improving with HF-energy up to CAM-B3LYP, which deviates by just 0.2%. Overall accuracy remains $\sim 78\%$ of CC3, similar to PSB3. On the other hand, excited-state μ_1 tells a different story. Accuracy improves by $\sim 50\%$ across the board. What was a 70 – 150% overestimation of CCSD in PSB3 is now reduced to ca. 18 – 120% in dm-PSB4, increasing with exact HF-energy contributions. This improvement also carries into $\Delta\mu$, where TD-DFT now underestimates CCSD by only 36–50% (growing with HF-

energy), as opposed to PSB3's 57 – 70%. In summary, TD-DFT functionals do deteriorate with polyene chain length extension but less so than CC methods in terms of δ^{2PA} . This is in part due to the modest absolute increase in μ_1 relative to the CC reference, allowing the reference itself to catch up to TD-DF. Consequently this also produces a larger absolute increase in $\Delta\mu$ values in TD-DFT while the CC increase remains mild, this time allowing TD-DFT to catch up to the reference. As a result, TD-DFT's overestimation of μ_{11} and underestimation of $\Delta\mu$ relative to the CCSD reference narrows, leading to a milder overall decrease in 2PA accuracy described above.

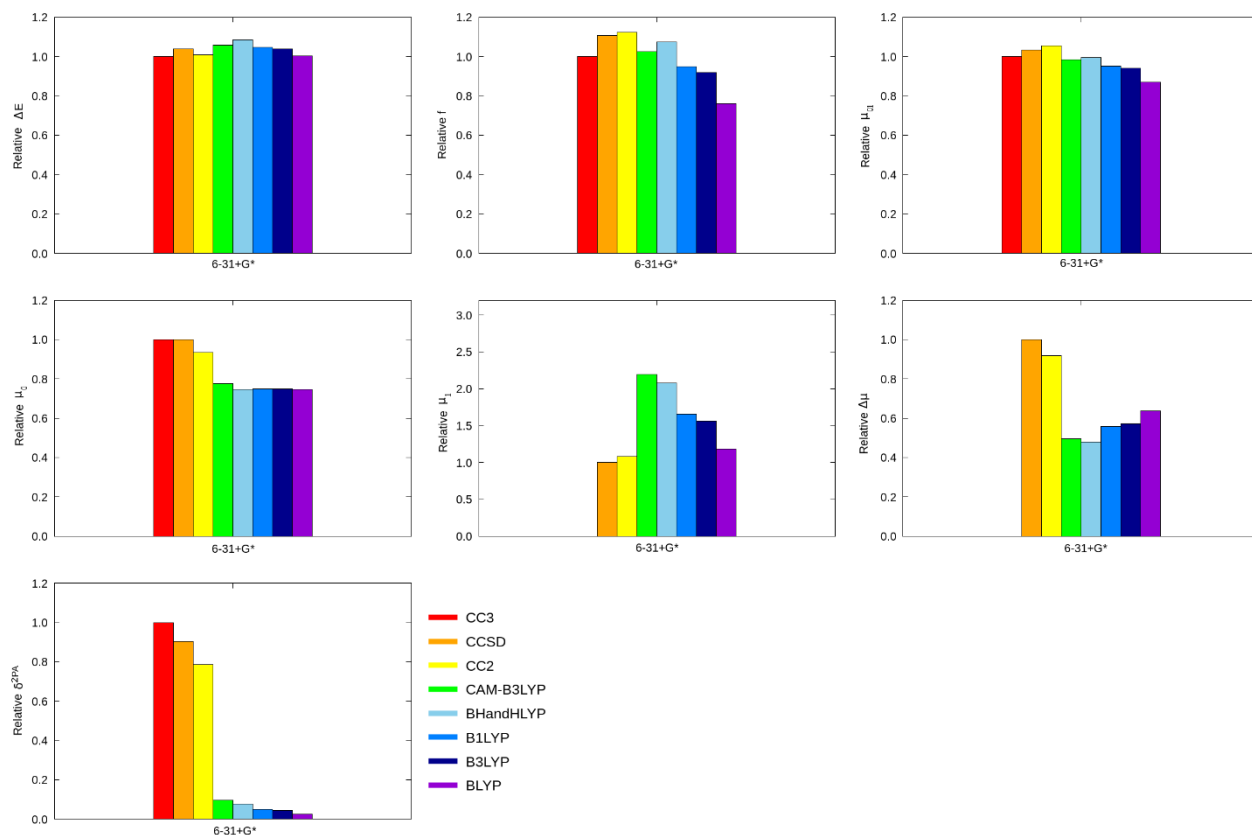


Figure 9. Relative S_1 electronic and two-photon absorption properties of dm-PSB4 calculated using CCx and TD-DFT methods with 6-31+G* b.s.. ΔE , f , μ_{01} , μ_0 , and σ^{2PA} are normalized to CC3 values, while μ_1 and $\Delta\mu$ are normalized to CCSD instead.

Restoring the methyl group to form the PSB4 model seems to positively shift all properties with the exception of excitation energies, ground-state μ_0 , and $\Delta\mu$. Across the board, all functionals display a decrease in excitation energies of less than 0.03 eV, or less than a percent reduction, contrary to CC3/CCSD's increase in energy levels. With increasing HF-energy contributions the reduction drops to 0.005 eV in CAM-B3LYP and 0.008 eV in BHandHLYP. Combined with the increase in energies in the reference, this brings the accuracy of TD-DFT in recovering excitation energies of CC3 up by half a percent across all functionals from its de-methylated counterpart. μ_0 saw a reduction of about 0.09 (2%) to 0.04 D, again decreasing with HF-energy contributions. Meanwhile excited-state μ_1

experienced an increase of 0.01 to 0.15 D, increasing with HF-energy as well. Similar to the previous two models, the total property shifts from the methyl substitution are less pronounced in CAM-B3LYP than for BHandHLYP in certain properties. Across all properties, the magnitude of change between dm-PSB4 and PSB4 either increases or decreases with the amount of HF-energy contributions. CAM-B3LYP consistently lies at the extremes, exhibiting either the largest or smallest overall shifts in these electronic properties. Previously, the shift from PSB3 to dm-PSB4 follows this same trend, with exceptions in oscillator strengths and μ_{01} , which show no clear pattern, as well as in ground-state μ_0 and $\Delta\mu$, where BHandHLYP lies at the extreme. Here in PSB4 the ground and excited-state μ , and consequently its $\Delta\mu$, are affected by this exception, causing BHandHLYP to be at the extremes.

When checking fidelity to the CC3 reference, CAM-B3LYP breaks from the expected increase in HF-energy trend, as it did in the previous models, for excitation energies, oscillator strengths, transition dipole moment μ_{01} , and $\Delta\mu$. In fact, for all the aforementioned properties in all models, CAM-B3LYP deviated from the trend of increasing HF-energy also increasing those properties, and instead becoming slightly less intense than its lower HF-energy counterpart; BHandHLYP. The only other case across all models/properties is μ_0 in PSB3, where CAM-B3LYP had also deviated from the trend in terms of fidelity to the reference, in all other cases CAM-B3LYP is more intense than BHandHLYP.

As expected TD-DFT's ability to recover excitation energies in PSB4 follows the same trend as the other models, increasing with HF-energy contributions. This results in BHandHLYP being the most deviated at 8% overestimated while BLYP being the most accurate, yet again, at sub 1% underestimated. It seems that with the larger models, all functionals increase in deviation by 2-3 percentage points with the exception of BLYP that actually improved by the same amount instead. On the other hand, the addition of a methyl group brings oscillator strengths of lower HF-energy functionals up in accuracy by one percentage point, while CAM-B3LYP's accuracy remains the same and BHandHLYP decreased in accuracy by a percent. Overall the accuracy of oscillator strengths scale well with model size, with CAM-B3LYP consistently performing the best out of all functionals. μ_{01} is slightly more complicated, previously in dm-PSB4 all functionals performed worse when extending the polyene chain, although only by one to two percentage points, with the exception of BHandHLYP that actually performed better by a percent. Restoring the methyl group to PSB4 improves all functionals again by one to two percentage points, bringing all functionals back to a similar accuracy as was in PSB3. In summary, some functionals perform better in the smaller models, others in the larger ones, but overall there is no systematic trend of decreasing or improving accuracy

with increasing the polyene chain in μ_{01} . The only notable exception being BHandHLYP that performed better but by merely one percent for each step in the model size increase. Ground-state μ_0 also display no discernible trends. Accuracy remains similar to the other two models, hovering at an underestimation of 22 – 26%, with CAM-B3LYP being the least underestimated. Functionals with lower (30% or less) HF-energies decrease in accuracy by ca. one to two percentage point, BLYP being hit the hardest by a decrease in accuracy of 4% in going from PSB3 to PSB4. The other two functionals either remained just as accurate as in smaller structures, like in CAM-B3LYP, or slightly more accurate by merely a single percent, like in BHandHLYP. The excited-state μ_1 is where large differences begin to appear. Earlier in dm-PSB4, a large increase in accuracy was observed. Here in PSB4, the accuracy has dropped back to the higher overestimation seen in PSB3. The difference here is that for functionals with HF-energy of 30% or lower (i.e. B1LYP and below), the accuracy did decrease by 20 – 30% from its demethylated counterpart, yet its accuracy remains higher than that of PSB3 by 16 – 36%. In fact, the functional with no HF-energy at all, BLYP, is merely 36% overestimated from CCSD in the PSB4 model. This makes BLYP twice as inaccurate as dm-PSB4 (18% overestimated), but also nearly twice as accurate as PSB3 (73% overestimated). On the other side are CAM-B3LYP and BHandHLYP that actually decreased in accuracy not just from dm-PSB4, but from PSB3 as well. Overall these increased HF-energy functionals decreased in accuracy by nearly 50% from dm-PSB4, and over 9% from PSB3 to a final overestimation of 155 – 165%, demonstrating their rather low tolerance to scaling of the model size. The likely culprit here can be seen in table 5 below, where the absolute increases from dm-PSB4 to PSB4 results in CC methods lowering their μ_1 values by 12% (CCSD) and 36% (CC2). At the same time all TD-DFT methods increased their μ_1 , as noted earlier within this PSB4 section. These absolute increases are small and constitute approximately half a percent to over 7%, increasing with HF-energy contributions but ending with BHandHLYP claiming the largest magnitude in shifts. The combined effect of the reduced CC μ_1 reference and the elevated TD-DFT results lead to an increased degree of overestimation. Interestingly, $\Delta\mu$ didn't actually decrease in accuracy by much compared to the demethylated variant, only by one to two percentage points across the board. This also means that, relative to PSB3, all functionals maintain good $\Delta\mu$ scaling, showing only minimal error increases with methyl addition and even an approximate 20% improvement in $\Delta\mu$ accuracy with polyene chain extension. The minimal drop in $\Delta\mu$ accuracy, despite the large decrease in excited-state μ fidelity, stems from the opposing nature of the errors in μ_0 and μ_1 . More importantly, the good $\Delta\mu$ performance across the larger structures primarily originates from the increase in ground-state dipole moments in the extension of the polyene chain. In going from PSB3 to dm-PSB4, CC methods show roughly a 20% rise in excited-state μ , while TD-DFT increases by only about 7% in CAM-B3LYP and less in the lower HF-energy functionals, as mentioned above. In contrast, the ground-state μ_0 increases by

nearly 50% across all functionals, effectively doubling the $\Delta\mu$ magnitude. This compounds with the error signs between μ_0 and μ_1 becoming less destructive and more additive, further stabilizing $\Delta\mu$ accuracy overall.

When it comes to δ^{2PA} in PSB4, TD-DFT methods are expectedly still very underestimated at 3 – 12% of the CC3 reference, increasing with HF-energy contributions. When compared to dm-PSB4, here δ^{2PA} improved by 1 – 3 percentage points, with the largest increases found in functionals with larger HF-energy contributions. On the other hand, compared to PSB3 and all functionals experience a decrease in accuracy by one to two percentage points across the board. It seems that the combined offset between the overestimated excited-state μ and the underestimated ground-state μ_0 , together with the improved excitation energies and stable μ_{01} performance, underlies the steady δ^{2PA} behavior across models. The 20% improvement in $\Delta\mu$ accuracy from PSB3 to PSB4 further supports this, as does the relative consistency of μ_0 , which varies by only a few percentage points across all models. Excitation energies improve slightly by about 2 – 3 percentage points from dm-PSB4, though they remain 2 – 3 points less accurate than in PSB3. Similarly, transition dipoles remain comparatively stable, worsening marginally by 2 – 3 percentage points from PSB3 to dm-PSB4 (with BHandHLYP improving by 1 point) and recovering by one to two points in PSB4. Altogether, this balance explains why δ^{2PA} accuracy remains minorly affected by the increase in model size, despite the significant deterioration in excited-state μ .

6-31+G*										
	ΔE	δ^{2PA}	σ^{2PA}	f	$\mu_{01}(\text{au})$	μ_0	μ_1	$\Delta\mu$	T1	Hr:Min
CC3	3.34	11877	48.52	1.048	3.580	6.568	-	-	86.58	1706:02
CCSD	3.47	10679	47.01	1.158	3.692	6.578	0.908	5.809	91.33	15:03
CC2	3.33	10796	43.84	1.175	3.795	6.148	0.713	5.631	91.84	24:33
CAM-B3LYP	3.52	1390	6.31	1.080	3.538	5.138	2.406	2.779	-	1:54
BHandHLYP	3.61	1180	5.63	1.132	3.579	4.934	2.302	2.688	-	1:32
B1LYP	3.47	753	3.32	1.008	3.444	4.926	1.810	3.150	-	1:24
B3LYP	3.44	668	2.90	0.978	3.405	4.922	1.701	3.251	-	1:23
BLYP	3.32	360	1.45	0.817	3.171	4.887	1.234	3.667	-	1:28
aug-cc-pVDZ										
	ΔE	δ^{2PA}	σ^{2PA}	f	$\mu_{01}(\text{au})$	μ_0	μ_1	$\Delta\mu$	T1	Hr:Min
CAM-B3LYP	3.501	1390	6.24	1.059	3.514	5.113	2.406	2.745	-	12:07
BHandHLYP	3.586	1170	5.52	1.110	3.555	4.920	2.298	2.669	-	8:12
B1LYP	3.446	788	3.44	0.990	3.424	4.897	1.851	3.072	-	7:33
B3LYP	3.419	707	3.03	0.960	3.386	4.889	1.749	3.164	-	7:32
BLYP	3.291	404	1.60	0.802	3.154	4.843	1.311	3.541	-	6:53

Table 5. Electronic properties of the PSB4 model computed at the CCx and TD-DFT level of methodology with the 6-31+G* (top) and aug-cc-pVDZ (bottom) b.s. for the S_1 state. ΔE is reported in eV, δ^{2PA} in a.u., σ^{2PA} in GM, μ_{01} in a.u., and μ_0 , μ_1 , and $\Delta\mu$ in Debye. T1% and computational time (where applicable) are also present.

Despite the lack of CC reference in aug-cc-pVDZ, one can still compare the absolute differences in this larger PSB4 model when switching basis sets. Most absolute increases remain the same as what

was seen in PSB3, where a reduction of usually less than a percent but in some cases up till 2 percent for all electronic properties except δ^{2PA} and μ_1 can be seen when switching over to the larger Dunning's b.s. Similarly to PSB3, δ^{2PA} and μ_1 here remained mostly the same or slightly increased up to 12% in lower exchange energy functionals. At the same time, μ_0 in PSB4 remains largely unchanged between the two basis sets for all investigated functionals. Similarly to PSB3, CAM-B3LYP displays the smallest difference between the two basis sets. This demonstrates that basis-set sensitivity is largely size-invariant when scaling the model, particularly for CAM-B3LYP, while μ_0 remains comparatively insensitive to the basis-set expansion.

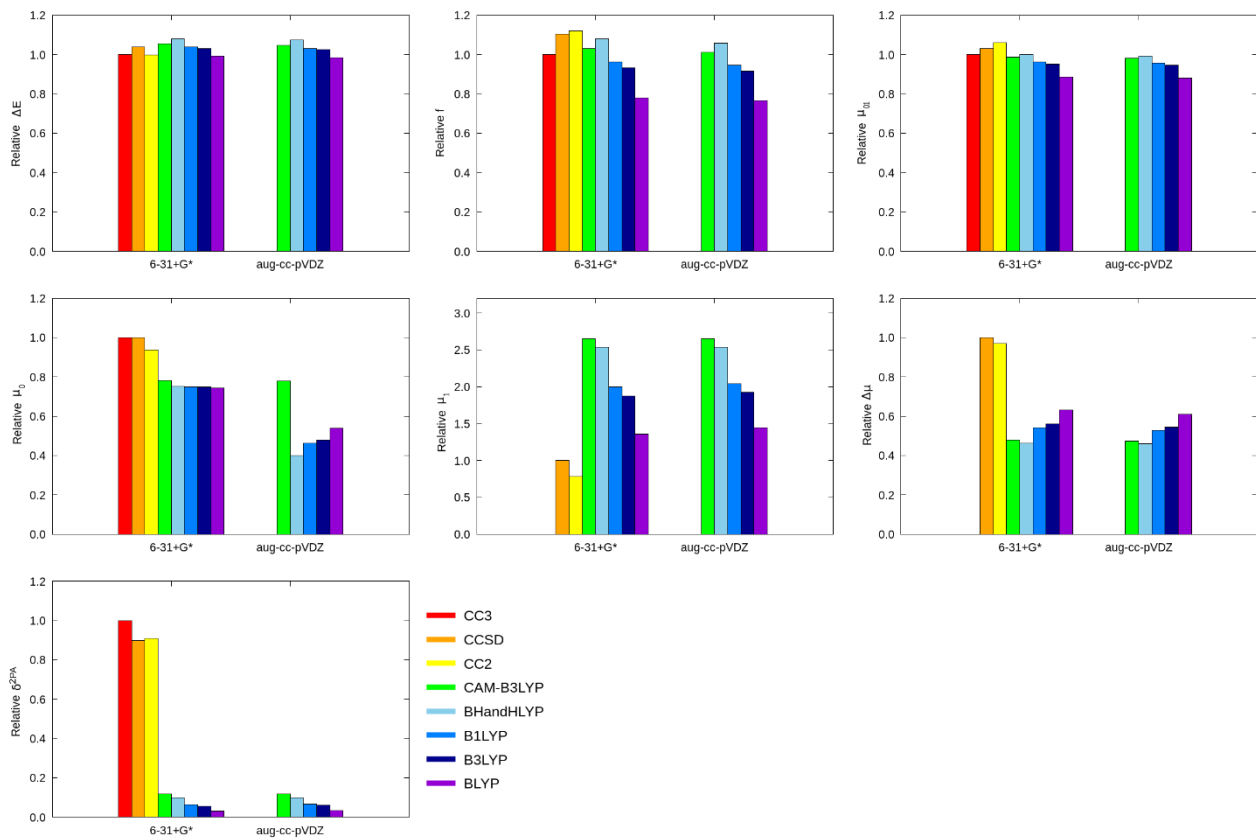


Figure 10. Relative S₁ electronic and two-photon absorption properties of PSB4 calculated using CCx and TD-DFT methods with 6-31+G* b.s.. ΔE , f , μ_{01} , μ_0 , and σ^{2PA} are normalized to CC3 values, while μ_1 and $\Delta\mu$ are normalized to CCSD instead.

Overall, the scaling from PSB3 to dm-PSB4 and PSB4 reveals that δ^{2PA} accuracy in TD-DFT is governed by compensating trends among the underlying electronic properties. The systematic underestimation of μ_{00} and overestimation of μ_{11} balance to produce a comparatively stable $\Delta\mu$ across models, meanwhile modest variations in excitation energies and μ_{01} further accomodate deviations in δ^{2PA} . Among the functionals, CAM-B3LYP and BHandHLYP scale most consistently; their higher exact HF-exchange energy helps preserve excitation energies and transition dipoles, while opposing deviations in μ_{00} and μ_{11} offset one another to stabilize $\Delta\mu$ and, consequently, δ^{2PA} fidelity. This internal error cancellation maintains δ^{2PA} performance nearly constant with increasing model size,

highlighting the relative robustness of these two hybrid functionals. Lower-exchange functionals such as BLYP may occasionally yield good excitation energies, but this comes at the expense of accuracy in $\Delta\mu$, δ^{2PA} , and oscillator strengths.

4.1.5 Higher Excited-states in TD-DFT

CC3							BhandHLYP						
states	ΔE	δ^{2PA}	σ^{2PA}	f	μ	$\Delta\mu$	states	ΔE	δ^{2PA}	σ^{2PA}	f	μ	$\Delta\mu$
ground	-	-	-	-	4.458	-	ground	-	-	-	-	3.298	-
1	4.1	2456	15.15	0.639	-	-	1	4.33	260	1.79	0.701	2.048	1.364
2	5.64	121	1.41	0.124	-	-	2	6.53	1220	19.06	0.142	0.878	2.499
3	6.49	2	0.02	0.000	-	-	3	6.57	1	0.02	0.001	2.799	0.722
4	6.77	2113	35.51	0.135	-	-	4	6.84	255	4.37	0.143	0.111	3.343
-	-	-	-	-	-	-	5	7.05	3	0.06	0.001	10.214	7.106
-	-	-	-	-	-	-	6	7.36	46	0.92	0.002	6.133	4.202
-	-	-	-	-	-	-	7	7.44	0	0.01	0.000	2.076	1.278
CAM-B3LYP							B3LYP						
states	ΔE	δ^{2PA}	σ^{2PA}	f	μ	$\Delta\mu$	states	ΔE	δ^{2PA}	σ^{2PA}	f	μ	$\Delta\mu$
ground	-	-	-	-	3.472	-	ground	-	-	-	-	3.419	-
1	4.23	313	2.05	0.663	2.115	1.465	1	4.14	181	1.14	0.588	1.718	1.782
2	6.14	1	0.02	0.000	2.841	0.822	2	5.79	1	0.01	0	3.163	0.802
3	6.43	861	13.04	0.133	0.939	2.618	3	5.83	168	2.09	0.194	0.604	3.330
4	6.77	151	2.54	0.151	0.219	3.686	4	6.52	5	0.08	0.001	11.158	7.815
5	7.04	3	0.05	0.001	8.329	5.129	5	6.65	98	1.59	0.12	0.055	3.432
6	7.05	2	0.03	0.000	3.672	0.430	6	6.74	0	0.01	0	2.521	1.051
7	7.33	67	1.31	0.002	6.244	4.219	7	6.92	50	0.87	0.001	6.104	3.424
B1LYP							BLYP						
states	ΔE	δ^{2PA}	σ^{2PA}	f	μ	$\Delta\mu$	states	ΔE	δ^{2PA}	σ^{2PA}	f	μ	$\Delta\mu$
ground	-	-	-	-	3.400	-	ground	-	-	-	-	3.473	-
1	4.17	197	1.25	0.608	1.786	1.702	1	3.99	117	0.68	0.489	1.435	2.096
2	5.94	3	0.04	0.002	3.058	0.785	2	5.21	1	0.01	0	3.559	0.946
3	5.95	245	3.18	0.186	0.601	3.229	3	5.31	38	0.4	0.238	0.626	3.594
4	6.56	4	0.07	0.001	11.053	7.740	4	6.01	9	0.12	0.001	11.719	8.270
5	6.67	109	1.78	0.122	0.069	3.359	5	6.18	0	0.01	0.000	2.856	0.948
6	6.88	0	0.00	0.000	2.433	1.091	6	6.42	0	0	0.001	2.082	1.532
7	6.95	48	0.85	0.001	6.104	3.518	7	6.47	64	0.97	0.106	0.472	3.942
-	-	-	-	-	-	-	8	6.51	35	0.54	0.001	6.170	3.070
-	-	-	-	-	-	-	9	7.07	11900	217.92	0.008	0.722	3.108

Table 6. Simulated one- and two-photon absorption spectra of PSB3 for all available excited states calculated using CC3 and selected TD-DFT functionals with the 6-31+G* basis set. ΔE is reported in eV, δ^{2PA} in a.u., σ^{2PA} in GM, μ_{01} in a.u., and μ_0 , μ_1 , and $\Delta\mu$ in Debye. T1% and computational time (where applicable) are also present.

Exploring TD-DFT's higher excited-state results allows one to determine how the distribution of bright and dark computed states compares with the CC3 reference. TD-DFT, limited by its adiabatic single-excitation framework, cannot properly capture states with significant double-excitation character, which may affect the relative energies and intensities of higher excited states. Furthermore, most xcfs suffer from self-interaction error (SIE), where electrons spuriously interact with themselves due to imperfect exchange cancellation. The functionals compensates for this artificial self-repulsion by allowing its electron density to become overly diffuse, inadvertently introducing delocalization errors [189, 190]. Together, these effects can distort the energetic spacing of nearby states and redistribute their calculated intensities.

The following analysis therefore focuses on how excitation energies and intensities are distributed among the computed roots, considering states with non-negligible δ^{2PA} or oscillator strengths. Comparison with CC3 reveals clear method-dependent differences in the placement of bright and

dark responses. In CC3, the dominant transitions are the computed S_1 and S_4 states, while S_2 contributes moderately to 1PA but only weakly to 2PA and S_3 is essentially dark in both. In BHandHLYP, the dominant higher-energy 2PA response instead appears at computed S_2 , while in CAM-B3LYP it appears at computed S_3 . This places the strongest higher-energy response at different computed roots than in CC3, but does not by itself establish a perfect correspondence or inversion of state character. BHandHLYP therefore produces a stronger higher-state δ^{2PA} response than CAM-B3LYP, while its S_1 response is slightly weaker. In terms of oscillator strengths, BHandHLYP reproduces the CC3 ordering of the first three computed states, $S_1 > S_2 > S_3$, while CAM-B3LYP instead gives $S_1 > S_3 > S_2$. In BHandHLYP, as in CC3, S_2 remains moderately bright while S_3 is nearly dark. CAM-B3LYP reverses this behavior, predicting a dark S_2 state and a brighter S_3 state. Comparing the same computed states directly, the S_3 and S_4 excitation energies obtained with both high-exchange functionals remain within approximately 1% of CC3, whereas S_2 is overestimated by approximately 9% with CAM-B3LYP and 16% with BHandHLYP. The lower-exchange functionals, including B1LYP, B3LYP, and BLYP, also generally give the ordering $S_1 > S_3 > S_2$, with the next bright higher-state feature appearing at S_5 for B1LYP and B3LYP and at S_7 for BLYP. Their S_3 excitation energies become increasingly underestimated as the HF-exchange contribution decreases, from approximately 8% in B1LYP to 18% in BLYP, while their 2PA strengths remain substantially below the CC3 values. BLYP additionally produces an unusually intense computed S_9 response of approximately 218 GM at 7.07 eV, which should be treated cautiously in the absence of further state-character analysis. These differences may reflect changes in state character and mixing across the methods, particularly for states with greater double-excitation contributions, such as the computed CC3 S_2 state. Several bright higher excited states across the TD-DFT functionals also exhibit relatively large $\Delta\mu$ values of approximately 2.5 – 4 D, suggesting charge-transfer character. Since TD-DFT is known to struggle with charge-transfer states because of self-interaction and delocalization errors, these effects may contribute to the method-dependent energetic shifts and redistribution of intensity observed here [185, 191, 192].

Within this study, aug-cc-pVDZ results for the higher excited states at the CC3 level were unrecoverable; however, the available TD-DFT results still provide insight into basis-set sensitivity within each functional. The complete higher-state aug-cc-pVDZ results are provided in Table S1. Excitation energies generally decrease by approximately 0.5 – 2.5% when switching from 6-31+G* to aug-cc-pVDZ, with only a few larger deviations. In contrast, δ^{2PA} is considerably more state dependent. For CAM-B3LYP, B1LYP, B3LYP, and BLYP, the computed S_3 response increases by approximately 7 – 13%, while the computed S_2 response decreases or remains negligible. More substantial redistribution is observed for BHandHLYP and B1LYP. In BHandHLYP, the dominant

higher-state δ^{2PA} response appears at computed S_2 with 6-31+G* but at computed S_3 with aug-cc-pVDZ. For B1LYP, the secondary bright higher-state response shifts from computed S_5 to S_4 . Same-root percentage comparisons for these cases should therefore be treated cautiously, since the basis-set change alters the computed root at which the intensity is concentrated. Oscillator strengths generally show a small decrease for S_1 and a modest increase for S_3 , while the permanent dipole moments remain strongly dependent on both the state and functional. Overall, 6-31+G* reproduces the excitation-energy trends reasonably well, but the higher-state δ^{2PA} values and distribution of intensity among the computed roots remain more sensitive to basis-set choice.

4.2 Tuning δ^{2PA} in 11-cis rPSB via De/Methylation Patterns: Benchmarks in RI-CC2 and TD-DFT Validation

4.2.1 Evidence of 2PA Tuning in rPSB

In this sub-chapter, the focus of study lies within determining a systematic way to tune δ^{2PA} transition strengths and its associated σ^{2PA} cross-sections of the 11-cis rPSB *in vacuo* via modulating the methyl group patterns along its polyene chain, which pertains to objective no. 2. The full and numbered structure of the 11-cis rPSB (rPSB5) can be seen below in figure 11 below. As noted in the introduction, the Rhodopsin protein contains a flexible enough binding pocket that can house modified retinal chromophore to form artificial visual pigments [77–81]. Some of these artificial visual pigments are distinguished from the native chromophore by only a deletion or insertion of a singular methyl group along retinal’s polyene chain. These modifications can directly affect the excited-state lifetime, quantum yield, and stability of the Rhodopsin photointermediates during photoisomerization [79, 84, 126, 193–196]. More specifically, in a femtosecond spectroscopic experiment by Wang and co-workers [79], the team identified the nonbonding interactions between the C13 methyl and C10 hydrogen to be the key interaction in the ultrafast photoisomerization mechanism of Rhodopsin. Accordingly, this sub-chapter examines de/methylation at C9, C10, and C13, using four analogues: 9-methyl, 13-methyl, 9,10-dimethyl, and 9,10,13-trimethyl. Additionally, two of these analogues (9,10-dimethyl and 9,10,13-trimethyl) as noted in the introduction, have equilibrium structures in both a planar and twisted conformation as shown by Walczak *et al.* [92]. Both conformers are included, bringing the total to 7 different structures analyzed here.

Building upon the results of 4.1, resolution-of-the-identity (RI) CC2 was employed here for the main investigation into spectral tuning on the models in fig. 11 below. In the previous benchmark study, CC2/6-31+G* consistently displayed accurate excitation energies with deviations $\leq 1\%$ in all model

sizes, while δ^{2PA} increases this value to $\sim 10\%$ in PSB3 and $\sim 9\%$ in PSB4. The aug-cc-pVDZ b.s. further improves the fidelity of both properties to 0.34% overestimated in energies and merely 4% underestimated in δ^{2PA} for the PSB3 structure. This demonstrates CC2's capacity in capturing structure–property trends across increasingly large model sizes, while also providing quantitatively competitive δ^{2PA} at lower computational costs. To this end, the RI-CC2 with Dunning's aug-cc-pVDZ b.s. is used here to investigate excitation energies, 2PA strengths and cross-sections, oscillator strengths, and dipole moments of the aforementioned rPSB and its analogues. The first goal (corresponds to objective no. 2 of this thesis) is to determine the impact of each methyl group along the polyene chain of RPSB5 on the 2PA properties to establish a rational design process for generating retinals with desired 2PA properties. In addition, this analysis also directly supports objective 3, as the observed changes in the 2PA signal are also interpreted through their relationships to the underlying relevant excited-state properties.

Furthermore, the two best functionals from 4.1 (CAM-B3LYP and BHandHLYP) are further benchmarked along with two new functionals; global-hybrid meta-NGA, MN15, and range-separated hybrid meta-GGA, M11. These two new functionals align with recent benchmarks showing that functionals beyond-GGA (meta-GGA and range-separated/hybrid meta-GGA) provide more reliable 2PA strengths and excited-state dipoles than conventional hybrids. A few studies have identified M06-2X, MN15, ω B97X/ ω B97X-V, M11, and QTP-type hybrids as top performers relative to RI-CC2 and even ahead of CAM-B3LYP for coumarin-based and other donor–acceptor chromophores. [197–199]. In fact, M11 was identified to be among the top overall performers in recovering σ^{2PA} with the lowest MAE of all functionals at 30.10 GM and the highest coefficient of determination, $R^2=0.960$ [199]. And therefore, the second goal is to determine whether MN15 and M11 provide a transferable improvement, over CAM-B3LYP or BHandHLYP, in recovering the electronic properties of interest, making this goal an expansion on objective no. 1.

4.2.2 Methodology

The 11-cis rPSB models studied within this sub-chapter has been reduced at the β -ionone ring, similar to the previous sub-chapter, to increase computational efficiency. All models were adapted from Walczak *et al.* [92] where geometry optimization was performed at the CASPT2 level containing 10 electrons in 10 π -orbitals within its active space, along with the ANO-L-VDZP basis set. As noted earlier, the 9,10-dimethyl and 9,10,13-trimethyl models exhibit a twisted counterpart due to the close proximity of its methyl group in its polyene chain causing steric hindrance. The resulting ground-state equilibrium structures therefore also displayed two separate global minima each, exhibiting a torsional deformation at the C10-C11 and C11=C12 bonds. For 9,10-dimethyl (twist), a torsion of

-167.9° and $+9.9^\circ$ was observed at these bonds with a 0.03 kcal/mol higher stability than its planar counterpart. While 9,10,13-trimethyl (twist), a -159.6° and $+27.0^\circ$ twist was seen for the same bonds with a much higher stability of 9.31 kcal/mol to its planar counterpart.

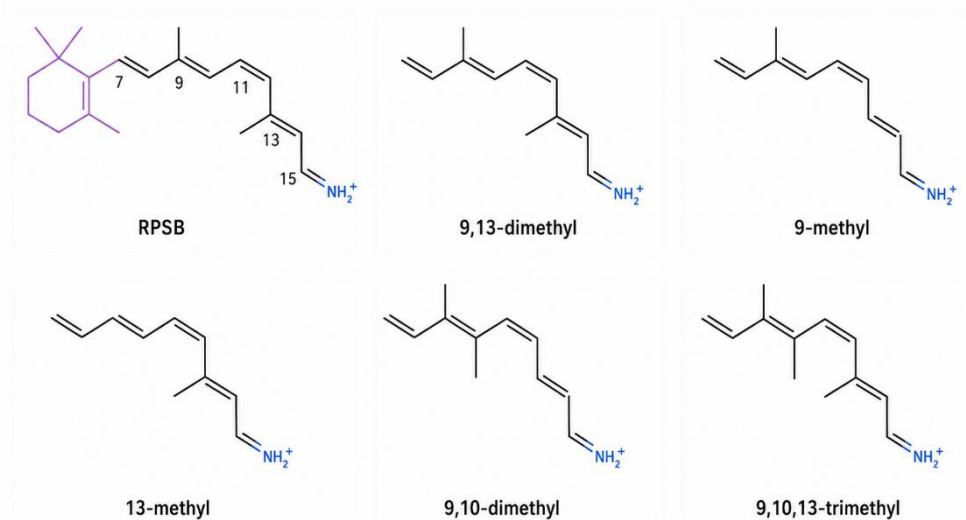


Figure 11. Native (9,13-demethyl) rPSB chromophore along with its truncated model (at the β -ionone ring) and de/methylated analogues.

All electronic properties including excitation energies, 1PA, 2PA, transition/permanent electric dipole moments and their differences were computed at the RI-CC2 and TD-DFT level. Functionals used in TD-DFT includes; the global-hybrid GGA functional BHandHLYP (50%), range-separated hybrid GGA functional CAM-B3LYP (19–65%), global-hybrid meta-nonseparable generalized approximation (meta-NGA) functional MN15 (44%), and range-separated hybrid meta-GGA (meta-GGA) functional M11 (44–100%). Included in the parentheses are the contributions from the exact HF-exchange energy. Both RI-CC2 and TD-DFT calculations with MN15/M11 functionals, both functionals taken from LibXC 5.0.0 library, were computed with Turbomole 2020 version 7.6.0. [200, 201] TD-DFT with the CAM-B3LYP and BHandHLYP functionals were computed with Dalton2020 [182, 183]. All calculations used the aug-cc-pVDZ basis set. In M11 and MN15 functionals, 1PA properties were computed using gauge-invariant current-dependent formalism, which was not implemented for 2PA properties yet, and thus δ^{2PA} are gauge-variant. All electronic properties were obtained for the 5 lowest excited-states, however, the discussion here will be mainly focused on the S_1 state. In the higher-lying excited-states, the main property of interest – 2PA strengths/cross-sections were either unrecoverable at the CC2 level of theory or suffer from resonance enhancement with the S_1 state in CC2 and TD-DFT. Additionally, to increase computational efficiency, frozen core approximation was applied to all RI-CC2 calculations, freezing a total of 11 core orbitals for the single methylated structures (9-methyl and 13-methyl), 12 core orbitals for the doubly methylated

structures (9,13-dimethyl and 9,10-dimethyl), and 13 core orbitals for triply methylated structures (9,10,13-trimethyl).

4.2.3 De/methylation Impact on Spectral Properties at the RI-CC2 level for the S_1 state

The S_1 singlet state is of π - π^* nature and has a significant charge-transfer character as seen from the large change in dipole moments below. It is dominated by a single-reference wavefunction where contributions from double-excitation are less than 12.06% according to the T1/T2 diagnostics, also seen in table 7 below. The findings related to 1PA properties and excitation energies has already been studied before by Walczak and co [92]. Nevertheless, the results below can help confirm trends established by the group as well as compare the accuracy of RI-CC2 to CASPT2 at the CASPT2 geometry optimization level. When comparing excitation energies and 1PA properties in CC2, in accordance to the table 7 below, to the same geometry optimized CASPT2 with CASPT2 excitations of Walczak most trends remain true. Excitation energies shift in the same red/blue direction with similar magnitudes across all model structures. $\Delta\mu$ trends are also agreeable to Walczak's with the only exception being the 13-methyl structure within this study increases in magnitude compared to the wild-type (9,13-dimethyl), while Walczak's results demonstrated a decrease. Oscillator strengths on the other hand agree only for the 9-methyl, 9,10-dimethyl(C_s), and 9,10,13-trimethyl structures in terms of direction but generally at slightly differing magnitudes. For 13-methyl, 9,10-dimethyl(twist), and 9,10,13-trimethyl(C_s), the CC2 trends here take the opposite direction to Walczak's CASPT2. Examining the accuracy of single-reference CC2 to that of multiconfigurational and multireference 2-root SA-CASSCF/CASPT2 one can see the gap in accuracy between the methods. Generally, excitation energies are extremely accurate and off by up to 1.50%. Oscillator strengths despite the difference in direction in some structures deviate by as little less than half a percent like in the wild-type and 9-methyl structures up to nearly 13% in 9,10,13-trimethyl(twist). Lastly, $\Delta\mu$ suffers the worse fate, averaging with a deviation of nearly 42%, or about 28 – 48% lower across the board. These results directly parallels Walczak's for the same method comparison but in MP2 geometries, with a notable difference in oscillator strengths here having a slightly higher max deviation.

Model	ΔE	λ	δ^{2PA}	σ^{2PA}	f	μ_{01}	μ_0	μ_1	$\Delta\mu$	T1
WT	2.72	910	21307	58.0	1.535	12.188	7.837	2.796	5.100	88.37
9meth	2.67	929	18170	47.4	1.382	11.684	8.102	3.663	4.818	88.57
13meth	2.80	885	27544	79.1	1.567	12.148	7.044	1.207	5.922	88.34
9,10-dimeth(C _s)	2.73	909	15097	41.1	1.587	12.388	8.312	4.188	4.261	88.37
9,10-dimeth(twist)	2.68	925	11451	30.1	1.488	12.102	7.991	4.424	3.752	88.48
9,10,13-trimeth(C _s)	2.60	953	31434	78.0	1.478	12.238	8.527	2.677	5.865	87.94
9,10,13-trimeth(twist)	2.51	990	16752	38.5	1.221	11.337	7.625	3.258	4.506	88.24

Table 7. Ground and S₁ excited-state electronic and two-photon absorption properties of all RPSB5 models calculated at the RI-CC2/aug-cc-pVDZ level of theory. ΔE is reported in eV, λ in nm, δ^{2PA} in a.u., σ^{2PA} in GM, and μ_{01} , μ_0 , μ_1 , and $\Delta\mu$ in Debye. The oscillator strength, f , is dimensionless, while T₁ single-excitation contribution is reported as a percentage.

The impact of structural changes on δ^{2PA} can also clearly be observed in table 7 above and figure 12 below. The results point to a proclivity for the planar analogs containing the C13 methyl group, similarly to what was observed by Wang and co-workers [79]. Starting the removal of the C9 methyl from the wild-type to produce the 13-methyl analogue results in an enhancement of δ^{2PA} by about 30%, going from approximately 21300 a.u. to 27500 a.u. At the same time, one can also consider that the C9 methyl group is seen as dampening δ^{2PA} as opposed to only C13 enhancing it. In reverse, removing the C13 methyl group to form the 9-methyl analogue from the wild-type leads to a reduction of over 3000 a.u. in δ^{2PA} . This results in a final δ^{2PA} of about 18000 a.u. and points to a direct enhancement by the C13 methyl. Yet, this is only half the overall change than just removing the C9 methyl which may hint at both methyl groups contributing in opposite directions. Replacing the C13-methyl group with the C10 one producing the 9,10-dimethyl structure instead drops the 2PA signal also by approximately 30% down to about 15000 a.u., among the lowest 2PA signals within this study. Restoring the C13 methyl to form the trimethyl structure enhances the δ^{2PA} back up to over 31000 a.u., a dramatic increase of over two times compared to its dimethyl counterpart. Interestingly, compared to the wild-type, this trimethyl planar structure represents a 50% increase and has the highest δ^{2PA} intensity within this study. On the other hand are the structures with the weakest 2PA signals. As noted above, 9,10-dimethyl(C_s) demonstrates the second weakest δ^{2PA} , the other structures that also demonstrate similar or even weaker δ^{2PA} signal belong to the twisted structures. In fact, when going from the planar structures to their twisted counterparts, a reduction of over 25% and 50% can be seen in the 9,10-dimethyl and 9,10,13-trimethyl structures respectively. It is clear from the above results that disruption in the planarity of the PSB structure also greatly dampens 2PA signal, even more so than simply lacking the C13 methyl group. This places the structures without the C13 methyl group and/or a twist in their polyene chain as the weakest in terms of 2PA intensity and worse than the wild-type, ergo placing 9,10-dimethyl(twist) as the weakest in the pack.

Similarly to 4.1, investigating the other properties that contribute to an enhanced 2PA signal in both microscopic δ 's and macroscopic σ 's can help further elucidate the pattern within the above enhancements. Again here, when removing the C13 methyl group from the wild-type to produce the 9-methyl analogue, along with the reduced 2PA signal, nearly all other properties also shift unfavorably compared to the wild-type. The attenuation of excitation energies ($\sim 2\%$ reduced) may make one assume that with two equal degenerate photons, reduction of excitation energies could help push δ^{2PA} higher, as seen in the quadratic response formalism. However, for the case of 9-methyl this was not sufficient enough to rescue it from the more exacerbated reduction of both transition ($\sim 4\%$ reduced) μ 's and change in permanent electric ($\sim 5\%$ reduced) μ 's. In fact, when looking at the intensity rankings of 9-methyl among other planar structures, it has the weakest μ_{01} at 11.68 D and second weakest $\Delta\mu$ at 4.82 D, which further supports the damping of its 2PA signal. On the flip side, removing C9 from the wild-type to produce 13-methyl resulted in a significant increase in δ^{2PA} as noted above, despite it having the highest excitation energy of all structures within this study at 2.8 eV. Interestingly, the 13-methyl analogue also have a slightly lower μ_{01} than the wild-type. In fact, nearly all structures exhibit a μ_{01} within a narrow range of 0.285 D, or between 12.10 to 12.39 D with the exception of 9-methyl and 9,10,13-trimethyl(twist). These two outliers increases the μ_{01} range by double or more to 0.5 and 0.85 D respectively, resulting in the lowest μ_{01} values among all structures. Yet despite the higher energies or marginal differences in μ_{01} , 13-methyl has the second highest δ^{2PA} signal consistent with its $\Delta\mu$. In fact, 13-methyl actually has the largest change in dipole moments at 5.92 D, or over 0.8 D larger than the wild-type. This further supports C13-methyl enhancing 2PA by favoring a larger $\Delta\mu$ and hence stronger charge-transfer character in S_1 state. Decomposing $\Delta\mu$, one finds that the ground-state dipole μ_{00} of 9-methyl increases by nearly 4%, whereas 13-methyl decreases by a little over 10%. At first glance, this would seem to favor 9-meth, since its larger μ_{00} would be expected to widen the $\mu_{00}-\mu_{11}$ gap. However, this trend is outweighed by changes in the excited-state dipole μ_{11} ; 9-methyl increases by over 30%, while 13-methyl instead decreases by nearly 60%. As a result, 13-methyl exhibits the smallest S_1 μ among all structures, and the reduction in μ_{11} ultimately shifts the balance in $\Delta\mu$ strongly toward 13-methyl, giving it the most favorable (largest) $\Delta\mu$ overall. Switching out the wild-type's C13-methyl for a C10-methyl does not seem to affect the excitation energies by much, a mere 0.01 eV increase, nor does it substantially change μ_{01} , which rises by 0.2 D ($\sim 2\%$). Nevertheless, this places μ_{01} of 9,10-dimethyl(C_s) as the highest within this study, yet it still exhibits a nearly 30% reduction in δ^{2PA} relative to the wild-type. Additionally, when compared to the 13-methyl structure, both the excitation energies and μ_{01} are even more δ^{2PA} favorable (excitation energies 0.07 eV smaller and μ_{01} 0.24 D larger), yet the δ^{2PA} of 9,10-dimethyl(C_s) is markedly worse, with a reduction of about 45%.

Looking to $\Delta\mu$ values, 9,10-dimethyl(C_s)'s is reduced by 16.5% relative to the wild-type and by nearly 30% relative to 13-methyl, scaling similarly to its δ^{2PA} . This again underscores the importance of the $\Delta\mu$ and μ_{01} balance in interpreting quadratic-response based 2PA. Although, Dalton2020 employs a full quadratic-response formulation, the behavior here is effectively two-state, with δ^{2PA} scaling more with $\Delta\mu$ than with μ_{01} . A slight saving grace is observed in the comparison to 13-methyl. Here, 9,10-dimethyl(C_s)'s δ^{2PA} still has a larger relative decrease than $\Delta\mu$ but the drop is less disproportionate ($\sim 2:3$ rather than $\sim 1:2$), consistent with partial compensation from the increased μ_{01} . Individually, the dipole moments take after the 9-methyl structure, with the ground and excited-state dipoles increasing by 6% and approximately 50%, respectively, relative to the wild-type. Although μ_0 increases, the substantially larger increase in μ_1 reduces the resulting vector difference between the two states, giving 9,10-dimethyl(C_s) the smallest $\Delta\mu$ among the planar structures. Returning the C13-methyl to the 9,10-dimethyl(C_s) produces the planar 9,10,13-trimethyl(C_s) structure. Relative to wild-type, the results show a reduction in excitation energies by almost 4.5%, a rather negligible increase in μ_{01} by almost half a percent, and an increase in $\Delta\mu$ by about 15% driven by a favorable increase in ground-state and decrease in excited-state μ of about 9% and 4% respectively. This results in the planar trimethyl structure having the second highest $\Delta\mu$ of 5.87 D, only ~ 0.06 D lower than 13-methyl (the largest $\Delta\mu$) or less than a percent. In essence, 9,10,13-trimethyl experiences a favorable increase in all properties which results in the highest δ^{2PA} . Interestingly, despite the largest δ^{2PA} , 9,10,13-trimethyl(C_s) actually has the second highest σ^{2PA} , only 1.1 GM lower than 13-methyl. This is due to the conversion from microscopic to macroscopic values, the higher excitation energy in 13-methyl bumps up the macroscopic σ to be slightly higher than the trimethyl structure.

As noted in 1.2.2, investigation into the effects of the twist along the C10-C11 and C11=C12 bonds resulted in a depressed charge-transfer and red-shifts. Therefore, it comes as no surprise that the results here would also be consistent with what was observed by Sugihara *et al.* [95] and Walczak *et al.* [92] and in similar fashion to the large reduction in δ^{2PA} noted above. Introducing a 10° and 27° twist to the C11=C12 bond of 9,10-dimethyl and 9,10,13-trimethyl structures respectively produces a red-shift in excitation energies by 0.05-0.09 eV, or about a 2% to 3.5% reduction from their planar counterparts. These red-shifts were not sufficient enough to rescue the twisted structures from the collective unfavorable shift of all other properties. μ_{01} is reduced by $\sim 2\%$ and 7% , being more amplified in the trimethyl structure. Similarly, $\Delta\mu$ is depressed by about 12% in the dimethyl structure and nearly double that at 23% for the trimethyl one, mirroring the results of Walczak *et al.* [92]. In tandem, δ^{2PA} falls by approximately 24% and 47% in the dimethyl and trimethyl structures respectively. In other words, 2PA attenuation is approximately twice the $\Delta\mu$ attenuation in both dimethyl and trimethyl structures, making their percent relative reduction follow a $\sim 1:2$ ratio,

similarly to the above 9,10-dimeth(C_s). In fact computing the percent relative reduction ratio of $\Delta\mu:\delta^{2PA}$ for all analogues against the wild-type demonstrated similar proportions ($\sim 1:1.8$) with 9-methyl and 9,10,13-trimeth(C_s) being the two outliers at 1:2.7 and 1:3.2, respectively, and a total average of $\sim 1:2.2$. Likely the additional damping from the reduction in μ_{01} observed in the twisted structures pushed the reduction of δ^{2PA} closer to 2. The magnitude of these changes further emphasizes the effect of the torsional twist around this double bond. In the trimethyl structure, the bond is proportionally more twisted to the reduction in these properties. The trimethyl's $\sim 27^\circ$ twist produces roughly double the $\Delta\mu$ loss and also about double the 2PA damping compared to the $\sim 10^\circ$ twist, a clear twist-response scaling consistent with Walczak's 1PA results.

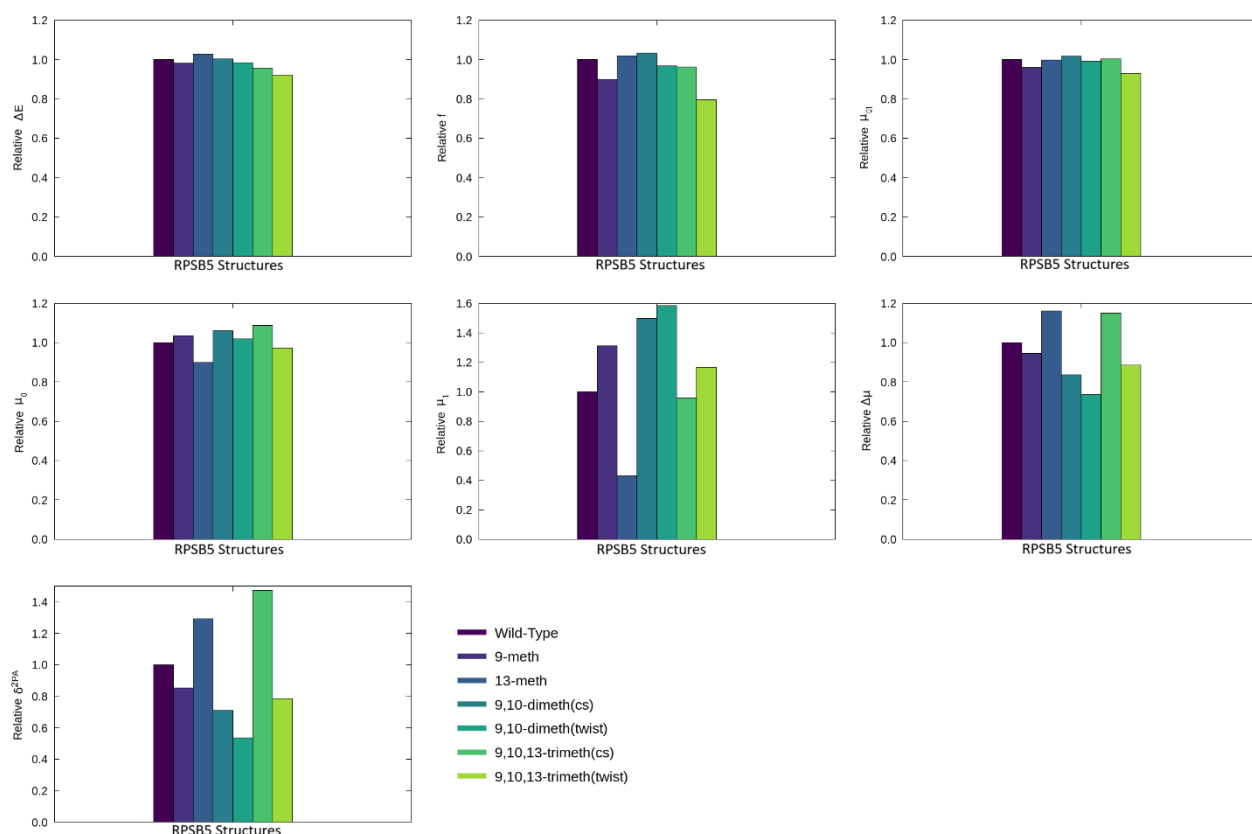


Figure 12. Wild-type normalized S_1 electronic and two-photon absorption properties of the RPSB5 analogues calculated at the RI-CC2/aug-cc-pVDZ level. All values are presented as dimensionless ratios relative to the wild-type (9,13-dimethyl structure).

Higher excited-states up to S_6 of RI-CC2 were also explored for all models, the results of which are presented in table 8 below. However, a quick glance over table 8 reveals that many of the excited-state δ^{2PA} are either negative, unrecoverable, or plagued by resonance enhancement with the S_1 state. In practically all cases the S_2 state exhibits a negative δ^{2PA} similarly to what was seen in 4.1, with the exception of 13-methyl that registers a near 31000 a.u. value. The S_3 and S_4 states for all models were unrecoverable in the Turbomole (2020 version 7.6.0) program, and therefore will be disregarded. Meanwhile, the S_5 state seems to be resonance enhanced (or at least near-resonance) since it is roughly twice the energy of S_1 . Equal-energy photons tuned to half of S_5 puts these transitions near two-

photon resonance, which boosts their 2PA signal. Despite its δ^{2PA} not exhibiting the dramatic divergences (as seen in the TD-DFT outliers discussed below), this proximity to being twice the energy of the S_1 state makes the computed state particularly sensitive to small energy shifts and to the absence of an explicit resonance/damping treatment in the response calculation, therefore, S_5 will also be disregarded from comparison. Nevertheless, these states display a δ^{2PA} signal of as low as about 6000 a.u. up to 27000 a.u., with a notable exception in 9-methyl and 9,10-dimethyl(C_s) that hovers around 150 to 460 a.u. One contributing factor is their separation from the S_1 -mediated two-photon resonance condition; i.e. detuning, or, how far the S_5 energy lies from twice the S_1 energy. For 9-methyl this detuning is relatively large (Δ resonance with $S_1 = +0.28$ eV), which likely reduced any near-resonant enhancement, whereas 9,10-dimethyl(C_s) is less detuned (Δ resonance with S_1 at +0.18 eV), implying that its unusually weak δ^{2PA} likely reflects additional damping from other factors like its state character or cancellations among contributing terms rather than detuning alone. The final state S_6 of the wild-type, interestingly is rather reminiscent to the computed S_6 state (figure 7 and table 2 in 4.1.3.2) of CC2/CCSD in 4.1, displaying a dimmer signal than the S_1 state. The small 2PA signal can also be similarly seen in nearly all structures except, again, 9-methyl and 9,10-dimethyl(C_s). For the most case, structures exhibit a δ^{2PA} signal of about 400 to 2000 a.u., with 9,10,13-trimethyl as a high outlier at over 7600a.u. Meanwhile, 9-methyl and 9,10-dimethyl(C_s) seems to shift this dimmer state from S_6 to the aforementioned S_5 , as noted by their weak δ^{2PA} . Nonetheless, hence all other structures are incomparable due to the issues mentioned above, it becomes impossible to determine methylation effects in these higher lying states here.

Wild-Type										9-methyl									
State	ΔE	λ	δ^{2PA}	σ^{2PA}	f	μ_0	μ	$\Delta\mu$	T1	State	ΔE	λ	δ^{2PA}	σ^{2PA}	f	μ_0	μ	$\Delta\mu$	T1
ground							7.837			ground							8.102		
1	2.72	910	21307	58.0	1.535	12.188	2.796	5.100	88.37	1	2.67	929	18170	47.4	1.382	11.684	3.663	4.818	88.6
2	4.36	569	-27059	-188.2	0.116	2.652	4.268	3.676	84.8	2	4.43	560	-226076	-1624.6	0.197	3.428	5.746	2.431	84.5
3	5.05	491	****	-	0.188	3.136	3.627	4.603	85.44	3	5.04	492	****	-	0.290	3.895	5.892	3.017	87.8
4	5.27	470	****	-	0.012	0.779	6.080	1.759	87.83	4	5.38	461	****	-	0.027	1.159	2.941	5.641	86.3
5	5.61	442	9013	104.0	0.000	0.103	4.891	12.727	90.73	5	5.62	441	146	1.7	0.000	0.064	6.525	11.422	91.0
6	5.98	415	2146	28.1	0.000	0.058	1.248	8.778	90.85	6	5.85	424	1	0.0	0.001	0.252	8.385	15.863	90.6
13-methyl										9-10-di-methyl(C ₅)									
State	ΔE	λ	δ^{2PA}	σ^{2PA}	f	μ_0	μ	$\Delta\mu$	T1	State	ΔE	λ	δ^{2PA}	σ^{2PA}	f	μ_0	μ	$\Delta\mu$	T1
ground							7.044			ground							8.312		
1	2.80	885	27544	79.1	1.567	12.148	1.207	5.922	88.34	1	2.73	909	15097	41.1	1.587	12.388	4.188	4.261	88.37
2	4.44	558	30798	222.9	0.106	2.510	3.415	3.652	85.07	2	4.33	572	-79196	-544.9	0.121	2.713	5.026	3.356	84.39
3	5.19	478	****	-	0.239	3.486	5.277	1.832	87.24	3	5.00	496	****	-	0.114	2.451	2.831	6.318	85.22
4	5.43	457	****	-	0.005	0.468	4.996	2.048	86.02	4	5.24	473	****	-	0.035	1.333	5.416	2.913	88.00
5	5.71	434	10630	127.2	0.000	0.054	6.711	13.755	85.52	5	5.63	440	462	5.4	0.001	0.257	6.504	14.007	90.51
6	6.09	407	545	7.4	0.000	0.033	1.636	7.079	87.75	6	5.90	420	-0.4	0.0	0.001	0.181	5.194	12.355	90.67
9-10-di-methyl(twist)										9-10-13-tri-methyl(C ₅)									
State	ΔE	λ	δ^{2PA}	σ^{2PA}	f	μ_0	μ	$\Delta\mu$	T1	State	ΔE	λ	δ^{2PA}	σ^{2PA}	f	μ_0	μ	$\Delta\mu$	T1
ground							7.991			ground							8.527		
1	2.68	925	11451	30.1	1.488	12.102	4.424	3.752	88.48	1	2.60	953	31434	78.0	1.478	12.238	2.677	5.865	87.94
2	4.30	577	-147989	-1002.6	0.117	2.680	6.098	1.912	84.06	2	4.19	591	-70112	-451.4	0.167	3.238	5.392	3.135	84.90
3	4.98	498	****	-	0.160	2.909	4.210	4.587	86.60	3	4.88	508	****	-	0.141	2.761	0.392	8.647	84.98
4	5.14	482	****	-	0.031	1.263	4.045	4.114	87.15	4	5.08	488	****	-	0.027	1.192	7.016	1.519	88.00
5	5.52	449	14076	157.4	0.001	0.235	1.952	6.437	88.63	5	5.49	452	5881	64.8	0.002	0.319	2.565	10.959	90.61
6	5.58	444	383	4.4	0.002	0.314	5.751	12.610	90.60	6	5.81	427	1118	13.8	0.001	0.221	4.071	12.593	90.34
9-10-13-tri-methyl(twist)										9-10-13-tri-methyl(twist)									
State	ΔE	λ	δ^{2PA}	σ^{2PA}	f	μ_0	μ	$\Delta\mu$	T1	State	ΔE	λ	δ^{2PA}	σ^{2PA}	f	μ_0	μ	$\Delta\mu$	T1
0							7.625			0									
1	2.51	989.509	16752	38.5	1.221	11.337	3.258	4.506	88.24	1	2.51	989.509	16752	38.5	1.221	11.337	3.258	4.506	88.24
2	3.98	622.616	-43015	-249.9	0.130	2.937	5.728	1.945	85.22	2	4.19	591	-70112	-451.4	0.167	3.238	5.392	3.135	84.90
3	4.84	512.200	****	-	0.117	2.528	2.512	8.750	84.85	3	4.88	508	****	-	0.141	2.761	0.392	8.647	84.98
4	4.95	500.763	****	-	0.046	1.569	6.732	0.899	88.10	4	5.08	488	****	-	0.027	1.192	7.016	1.519	88.00
5	5.24	473.653	269271	2703.2	0.008	0.648	5.789	1.889	85.26	5	5.49	452	5881	64.8	0.002	0.319	2.565	10.959	90.61
6	5.40	458.899	7678	82.1	0.008	0.629	1.323	6.528	88.02	6	5.81	427	1118	13.8	0.001	0.221	4.071	12.593	90.34

Table 8. Ground and S₁ - S₀ excited-state electronic and two-photon absorption properties of all RPSB5 models calculated at the RI-CC2/aug-cc-pVDZ level of theory. Excitation energies (ΔE) are reported in eV, wavelengths (λ) in nm, δ^{2PA} in a.u., and σ^{2PA} in GM. Transition dipole moments (μ_0), ground and excited-state permanent dipole moments (μ_0 and μ_n), and changes in permanent dipole moment ($\Delta\mu$) are reported in Debye. Oscillator strengths (f) are dimensionless, and T₁ single-excitation contributions are reported as percentages. Unrecoverable values are indicated by asterisks.

4.2.4 Performance of TD-DFT Functionals in the S₁ Excited-state

Starting with the excitation energies of the S₁ state, observing table 9 below, one can see that all functionals are able to reproduce RI-CC2 values fairly well. Most trends in the DFT functionals investigated here follow that of RI-CC2 closely, whereby all analogues gain/lose excitation energy in their transformation from the wild-type in the same direction as RI-CC2 with the only exception being the 9,10-dimethyl(C_s) structure, as seen in table 9 below. In this dimethyl structure RI-CC2 actually increases by 0.007 eV, while all TD-DFT functionals displayed a relative decrease from the wild-type by 0.002 – 0.03 eV. Interestingly, in terms of absolute values, the meta-GGA/meta-NGA functionals outperform the previously tested ones from section 4.1 by about 20 – 40%. Specifically, CAM-B3LYP and BHandHLYP demonstrated an MAE of 0.215 and 0.293 eV respectively, while M11 and MN15 brings this value down to 0.168 and 0.175 eV respectively.

Transition dipole moments, μ_{01} , also tell a similar story. Observing table 9 below, one can notice that the trends in TD-DFT mirror that of RI-CC2 exactly again. All structures in all functionals, with the exception of 9-methyl, and 9,10,13-trimethyl(twist) have a μ_{01} that lie along a narrow spread of 0.32 – 0.33 D, similarly to RI-CC2's of 0.29 D. Additionally, the outliers within TD-DFT methods also follow the same scale as outliers of RI-CC2, 9-methyl underestimates the wild-type by about 4% or 0.42 – 0.5D. Whereas the trimethyl(twist) is slightly more exaggerated here than in RI-CC2 (7% underestimated) and brings this value down even lower to 9% or about 0.85 – 0.99 D lower than the wild-type. Despite the agreeable trends, TD-DFT still underestimates μ_{01} absolute values by about a full Debye as clearly seen in table 9 below. Observing table 9 below will also show that all functionals display an MAE of 0.954 to 1.21 D, with the best functionals being BHandHLYP, followed closely by M11.

The top row of figure 13 also summarizes the δ^{2PA} results of RI-CC2 and DFT functionals considered within this sub-chapter. The intensity ranking order shown among all methods clearly mirror one another with the exception of 9,10,13-trimethyl(twist), the only other caveat is also the dampened absolute intensities, which are unsurprising given the results of the previous sub-chapter. On closer inspection of table 9, DFT functionals clearly exaggerate the percent gain and loss upon transformation to each of the analogues. In RI-CC2, all analogues shift by ~15% (9-methyl) to 47.5% (trimethyl cs) from the wild-type. In DFT, notwithstanding the trimethyl(twist) outlier, these values increase by a fair amount up to about 26% on the low end and 69% on the higher end in M11, where the most severely impacted functional, BHandHLYP climbs even higher to 31% till ~95% shifts for the same structures. This makes the M11 functional the least shifted DFT method by raw percentage

alone. Again, the only exception here where percent δ^{2PA} shifts of the analogues are not exaggerated from RI-CC2 but instead underestimated is for the 9,10,13-trimethyl(twist) structure. In RI-CC2, this structure shifts down from the wild-type by about 21%, while among the DFT functionals, this shift is only down by ~4% in BHandHLYP, ~6% in M11, ~7% in CAM-B3LYP, and 14% in MN15. Thus making the trimethyl(twist) structure display the 4th most intense δ^{2PA} signal in DFT functionals instead of RI-CC2 which places it at the 5th most intense. Normalizing DFT functional's δ^{2PA} percent shifts against RI-CC2's to determine which functional generally captures the best observed trends show that M11 keeps this ratio between 0.28 (trimethyl twist) up to 1.76 (9-methyl), with an average of 1.18 among all analogues. In other words, on average, M11 predicts an 18% larger shift in magnitude from the wild-type to the analogue structures than RI-CC2. On the other hand CAM-B3LYP brings this averaged ratio up to 1.42, while MN15 and BHandHLYP ties at 1.53. However, these averages are slightly misleading, giving DFT functionals a higher fidelity than what would be expected. As noted earlier, the 9,10,13-trimethyl(twist) structure is an outlier in both magnitude and intensity ranking when compared to RI-CC2 and therefore, due to the underestimated percent shift in this structure, the averages of the DFT functionals are skewed lower. Leaving out this structure brings all the averaged ratios up to 1.36 in M11, 1.65 in CAM-B3LYP, 1.70 in MN15, and the most overestimated functional still being BHandHLYP at a ratio of 1.80. Overall, M11 most closely reproduces both the absolute percent-change magnitudes and the RI-CC2-normalized shift scaling, outperforming the other functionals in terms of qualitative fidelity.

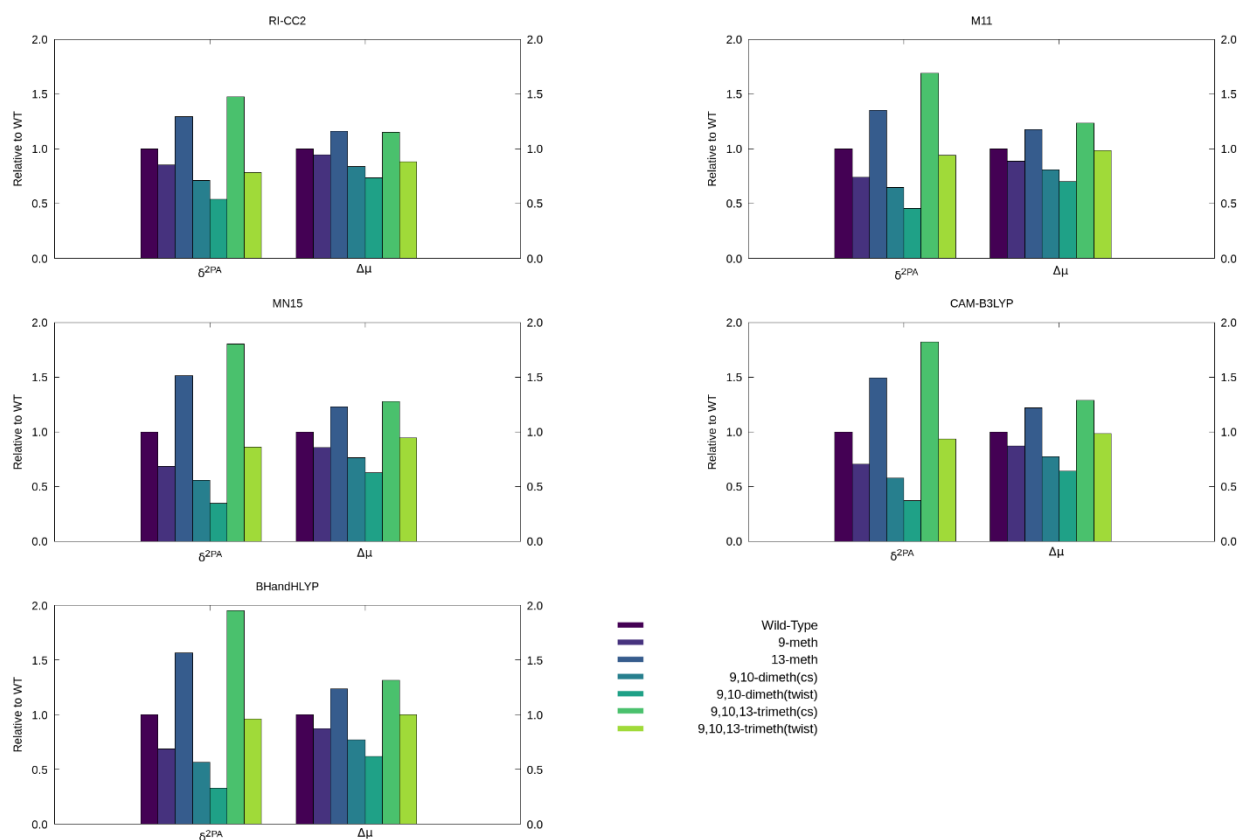


Figure 13. Wild-type-normalized two-photon transition strengths (δ^{2PA}) and changes in permanent dipole moment ($\Delta\mu$) for the RPSB5 analogue series calculated using RI-CC2, M11, MN15, CAM-B3LYP, and BHandHLYP with the aug-cc-pVDZ basis set. Each panel corresponds to one electronic-structure method, and all values are presented as dimensionless ratios relative to the wild-type 9,13-dimethyl structure.

Within DFT, the trimethyl(twist) structure demonstrates only a slightly lower μ_{01} (by $\sim 2\%$) shift than RI-CC2 as noted above. And thus, the increase in intensity ranking of the trimethyl(twist) in DFT is largely owed to its $\Delta\mu$. Analogously to the δ^{2PA} , percent decrease in $\Delta\mu$ for DFT functionals in this structure compared to the wild-type is only fractions of RI-CC2. RI-CC2 displays a decrease of about 11.7% while DFT functionals brings this down to $\sim 5.6\%$ in MN15, $\sim 1.5\%$ in either M11 or CAM-B3LYP, and as low as 0.05% in BHandHLYP. The higher δ^{2PA} ranking of trimethyl(twist) in DFT is also consistent with its $\Delta\mu$ behavior. RI-CC2 predicts a $\Delta\mu$ decrease of $\sim 11.7\%$ from wild-type, but DFT dampens this to $\sim 5.6\%$ in MN15, $\sim 1.5\%$ in both M11 and CAM-B3LYP, and as low as 0.05% in BHandHLYP. Normalized against RI-CC2, this corresponds to only $\sim 48\%$, $\sim 12\text{--}13\%$, and $\sim 0.4\%$ of the expected $\Delta\mu$ drop, respectively, thus making the trimethyl(twist) in DFT not nearly as dampened as in RI-CC2. Furthermore, the 9-methyl structure percent $\Delta\mu$ shifts from the wild-type is the most exacerbated model in DFT functionals. Wherein RI-CC2 shifts this analogue down from the wild-type by $\sim 5.5\%$, DFT functionals doubles this or more at $\sim 11\%$ (M11) to over $\sim 14\%$ (MN15). Consistent with this amplified $\Delta\mu$ shift, the corresponding δ^{2PA} shift for 9-methyl is likewise enlarged in DFT; as noted above, M11 reaches 1.76x the RI-CC2 shift, while the most exaggerated case is MN15 at 2.13x, reinforcing that enhanced $\Delta\mu$ attenuation correlates to amplified δ^{2PA} damping. Whereas trimethyl(twist) is the slightly unique case where DFT nearly eliminates the $\Delta\mu$ penalty and thus, as expected, underestimate the δ^{2PA} loss, resulting in the re-ordering of this analogue's intensity in DFT. On average functionals predict $\Delta\mu$ shift amplification from RI-CC2 of about 1.18x in M11, 1.41 in CAM-B3LYP, 1.43 in BHandHLYP and, 1.53 in MN15. Excluding the trimethyl(twist) outlier brings the ordering of this amplification close to mirroring δ^{2PA} above save for the difference in MN15 and BHandHLYP, at 1.39 (M11), 1.67 (CAM-B3LYP), and 1.74 (both MN15 and BHandHLYP). Extending the analysis to the $\Delta\mu:\delta^{2PA}$ ratio provides a direct test of whether each DFT functional reproduces the $\Delta\mu - \delta^{2PA}$ scaling observed at the RI-CC2 level. As noted earlier, RI-CC2 averages the ratio at $\sim 1:2.2$ (1:2.17), the best scaling DFT functional with a ratio of $\sim 1:2.25$ surprisingly belongs to MN15, with M11 and CAM-B3LYP following closely at 1:2.56 and 1:2.60 respectively. And as usual, BHandHLYP brings this ratio up by a huge amount to 1:13.99. As noted above, the trimethyl(twist) structure is a huge outlier with its extremely low $\Delta\mu$ shift from the wild-type in DFT functionals, particularly in BHandHLYP. Neglecting this outlier brings all ratios for all functionals to 1:2.18 (M11) to 1:2.20 (CAM-B3LYP), with BHandHLYP still being the most overestimated at 1:2.30. In summary, the $\Delta\mu:\delta^{2PA}$ scaling is broadly conserved across the DFT functionals, with M11 showing a slight edge in reproducing the RI-CC2 proportions. Viewed purely in terms of $\Delta\mu$ shift

amplification, M11 again performs the best, overestimating RI-CC2 shift magnitudes by only ~18% (1.18x) on average, or about 23% less than the next best functional (CAM-B3LYP at 1.41x).

RI-CC2							
Structure	ΔE [eV]	δ^{2PA} [a.u.]	σ^{2PA} [GM]	μ_{01} [D]	μ_0 [D]	μ_1 [D]	$\Delta\mu$ [D]
9,13-dimethyl	2.720	21307	58.0	12.188	7.837	2.796	5.100
9-methyl	2.670	18170	47.4	11.684	8.102	3.663	4.818
13-methyl	2.800	27544	79.1	12.148	7.044	1.207	5.922
9,10-dimethyl(C _S)	2.727	15097	41.1	12.388	8.312	4.188	4.261
9,10-dimethyl(twist)	2.680	11451	30.1	12.102	7.991	4.424	3.752
9,10,13-trimethyl(C _S)	2.600	31434	78.0	12.238	8.527	2.677	5.865
9,10,13-trimethyl(twist)	2.510	16752	38.5	11.337	7.625	3.258	4.506
MN15 (2PA gauge variant)							
9,13-dimethyl	2.906	1655	5.12	11.024	6.497	4.771	1.739
9-methyl	2.849	1136	3.38	10.549	6.690	5.289	1.492
13-methyl	2.976	2504	8.12	11.058	5.771	3.651	2.133
9,10-dimethyl(C _S)	2.904	920	2.84	11.148	6.993	5.728	1.327
9,10-dimethyl(twist)	2.849	573	1.70	10.820	6.663	5.639	1.088
9,10,13-trimethyl(C _S)	2.779	2986	8.45	10.965	7.259	5.034	2.224
9,10,13-trimethyl(twist)	2.668	1423	3.71	10.037	6.483	4.899	1.642
MAE	0.175	18651	48.4	1.212	1.298	1.828	3.226
M11 (2PA gauge variant)							
9,13-dimethyl	2.900	4713	14.52	11.234	6.942	4.053	2.899
9-methyl	2.836	3494	10.30	10.815	7.081	4.641	2.576
13-methyl	2.991	6378	20.90	11.245	6.336	2.939	3.407
9,10-dimethyl(C _S)	2.874	3055	9.24	11.378	7.357	5.070	2.344
9,10-dimethyl(twist)	2.815	2156	6.26	11.056	6.991	5.045	2.030
9,10,13-trimethyl(C _S)	2.791	7959	22.71	11.281	7.804	4.231	3.583
9,10,13-trimethyl(twist)	2.674	4432	11.61	10.262	7.008	4.192	2.860
MAE	0.168	15653	39.5	0.973	0.846	1.137	2.075
CAM-B3LYP							
9,13-dimethyl	2.950	2187	7.0	11.179	6.533	4.537	2.007
9-methyl	2.880	1541	4.7	10.710	6.704	4.975	1.749
13-methyl	3.020	3267	10.9	11.212	5.857	3.417	2.449
9,10-dimethyl(C _S)	2.930	1266	4.0	11.307	7.000	5.451	1.549
9,10-dimethyl(twist)	2.880	810	2.5	10.979	6.652	5.370	1.285
9,10,13-trimethyl(C _S)	2.830	3987	11.7	11.173	7.332	4.753	2.580
9,10,13-trimethyl(twist)	2.720	2037	5.5	10.222	6.567	4.599	1.977
MAE	0.215	18094	46.6	1.043	1.256	1.556	2.947
BhandHLYP							
9,13-dimethyl	3.020	1660	5.5	11.263	6.334	4.495	1.849
9-methyl	2.960	1143	3.7	10.785	6.500	4.908	1.612
13-methyl	3.100	2597	9.2	11.313	5.606	3.331	2.284
9,10-dimethyl(C _S)	3.010	939	3.1	11.384	6.812	5.383	1.430
9,10-dimethyl(twist)	2.960	547	1.8	11.063	6.460	5.315	1.147
9,10,13-trimethyl(C _S)	2.910	3239	10.1	11.283	7.115	4.680	2.435
9,10,13-trimethyl(twist)	2.800	1595	4.6	10.315	6.360	4.524	1.848
MAE	0.293	18576	47.8	0.954	1.464	1.489	3.089

Table 9. Ground and S_1 electronic and two-photon absorption properties of all RPSB5 models calculated using RI-CC2 and the M11, MN15, CAM-B3LYP, and BHandHLYP functionals with the aug-cc-pVDZ basis set. Excitation energies (ΔE) are reported in eV, δ^{2PA} in a.u., σ^{2PA} in GM, and the transition dipole moment (μ_{01}), ground- and excited-state permanent dipole moments (μ_0 and μ_1), and change in permanent dipole moment ($\Delta\mu$) in Debye. Mean absolute errors (MAEs) are reported relative to RI-CC2. The 2PA quantities obtained using M11 and MN15 are gauge variant.

In terms of twisted structure compared against their planar counterparts, all four DFT functionals reproduce the same qualitative twist-response established at the RI-CC2 level for both methylation states. That is, introducing the twist consistently induces a small red-shift in the excitation energy, dampens μ_{01} , reduces $\Delta\mu$, and attenuates δ^{2PA} . For 9,10-dimethyl, the planar to twist red-shift is comparable to RI-CC2 (~ 1.7 – 2.1% in DFT vs $\sim 1.7\%$ in RI-CC2), accompanied by dampening of μ_{01} by $\sim 3\%$ (vs 2.3%), $\Delta\mu$ by ~ 13 – 20% (vs 12%), and δ^{2PA} by ~ 29 – 42% (vs 24%). For 9,10,13-trimethyl, DFT similarly preserves the stronger twist sensitivity; excitation energies drop by ~ 3.8 – 4.2% (vs 3.5%), μ_{01} by ~ 8.5 – 9.0% (vs 7.4%), $\Delta\mu$ by ~ 20 – 26% (vs 23%), and δ^{2PA} by ~ 44 – 52% (vs 47%). Importantly, the $\Delta\mu:\delta^{2PA}$ attenuation ratio remains essentially conserved across all functionals for the planar to twist comparisons, with a ratio staying near ~ 2 for both dimethyl and trimethyl structures (~ 2.0 – 2.2), in direct agreement with the RI-CC2's ($\sim 1:2$) behavior. Quantitatively, M11 tends to track RI-CC2 closest in the δ^{2PA} damping magnitudes (reduction of $\sim 29\%$ vs 24% in dimethyl, and $\sim 44\%$ vs 47% in trimethyl), whereas the other functionals generally over-amplify the dimethyl damping but still preserve the overall twist-response ordering.

Despite M11's good qualitative prediction of RI-CC2 results, DFT functional's absolute values still leave a lot to be desired as seen in Table 9 above. Across the full analogue set, all TD-DFT functionals systematically underestimate δ^{2PA} , but to very different extents. M11 is consistently the least severe yet still multiple times smaller, with δ^{2PA} signals spanning ~ 3.8 – $5.3\times$ (a mean of about $4.6\times$) less intense than RI-CC2 among all analogues. In contrast, CAM-B3LYP underestimates RI-CC2 by ~ 7.9 – $14.1\times$ (with a mean of $\sim 10.3\times$), while BHandHLYP and MN15 typically stay between ~ 9.7 – $20.9\times$ (both means between 13.8 and $14.1\times$, respectively), where MN15 remained the worst overall. This absolute δ^{2PA} fidelity ranking echoes the trend analysis above; when going from the wild-type to the respective analogues, functionals that over-amplify the percent changes in δ^{2PA} (e.g., MN15/BHandHLYP) are likewise the ones that show the most severe suppression in absolute values as well. Meanwhile, M11 shows the weakest exaggeration and the smallest absolute suppression. The same (similar?) ordering also emerges from the Mean Absolute Error metrics. The δ^{2PA} MAE is lowest again for M11, followed by CAM-B3LYP, then BHandHLYP and MN15, reinforcing that M11 is the most defensible functional here for retinal. This is contrary to the performance trend reported for other organic molecule sets by Ahmazadeh *et al.* [184] and separately, Grotjahn & Furche [197], but instead aligning with the earlier recommendation of M11 by Brown and co-workers for coumarin dyes [199]. A closer look at dipoles and their differences again help explain why this absolute δ^{2PA} damping is so persistent. In line with Jacquemin's [202] observations where permanent electric dipole moments are highly functional dependent, particularly for charge-transfer states, DFT here underestimates ground-state dipoles (μ_0) while overestimating excited-state dipoles (μ_1) relative to

RI-CC2. Quantitatively, μ_0 is shifted downward by roughly 0.6 – 1.6 D across all models and all functionals where on average M11 is underestimated the least by only -0.85 D, -1.26 D for CAM-B3LYP, -1.30 D for MN15, -1.46 D for BHandHLYP. Meanwhile μ_1 is shifted in the opposite direction, upwards, by as much as 0.6 – 2.4 D, again on average the least up shifted functional is M11 at $+1.14$ D, up to the most up shifted being MN15 at $+1.83$ D. The net effect is a strong damping of $\Delta\mu$, with $\Delta\mu$ being roughly 1.6 – 1.9x smaller in M11 than in RI-CC2, versus being 2.3 – 3.4x smaller in the remaining functionals. Understandably this puts the $\Delta\mu$ MAEs in a similar order with M11 having the smallest (2.075 D), followed by CAM-B3LYP (2.947 D), BHandHLYP (3.089 D), and the largest, MN15 (3.226 D). Overall, M11 is the most consistent performer across both trend fidelity and absolute-value metrics, with CAM-B3LYP typically second. MN15 and BHandHLYP are more property-dependent, where MN15 is closer in μ_0 , whereas BHandHLYP is closer in μ_1 , $\Delta\mu$, and (slightly) δ^{2PA} absolute values. Since δ^{2PA} tracks $\Delta\mu$ in this series, TD-DFT's consistent bias systematically dampens $\Delta\mu$ and therefore yields a predictable underestimation of absolute δ^{2PA} .

4.2.5 Performance of TD-DFT Functionals in the Higher Excited-states

The complete higher excited-state TD-DFT and RI-CC2 results for all RPSB5 analogues are provided in Table S2, while the corresponding RI-CC2 results are reported in Table 8. Within the wild-type structure at the RI-CC2 level, S_2 is negative, $S_3 - S_4$ were unrecoverable, while S_5 lies near twice the S_1 energy and is therefore treated as near resonance, both of which are excluded from quantitative comparison. This leaves S_6 as the only higher-lying reference feature that is clearly non-resonant and remains weaker than the S_1 state ($\sigma^{2PA} = 28.1$ vs 58.0 GM). In contrast, the TD-DFT results consistently show that substantial 2PA signal is redistributed into a higher-energy window spanning roughly 4.8 – 5.7 eV, however, this feature is more accurately described as a multi-state cluster as opposed to the singular second peak validated in RI-CC2. Depending on the functional and structure, two or more states between the $S_2 - S_5$ can exhibit comparable or even exceed σ^{2PA} intensity of the S_1 state. For example, in the wild-type, BHandHLYP already indicates that higher state activity can exceed S_1 well below 6 eV, which correlates to the valid S_6 state of RI-CC2 (5.98 eV) as well as the resonant enhanced region of this functional and structure (~ 6.04 eV). More specifically, BHandHLYP's S_2 state demonstrates a 2PA signal of ~ 34.9 GM at 4.64 eV, which is about 6 – 7x more intense than its S_1 (~ 5.5 GM) state. While its strongest contribution occurs at S_3 (~ 1188 GM) with an excitation energy of 5.22 eV (Δ resonance with $S_1 = -0.82$ eV, i.e. not S_1 -mediated near/resonance). Both S_2 and S_3 states are also accompanied by a dim oscillator signal of 0.084 and 0.144 respectively. CAM-B3LYP and MN15 similarly places both these states in the S_2 and S_3 region

as well. With their dominant higher state responses sitting near 5.0 ~ 5.2 eV with a σ^{2PA} of about 1900 ~ 2700 GM, and also accompanied by a dim oscillator signal of 0.149 and 0.142, respectively. The weaker S_2 state of CAM-B3LYP is closely reminiscent of BHandHLYP at 32.8 GM and a dim oscillator of 0.078, while MN15 drops the 2PA signal down to ~10 GM while maintaining oscillators at 0.089. In contrast, M11 tends to flip the S_2 and S_3 states when compared to the other functionals. This is supported by the higher, yet more modest than the other functionals, 2PA intensity of 282 GM in the S_2 state but is accompanied by the weaker oscillator signal ($f \sim 0.063$). Whereas its adjacent S_3/S_4 states, both being nearly degenerate in excitation energy, carry the larger oscillator strength ($f \sim 0.159$) yet remain dimmer in terms of σ^{2PA} (~23 and 2 GM, respectively). This illustrates that the computed state index is not transferable across functionals even when the energy window is similar.

Perhaps more importantly, the higher state intensity ratios are also not proportional to the RI-CC2 non-resonant state as noted above. RI-CC2 retains its only reliable higher reference (S_6) as fractions of S_1 's intensity (~0.48x), while TD-DFT often predicts the 5.0 – 5.2 eV manifold to exceed S_1 by orders of magnitude, making direct comparisons to RI-CC2 reference practically impossible. Furthermore, exploring whether this expanded PSB5 structure can draw any meaningful parallels to the smaller PSB3 structure 2PA patterns in section 4.1 resulted purely superficial similarities. PSB3's CC2/aug-cc-pVDZ shows a moderate S_1 (12.3 GM), and a split upper manifold where one near-degenerate component is essentially dark while the other is bright (S_2 : 55GM and S_3 : 0.02 GM). Expanding to the larger PSB5 wild-type structure in TD-DFT functionals likewise provides a useful probe of whether system-size expansion can, or at least partially, converge a lower-level method toward the CC2-like 2PA pattern. And although TD-DFT here produces a brighter band in the S_2 or S_3 states than the S_1 state, it is followed by the second brightest band in either the same S_2 or S_3 state (functional dependent), thus making the CC2-like (from section 4.1) intensity proportions (and total intensity rankings) also not recovered here. Additionally, a subset of analogues further exhibit clear S_1 -mediated near/at-resonant amplification, most dramatically 9,10,13-trimethyl(twist) in M11 (where S_5 's Δ resonance with $S_1 \sim +0.02$ eV, i.e. resonance enhanced), and again such cases should therefore be treated as diagnostically useful but not quantitatively, and rarely qualitatively, comparable.

4.2.6.1 2-State Model as a Good Qualitative Predictor of QR

Structure	RI-CC2		M11		MN15		CAM-B3LYP		BHandHLYP	
	δ^{2PA}	σ^{2PA}	δ^{2PA}	σ^{2PA}	δ^{2PA}	σ^{2PA}	δ^{2PA}	σ^{2PA}	δ^{2PA}	σ^{2PA}
9,13-dimethyl	29639	80.3	7158	22.1	2471	7.6	3283	10.5	2698	9.0
9-methyl	25119	65.6	5469	16.2	1726	5.1	2395	7.3	1951	6.3
13-methyl	37398	107.4	9311	30.5	3566	11.6	4692	15.7	3939	13.9
9,10-dimethyl (Cs)	21199	57.8	4872	14.7	1451	4.5	2006	6.3	1640	5.4
9,10-dimethyl (twist)	16209	42.7	3587	10.5	951	2.8	1341	4.1	1023	3.3
9,10,13-trimethyl (Cs)	43253	107.1	11909	34.0	4366	12.4	5887	17.3	5060	15.7
9,10,13-trimethyl (twist)	23316	53.8	6766	17.7	2095	5.5	3053	8.3	2543	7.3

Table 10. δ^{2PA} and σ^{2PA} (in a.u. and GM, respectively) for the lowest excited state in RPSB5 models, calculated using two-state model (2SM).

Across all seven RPSB5 analogues and all five methods considered, the two-state model (2SM) tracks the full quadratic-response δ^{2PA} remarkably well for the S_1 state. Figure 14 below shows a near one-to-one relationship between the quadratic response formalism based δ^{2PA} and the 2SM based one for each method. All methods display correlation coefficients that are effectively one ($r = 0.9991 - 0.9997$ across RI-CC2 and all TD-DFT functionals), with RI-CC2 slightly in the lead at $r=0.9997$. In practice, this means that the 2SM expression captures the relative impact of de/methylation on the S_1 two-photon strengths extremely reliably. Structures that were seen to enhance or dampen 2PA signal in quadratic response are also enhanced or dampened by the same amount in the 2SM, while also preserving the ordering across analogues. For example, in TD-DFT (M11), the strongest enhancements in quadratic response belong to 13-methyl at +44%, and 9,10,13-trimethyl(C_s) at +56%, are mirrored closely by 2SM at +38.0% and +53.8%, respectively. Similarly, the largest attenuation in quadratic response, 9,10-dimethyl(twist) dampens by -56.9%, which is reproduced by 2SM at -52.5%. The same behavior is also seen at the RI-CC2 level; 13-methyl rises by +36% in quadratic response and +34% in 2SM, while 9,10-dimethyl(twist) drops by -48% in quadratic response versus -47% in 2SM. Even so, within each method, 2SM systematically overestimates the absolute 2PA intensities while reproducing the same structure-to-structure pattern. For RI-CC2, the mean offset is approximately +7800 a.u., corresponding to about +20 GM. The TD-DFT offsets are smaller and remain fairly uniform within each functional: approximately +2400 a.u. or +7 GM for M11, +1080 a.u. or +3.3 GM for CAM-B3LYP, +1020 a.u. or +3.3 GM for BHandHLYP, and +776 a.u. or +2.3 GM for MN15. Overall, the near-perfect linearity indicates that the 2SM ingredients, μ_{01} and $\Delta\mu$, are sufficient to reproduce the structure-to-structure variation observed in the full quadratic-response results. Thus, δ^{2SM} provides a simple and robust descriptor of relative 2PA trends across the 11-cis rPSB models, although it should not be treated as a quantitative replacement for the full quadratic-response values.

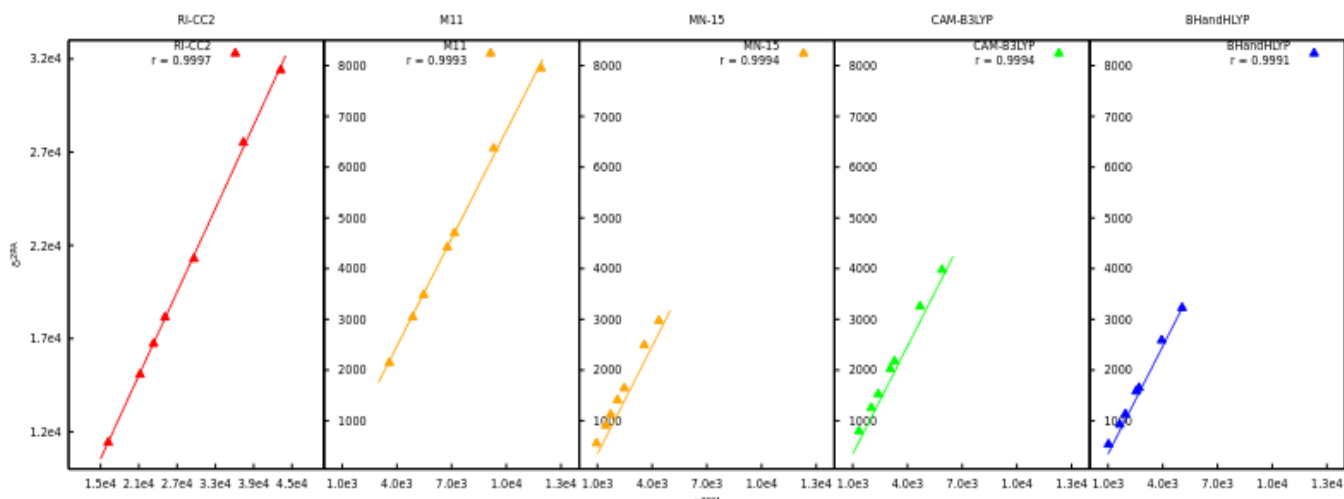


Figure 14. Comparison of two-photon transition strengths from response theory (δ^{2PA}) and two-state model (δ^{2SM}) for RPSB5 models obtained at different levels of theory.

4.2.6.2 Retinal Protonated Schiff Base with and without β -ionone Ring

M11 (gauge invariant 1PA and gauge variant 2PA)							
Structure	ΔE [eV]	δ^{2PA} [a.u.]	σ^{2PA} [GM]	μ_{01} [D]	μ_0 [D]	μ_1 [D]	$\Delta\mu$ [D]
9,13-dimethyl	2.576	46084	111.0	12.846	17.852	11.171	6.799
9-methyl	2.558	34837	82.7	12.145	17.403	11.416	6.222
13-methyl	2.605	57791	142.2	13.584	18.210	11.157	7.158
9,10-dimethyl(twist)	2.511	31090	71.1	12.863	16.934	11.720	5.410
9,10,13-trimethyl(twist)	2.417	41388	87.6	12.532	16.951	11.084	5.988

Table 11. M11/aug-cc-pVDZ values for 5 different RPSB5 models. 1PA properties were performed in gauge-invariant formalism, while 2PA properties were performed in gauge-variant formalism.

One may also wonder whether introducing the β -ionone ring into the polyene chain (RPSB) alters the de/methylation trends established for the RPSB5 models. To determine this, M11/aug-cc-pVDZ calculations were performed on five RPSB analogues. Figure 15 below shows that, on a logarithmic scale, the RPSB and RPSB5 curves display essentially the same shape (and therefore the same analogue intensity ordering), but with the RPSB series shifted upward by roughly one log unit, indicating a systematic enhancement of the absolute 2PA scale upon ring inclusion. Qualitatively, the WT-relative 2PA shifts remain comparable between the two sets of analogues. For the RPSB5 model with the M11 functional, δ^{2PA} shifts relative to the WT are -26% in 9-methyl, $+35\%$ in 13-methyl, -54% in 9,10-dimethyl(twist), and -6% in 9,10,13-trimethyl(twist). Meanwhile, the corresponding RPSB shifts are -24% , $+25\%$, -33% , and -10% , respectively. Thus, both systems preserve the same

qualitative ranking, where 13-methyl remains the most enhanced, while the twisted dimethyl analogue is the most attenuated, although the magnitude of the twist penalty is somewhat reduced in the larger RPSB structures. Quantitatively in terms of absolute intensities, however, increase dramatically in RPSB. Relative to the RPSB5 set at the same M11 level, the RPSB excitation energies are uniformly lower by $\sim 0.26 - 0.39$ eV, consistent with increased π -delocalization expected upon extending the conjugated chain. At the same time, $\Delta\mu$ is substantially larger in RPSB by $\sim 3.1 - 3.9$ D ($\sim +109\%$ to $+167\%$) across all analogues, while μ_{01} increases more modestly by $\sim 1.3 - 2.3$ D ($\sim 12 - 21\%$). The combination of a much larger $\Delta\mu$, modest increase in μ_{01} , and reduced excitation energies rationalizes the near order-of-magnitude increase in the absolute 2PA magnitude. Overall, adding the β -ionone ring strongly boosts absolute 2PA strengths primarily through enhanced charge-transfer character (larger $\Delta\mu$) and lower EE, while leaving the relative methylation/twist response largely intact.

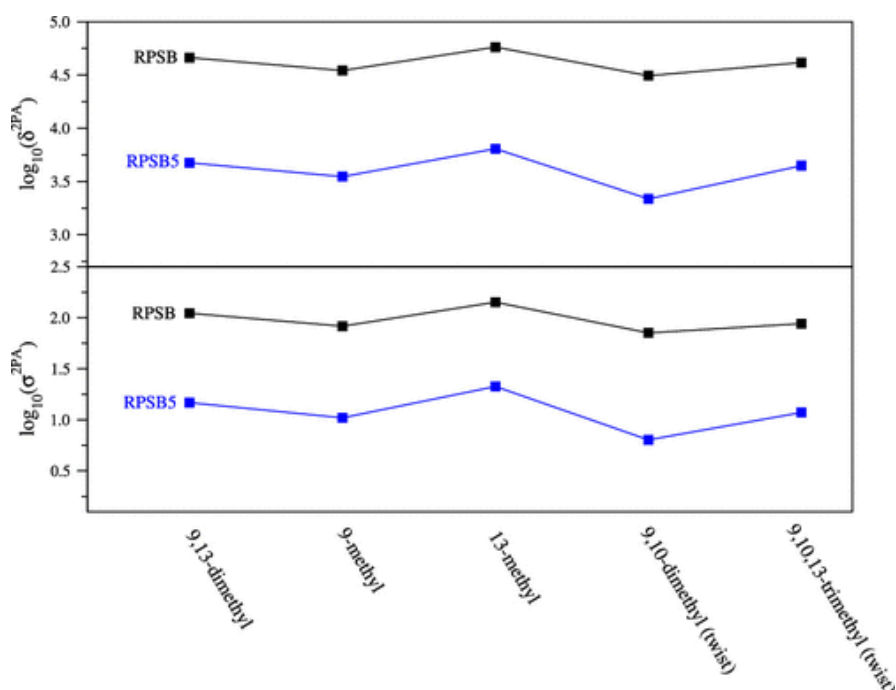


Figure 15 Comparison of the microscopic two-photon transition strengths, δ^{2PA} (top), and macroscopic two-photon absorption cross-sections, σ^{2PA} (bottom), of the RPSB and truncated RPSB5 analogue series calculated at the M11/aug-cc-pVDZ level. The RPSB models contain the β -ionone ring, whereas the RPSB5 models represent the corresponding truncated chromophores. Values are shown on base-10 logarithmic scales, with δ^{2PA} reported in a.u. and σ^{2PA} in GM.

5. Results and Discussion (Rhodopsin Protein)

5.1 QM region size convergence for 1PA/2PA in Rhodopsin

In Chapter 4, the author established a computationally feasible and internally consistent protocol for retinal-based 2PA calculations *in vacuo* by climbing the coupled-cluster hierarchy on reduced rPSB models and benchmarking multiple exchange–correlation functionals within the TD-DFT methodology against higher-level reference data. As noted in the introduction, the spectral response, for e.g. the 2PA intensity, interpreted from the chromophore structure alone may be misleading once the opsin electrostatics, local hydrogen-bond network, and embedding field begin to reshape it and its associated properties like ΔE , μ , and $\Delta\mu$. Therefore, in this chapter (Chapter 5), the author analyses the 2PA signal and associated excited-state electronic properties for full Rhodopsin models. Objective no. 4 is addressed by determining the QM region size and composition required for spectral convergence, while the influence of the QM/MM coupling scheme is also examined through selected PE, EE, and NE comparisons. Together, these modeling choices are required to obtain stable 2PA trends that can be meaningfully compared across computational methods and, where available, against experiments.

5.1.1 The Rhodopsin model

The full Rhodopsin protein models used in this chapter were adapted from the published QM/MM construction protocol reported in the Supporting Information (SI) of Olivucci *et al.* 2019 [97]. Starting from the bovine Rhodopsin crystal structure (PDB ID 1U19, chain A), the team used their proposed Automatic Rhodopsin Modelling (ARM) workflow [102] to define the chromophore cavity, assign protonation states consistent with the crystallographic pH = 6.0, place external counterions to neutralize the extracellular and intracellular surfaces (for a total of 2 Na⁺ and 6 Cl⁻). The resulting model produced neutral forms of Asp83, Glu122, Glu181 and His211 (with His211 specified as ‘HSD’ a δ -protonated form), and crystallographic waters were optimized using DOWSER. The retinal binding site, and ergo QM region, was represented with an rPSB11-Lys296 fragment (RET) comprised of the rPSB11 plus the C ϵ atom of Lys296 (54 atoms total) with a standard hydrogen link-atom treatment at its C ϵ -C δ boundary, and the remaining Lys296 atoms are fitted on the MM side. To generate an ensemble for the team’s subsequent response calculations, 10 independent replicas were produced (distinct initial velocities) via MD at room temperature in GROMACS. This is then followed by a QM/MM geometry optimization of the chromophore and protein cavity at the CASSCF/AMBER-FF94 level using MOLCAS-TINKER, electrostatic embedding (EE) of QM into MM was treated with the electrostatic potential fitted method.

One of these refined structures (specifically model 18273) by Olivucci and co-workers was subsequently used as the starting point for an additional solvated QM/MM MD simulation. This step helps generate a consistent set of equilibrated snapshots for the multiphoton response calculations reported in this chapter. The structure was converted and parameterized within GROMACS and prepared for explicit solvation and ion neutralization prior to transfer to CP2K for QM/MM MD (further MD details below). In CP2K, the QM subsystem was defined to include RET (retinal chromophore and Lys296 residue), together with the C and O atoms of Ala295 and the N, H, C α (CA), and H α (HA) atoms of Thr297, giving the QM region a total of 77 atoms. All remaining protein, solvent, and ions were thus assigned to the MM region. The QM/MM boundary was placed at the Ala295 and Thr297 cut, where the first set of frontier (M1) atoms that are directly bonded to the QM region, more specifically the C α (CA) of Ala295, and C and C β (CB) of Thr297, were treated using the standard dummy hydrogen link-atom representation in the QM description. The QM subsystem was assigned a total charge of +1, with the MM region carrying the compensating charge (ca. -1) arising from the full solvated and ionized environment.

Following the QM/MM MD simulation, ten snapshots were selected from the trajectory. For each snapshot, the simulation box was stripped to retain the complete Rhodopsin protein together with a 3 Å hydration shell of the nearest water molecules and ions, and the resulting system was transferred to the MOLCAS/LoProp [203, 204] fragmentation workflow to generate embedding potentials for subsequent response calculations in Dalton2020 [182, 183]. After the equilibration period in QM/MM MD the final 10 frames (14–23 ps) were selected as independent starting points for embedding generation, where the system was fragmented via the PyFrame 0.2.0 interface [205]. The protein was fragmented using the Molecular Fractionation with Conjugate Caps (MFCC) scheme, where each fragment contains the target residue (e.g. residue A) and a small backbone segment of the neighboring residue (e.g. residue B) while also being terminated with conjugate caps across peptide-bond cuts, minimizing artificial polarization at the fragmentation boundary [206]. The fragment's electronic structure themselves were calculated in MOLCAS at the DFT/B3LYP level with the ANO-L-VDZP basis set, followed by LoProp which derive the PE force-field terms, namely distributed permanent electrostatic moments (multipoles) and site polarizability tensors (α) for the MM region [203, 207]. The fragment properties were recombined to form the full-system PE potential for each snapshot, which was then inserted together with their corresponding coordinates into Dalton2020 for embedded response calculations. To prevent over polarization of the QM region by close MM sites, Dalton's default PE handling was employed. That is, all higher-order multipoles and polarizabilities within 1.4 Å of any QM atom were removed (set to zero), and their corresponding point charges were redistributed to nearby MM sites so that the overall MM charge was preserved. Additionally, to

account for the lack of short-range Pauli repulsion between the QM electron density and the MM environment in classical polarizable force fields, pseudopotential-type [166] ECPs were applied to all atoms within exactly 11 Å of the QM region. This prevents electron spill-out (ESO) into the MM region, which can otherwise introduce artefacts such as unphysically low-lying excited-states with orbitals extending into the embedded MM region [166, 168]. Furthermore, a 1.4 Å buffer from the QM boundary was excluded from the ECP selection, and thus the M1 frontier atoms were not assigned pseudopotentials. This buffer region helps avoid an excessively strong repulsive wall directly at the QM/MM boundary as noted by Grabarek and Andruniow [167]. Each snapshot was then processed in Dalton2020 through two distinct stages; ground-state convergence using Direct-SCF (DSCF), followed by the actual response-property evaluation, using BHandHLYP/aug-cc-pVDZ, with excitation and response properties computed up to the sixth lowest excited-state.

To assess sensitivity to QM-subsystem size, the QM region was subsequently expanded beyond M1 following the same systematic build-up scheme as Valsson *et al.* [158], using the 18 ps frame as the reference structure for defining the successive partitions. In M2, the QM region was increased from 77 to 122 atoms by adding seven nearby waters (SOL349, SOL350, SOL5782, SOL5798, SOL5812, SOL6358, SOL6402), the entirety of Glu113, and partial backbone atoms from Gly114 (N, H, C α , H α) and Leu112 (C, O), with frontier atoms at C(Gly114) and C α (Leu112). M3 extends M2 up to 128 atoms by adding a partial neutral protonated Glu181 (GLH181) (C, O, C α , H α , N, H), leaving the remainder in MM with a frontier at C β (GLH181). Finally, M4 contains 152 atoms and adds the entirety of Ala292 together with partial Pro291 (C, O) and Phe293 (C α , H α , N, H), assigning frontier atoms at C α (Pro291) and C/C β (Phe293).

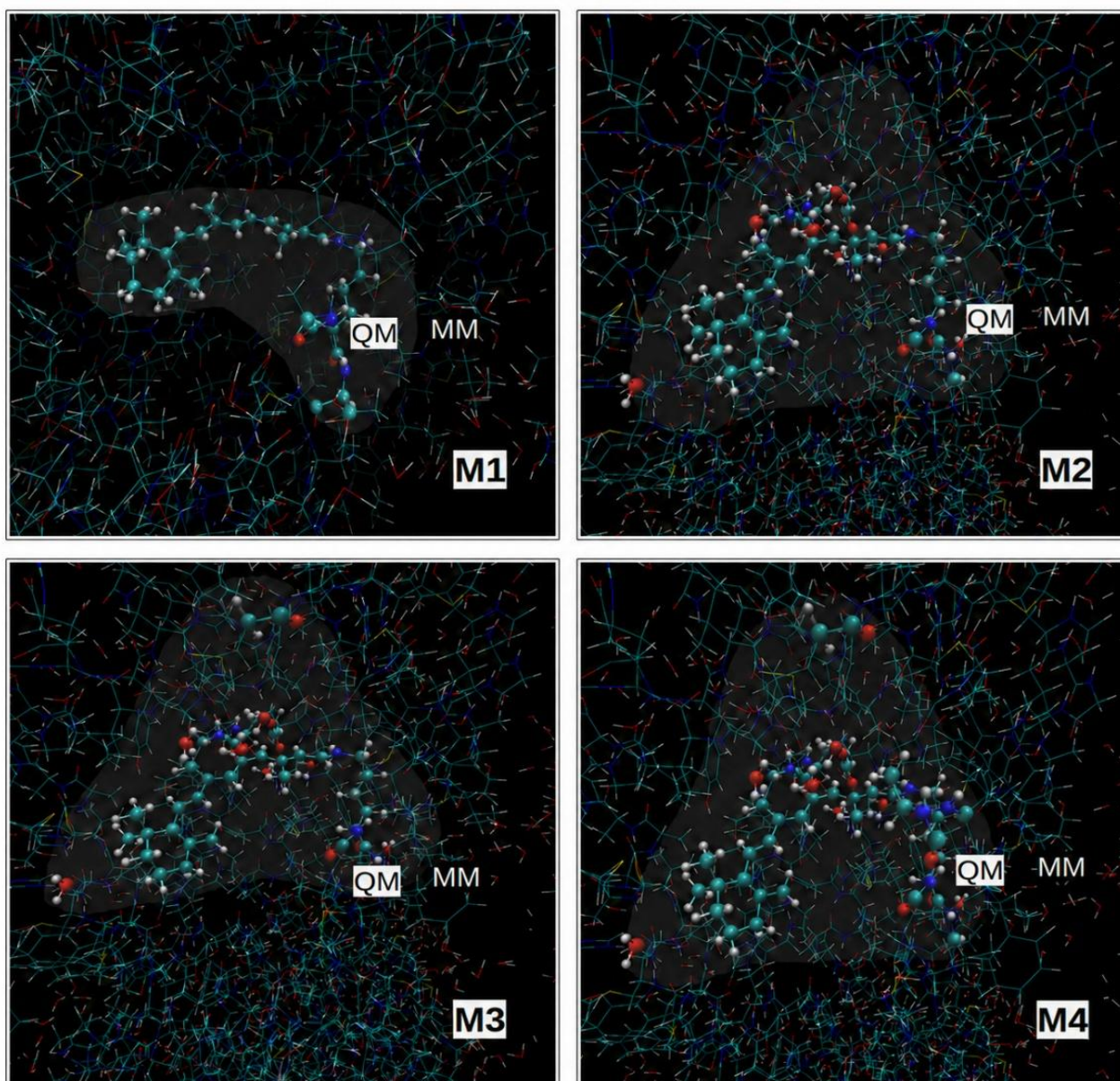


Figure 16. QM/MM partitioning schemes for the M1–M4 Rhodopsin models constructed from the 18 ps QM/MM molecular-dynamics snapshot. The QM subsystems are shown in ball-and-stick representation, while the surrounding MM environment is shown as lines. The successive panels illustrate the systematic expansion of the QM region from M1 to M4. M1 (77 atoms), M2 (122 atoms), M3 (129 atoms), M4 (148 atoms).

5.1.2 Molecular Dynamics Calculations Details

As noted above, model 18273 from the Olivucci group [97] was placed into an explicit-solvent box generated through standard GROMACS preparation steps (`editconf`, `solvate`, `grompp`, `genion`) using the AMBER94 parameters [208] and the TIP3P [209] water model. The resulting solvated system contained 13,707 water molecules together with 45 Na^+ and 49 Cl^- ions and a net MM charge of ca. -1 . To ensure force-field consistency across software and enable QM/MM in CP2K(2.6.1) [210], ParmEd(3.4.0) [211] was used (i) to convert the GROMACS AMBER94 description into a CP2K-readable AMBER94 format and (ii) to add/adjust missing Lennard–Jones

(van der Waals) parameters for the TIP3P hydrogen atom types in a manner consistent with the General Amber Force Field 2 (GAFF2) [212]. QM/MM MD was then carried out in CP2K under the canonical (NVT) ensemble, treating the QM region at the DFT/BLYP/6-31G* level within the Gaussian and Augmented Plane Waves (GAPW) formalism [213], while the MM region was described by the AMBER94 force field. The simulation protocol comprised four stages: (1) heating to 200 K for 300 steps using a CSVR thermostat, (2) heating to 250 K for an additional 300 steps, (3) heating to 300 K for yet another 300 steps, and finally (4) an NVT production run at 300 K for 23,000 steps ($\Delta t = 1$ fs for all steps) using a Nosé–Hoover thermostat; due to restarts and failed segments, the accumulated trajectory length used for snapshot selection amounted to approximately 23.5 ps.

5.1.3 Electronic Properties and Spectral calculations

Within this chapter, the three interaction treatments considered includes polarizable embedding (PE), electrostatic embedding (EE), and no embedding (NE), and differ in how the protein environment enters the electronic-structure evaluation of 2PA and associated excited properties. As noted in the methodology section, within the EE model the MM subsystem is represented purely by a set of fixed electrostatic point charges acting on the QM region. Whereas PE augments this description by including distributed multipoles and site polarizability tensors that allow the MM region to respond to the QM electric field (mutual polarization) via induced dipoles computed self-consistently during the ground-state SCF (DSCF) and subsequent response calculation (see 5.1.1/methodology). In all embedded calculations, a dispersion correction was included to improve the description of noncovalent interactions at the DFT level of theory which doesn't account for dispersion energies inherently. More specifically, Grimme's D3 correction together with Becke-Johnson (BJ) damping was applied [214], providing an additional London-dispersion contribution and improving short-range behavior via BJ damping. Operationally, this correction contributes to the total energy and therefore affects the SCF solutions of the Kohn–Sham (KS) equations in DFT. Therefore, D3 corrections with BJ damping is relevant for obtaining a physically consistent electronic structure in a complex protein environment like Rhodopsin. An effective external field (EEF) correction was also additionally applied. EEF accounts for the polarization response of the MM subsystem induced by the external electromagnetic field (e.g., the light field driving OPA or TPA), thereby incorporating local-field effects in which the environment modifies the effective perturbation experienced by the QM region. Because the correction modifies the induced electrostatic field intensity rather than introducing a frequency-dependent shift, it does not affect the excitation-energy scale (ΔE), but it can influence computed transition strengths.

To ensure that EE is compared to PE on a strictly consistent footing with respect to boundary handling, the EE charges were constructed by reusing Dalton’s default PE boundary treatment. After initiating the PE job, the Dalton output from the DSCF stage was used to extract the list of MM sites whose charges were redistributed and whose higher-order multipoles/polarizabilities were removed within 1.4 Å of the QM region. This same redistribution scheme was then copied into the EE workflow (via a custom script), along with the rest of the environment which was represented solely by fixed atomic point charges at the corresponding MM coordinates, without any explicit polarization response. NE systems on the other hand removes all MM electrostatic sites, that is, the NE considers purely the QM region for the electronic-structure calculations and treats the QM as *in vacuo*. Within EE and NE all other system setup for the QM remain the same, including the expansion of the QM region size in the M1 – M4 series described in 5.1.1. For the EE and NE variants, spectra calculations were performed in Turbomole 7.6 using RI-DFT with BHandHLYP and the aug-cc-pVDZ basis set, with D3(BJ) dispersion and no symmetry constraints (C1 symmetry). A gridsize m5 numerical integration grid was employed for exchange–correlation quadrature (a fine integration grid choice intended to reduce grid-related noise in response properties), while the RI approximation accelerates the DFT stages by using an auxiliary basis representation of electron-repulsion contributions. In practice, the workflow was organized to first a single ground-state density was generated (within the ‘ridft’ module), and state-specific properties were then evaluated using ‘escf/egrad’ for the 6 lowest excited-states.

5.2.1 The M1 Ensemble and its Averages in Polarizable Embedding

State	14PS	15PS	16PS	17PS	18PS	19PS	20PS	21PS	22PS	23PS
1	2.71 (915)	2.41 (1029)	2.77 (895)	2.65 (936)	2.93 (846)	2.68 (925)	2.53 (980)	2.59 (957)	2.57 (965)	2.52 (984)
2	3.81 (651)	3.81 (651)	4.19 (592)	3.97 (625)	4.13 (600)	4.18 (593)	4.03 (615)	4.03 (615)	3.90 (636)	4.07 (609)
3	4.60 (539)	4.35 (570)	4.40 (564)	4.48 (553)	4.77 (520)	4.40 (564)	4.23 (586)	4.57 (543)	4.24 (585)	4.09 (606)
4	4.82 (514)	4.60 (539)	4.97 (499)	4.73 (524)	4.91 (505)	4.84 (512)	4.95 (501)	4.79 (518)	4.90 (506)	4.95 (501)
5	5.50 (451)	4.93 (503)	5.53 (448)	5.42 (457)	5.12 (484)	5.41 (458)	5.13 (483)	4.87 (509)	5.02 (494)	5.09 (487)
6	5.56 (446)	5.04 (492)	5.59 (444)	5.57 (445)	5.56 (446)	5.53 (448)	5.31 (467)	5.02 (494)	5.14 (482)	5.35 (463)

State	Mean	Stdev
1	2.65 (936)	0.15
2	3.95 (628)	0.10
3	4.45 (558)	0.20
4	4.80 (517)	0.10
5	5.20 (477)	0.18
6	5.35 (464)	0.17

Table 12. S_1 - S_6 excitation energies for the ten frames of M1 ensemble in polarizable embedding, calculated at the BHandHLYP/aug-cc-pVDZ level of theory. Excitation energies, ΔE , are reported in eV, with the corresponding degenerate two-photon excitation wavelengths in nm given in parentheses. Ensemble means and standard deviations of ΔE are summarized below.

The excitation energies obtained for the M1 QM-region exhibits only a modest variation across all ten QM/MM MD frames (14 – 23 ps). Within the lowest excited-state, S_1 , excitation energy spans from 2.41 eV to 2.93 eV, or a 0.52 eV spread across all frames as seen in table 12 above. Interestingly, the higher states here are not uniformly more variable. The higher S_2 – S_6 states exhibit a spread of 0.37 eV to 0.68 eV, with the smallest spread being S_4 and the largest being S_3 . Standard deviation (σ) across all frames further supports this lack of increase in variability in these higher lying states. Although S_1 is not the most stable of states, it records a rather low σ of 0.15 eV, meanwhile S_4 and S_3 records a σ of 0.10 and 0.20 eV respectively. This means across all states and all frames, standard deviation spans only ~ 0.10 – 0.20 eV, below the typical TD-DFT excitation energy errors of ~ 0.20 – 0.40 eV, as noted by others across a range of xcf [142, 215, 216]. And hence these low fluctuations demonstrated in the lower-lying excited-states like S_1 are preserved across the sampled ensemble. In addition, the energetic ordering of higher states remains unchanged, indicating that thermal fluctuations of the protein environment introduce only weak perturbations to the electronic structure. Furthermore, in table 12 the averaged S_1 state recorded a ΔE of 2.65 eV (468 nm for single-photon excitations and 936 nm for degenerate double-photon) and when directly compared to experimental results of about 490 – 500 nm [12, 217], a blue-shift of only about 20-30 nm can be observed. This average is also in agreement with experimental 2 photon-activation in the near-IR spectrum (~ 900 – 1000 nm) [16]. In comparison to other theoretical works in QM/MM Rhodopsin, accurate reproduction of experimental absorption maximum depends both an appropriate QM region and configurational sampling rather than reliance on a single structure [158, 218]. The next state, S_2 averages at 3.95 eV and therefore has quite a large gap of 1.30 eV from the previous S_1 state. In fact, among all the energy levels investigated here (up to S_6), this jump is by far the largest with S_6 averaging out at 5.35 eV, or a mere 1.40 eV jump from S_2 . The energy gaps between all these higher states are rather compressed, as seen in table 12 above, exhibiting only a 0.15 eV to 0.50 eV gap between each of them. On closer inspection of the higher-lying manifold reveals that for a subset of configurations (notably 14–19PS), the energy spacing tends to favor a loose S_3 – S_4 and S_5 – S_6 pairing, whereas in later frames this separation becomes less distinct, with S_4 – S_6 forming a more collectively compressed group and S_2 and S_3 remaining more energetically isolated. This reduced spacing between the higher-lying excited-states indicates that this region of the manifold is more susceptible to small structural perturbations and state mixing than the well-isolated S_1 state. In the present ensemble, however, this increased density does not lead to any change in state ordering as noted earlier.

OPA State	14PS	15PS	16PS	17PS	18PS	19PS	20PS	21PS	22PS	23PS
1	0.983/9.770	1.090/10.925	1.001/9.770	1.096/10.452	0.949/9.238	0.978/9.817	1.174/11.053	1.089/10.525	1.080/10.538	1.041/10.442
2	0.160/3.330	0.149/3.210	0.193/3.485	0.101/2.593	0.156/3.159	0.152/3.095	0.222/3.810	0.132/2.936	0.214/3.809	0.269/4.177
3	0.202/3.408	0.093/2.375	0.219/3.628	0.044/1.602	0.118/2.554	0.229/3.700	0.207/3.593	0.112/2.537	0.146/3.013	0.096/2.493
4	0.069/1.944	0.088/2.248	0.031/1.292	0.180/3.168	0.152/2.862	0.046/1.588	0.021/1.046	0.089/2.220	0.026/1.177	0.071/1.949
5	0.039/1.372	0.000/0.042	0.001/0.219	0.000/0.109	0.000/0.071	0.001/0.237	0.001/0.237	0.000/0.081	0.000/0.087	0.001/0.255
6	0.036/1.312	0.021/1.053	0.041/1.386	0.070/1.826	0.039/1.365	0.041/1.393	0.000/0.058	0.000/0.063	0.000/0.048	0.000/0.066

State	Mean	Stdev
1	1.048/10.253	0.070/0.575
2	0.175/3.360	0.049/0.471
3	0.147/2.890	0.064/0.691
4	0.077/1.949	0.053/0.704
5	0.004/0.271	0.012/0.396
6	0.025/0.857	0.024/0.711

Table 13. Oscillator strengths, f , (in length) and transition dipole moments, μ_{on} , for the S_1 - S_6 states of the ten frames in the M1 ensemble in polarizable embedding, calculated at the BHandHLYP/aug-cc-pVDZ level of theory. Each entry is reported as f/μ_{on} , where f is dimensionless and μ_{on} is given in Debye. Ensemble means and standard deviations are summarized below.

In contrast to the relatively small variation observed for the excitation energies, one-photon transition properties exhibit a greater sensitivity to the sampled configurations within the M1 ensemble. The lowest excited-state S_1 remains the dominant transition across the samples, at an average oscillator strength of 1.05 with a standard deviation of 0.07. Correspondingly, the transition dipole moment is at 10.25 D with its standard deviation at 0.58 D, as seen in table 13 above. This shows that the bright one-photon state is consistently localized in S_1 for all sampled configurations which corresponds to a π - π^* transition of the retinal chromophore. The higher-lying states on the other hand are substantially weaker and display larger relative fluctuations in both oscillator strengths and consequently transition dipole moments. S_2 and S_3 retain moderate intensities while S_4 - S_6 become weak to nearly dark in most of the samples. At the S_2 - S_4 states, variability is on the order of tens of percent relative to their mean values (coefficient of variation), on the other hand variability is most pronounced for S_5 and S_6 states in which the standard deviation matches or even exceeds the mean by approximately three times. A closer inspection of Table 13 further reveals that the relative ordering of oscillator strengths within the S_2 - S_4 manifold also varies across the ensemble. Depending on the sampled configuration, the dominant contribution (second to the S_1 of course) among these states shifts, with S_2 , S_3 , or occasionally S_4 (like in 17PS) carrying comparable intensity at approximately ~ 0.200 . This behavior reflects a redistribution of transition intensity within a group of closely spaced excited-states. Furthermore, this redistribution is also consistent with the reduced spacing within the higher-lying manifold as seen in ΔE separation, which increases the sensitivity of these states to local structural variations and changes in transition character.

In summary, similarly to ΔE the protein environment introduces only moderate perturbations to the 1PA properties without altering the qualitative structure of the low-lying excited-state manifold. The

preservation of state ordering, that is, the consistent localization of the bright transition in S_1 , shows that the intrinsic electronic structure of the retinal chromophore remains the dominant factor in governing the dominant spectral response. This is consistent with the trends established in Chapter 4 and previous Rhodopsin QM/MM [158, 218], where the excitation energies of reduced retinal models were shown to depend primarily on the chromophore geometry and electronic structure *in vacuo*, while changes in electronic-structure methodology produced mainly systematic quantitative shifts rather than qualitative changes in the low-lying excited-state behavior.

State	14PS		15PS		16PS		17PS		18PS	
	δ^{2PA}	σ^{2PA}	δ^{2PA}	σ^{2PA}	δ^{2PA}	σ^{2PA}	δ^{2PA}	σ^{2PA}	δ^{2PA}	σ^{2PA}
1	19900	53.5	7390	15.7	13500	38.0	4730	12.2	13800	43.4
2	65800	349.9	126000	670.1	79700	512.6	64800	374.2	40100	250.6
3	67100	520.2	1130000	7833.6	120000	851.1	15200	111.8	59900	499.3
4	26000	221.3	929000	7201.7	70300	636.2	11300	92.6	17900	158.1
5	58700000	650526.8	1320000	11753.6	4.18E+11	4.68E+09	145000	1560.5	147	1.4
6	6610000	74860.5	80400	748.2	98800000	1131051.2	27500	312.6	161000	1823.4

State	19PS		20PS		21PS		22PS		23PS	
	δ^{2PA}	σ^{2PA}	δ^{2PA}	σ^{2PA}	δ^{2PA}	σ^{2PA}	δ^{2PA}	σ^{2PA}	δ^{2PA}	σ^{2PA}
1	12200	32.1	13400	31.4	8310	20.4	10800	26.13	16000	37.2
2	83800	536.4	134000	797.3	116000	690.2	107000	596.23	13700	83.1
3	50700	359.6	52100	341.5	64600	494.3	86500	569.7	200000	1225.7
4	93400	801.6	14000000	125672.5	69900	587.6	3660000	32194	242000000	2172338.3
5	52200000	559714.7	65100000	627649.8	14000	121.6	1990000	18372.24	158000000	1499665.3
6	527000	5904.2	149000	1539.1	671000	6194.9	9040000000	87497644.4	31900	334.5

State	ΔE	δ^{2PA}		λ	σ^{2PA}		Removed States
	Mean(e.V)	Mean(a.u.)	Stdev	Mean(nm)	Mean(GM)	Stdev	
1	2.65	12003	4410	936	31.01	12.80	
2	3.95	83090	38581	628	486.06	220.26	
3	4.45	184610	335949	558	1280.67	2322.29	
4	4.80	173971	334379	517	1385.57	2579	20/22/23PS
5	5.20	53049	79933	477	561.19	868	17/18/21PS*
6	5.35	235400	257047	464	2408.12	2553	14/16/22PS

Table 14. Two-photon absorption properties for the S_1 - S_6 states of the ten frames in the M1 ensemble under polarizable embedding, calculated at the BHandHLYP/aug-cc-pVDZ level of theory. Two-photon transition strengths, δ^{2PA} , are reported in a.u., and two-photon absorption cross-sections, σ^{2PA} , in GM. The lower table summarizes the mean excitation energies in eV, the corresponding degenerate two-photon excitation wavelengths in nm, and the ensemble means and standard deviations of δ^{2PA} and σ^{2PA} . The “Removed States” column identifies frames excluded from the state-specific ensemble averages. For S_5 , the asterisk indicates that only frames 17PS, 18PS, and 21PS were included in the average, while all remaining frames were excluded.

The 2PA response of the M1 ensemble demonstrates significantly larger variability across all configurations than what was observed in excitation energies or 1PA properties as seen from the standard deviation in table 14 above. Despite this variability, the averaged spectrum develops a well-defined multi-peak structure spanning approximately 550 – 950 nm, indicating that the observed features are consistently recovered across the ensemble. So, while individual transitions are highly configuration-dependent, the averaged spectrum is a structured and physically meaningful. The lowest excited-state (S_1) produces a weak but clearly resolved contribution in the near-IR region, with an averaged σ^{2PA} of ~ 31 GM at ~ 936 nm. Although it is small in magnitude, the S_1 band is consistently present across the configurations and aligns well with the reference position at 950 nm

[97] as seen in fig. 18 below. Furthermore, the S_1 is also the least variable among all averaged states with a coefficient of variation of $\sim 41\%$, lowest in both relative and absolute values. The S_2 state gives rise to a significantly stronger but slightly less stable response, with an averaged σ^{2PA} of ~ 486 GM at ~ 628 nm and slightly higher variation across frames as seen from the coefficient of variation being $\sim 45\%$. The S_3 state forms one of the dominant averaged peaks in the structured spectral region, reaching ~ 1281 GM at ~ 558 nm. However, S_3 exhibits among the highest configuration dependence of all states, as reflected by the large standard deviation resulting in a coefficient of variation of over 180% , indicating a very broad distribution of intensities across the ensemble. Interestingly, this dominant intensity of the S_3 average is primarily driven by a few specific configurations, most notably the 15PS frame and, to a lesser extent the 23PS frame, which exhibit a significantly enhanced σ^{2PA} signal. The statistical amplification by these high-intensity frames is consistent with configuration-dependent state mixing within the compressed excited-state manifold observed in ΔE , and 1PA's increased variability in higher excited-states. Across the ensemble, S_4 to S_6 states contribute primarily to spectral peaks beyond ~ 550 nm and display increasing variability without forming consistently dominant averaged features. Therefore, these higher states unsurprisingly also show reduced statistical reliability, with coefficient of variations exceeding 100% , where S_4 reaches the highest at just over 186% , suggesting they are also strongly affected by configuration-dependent state mixing like in S_3 , as well as potential resonance enhancements. Fig. 17 below illustrates this variability across configurations, with individual frames shown alongside the averaged spectrum (black).

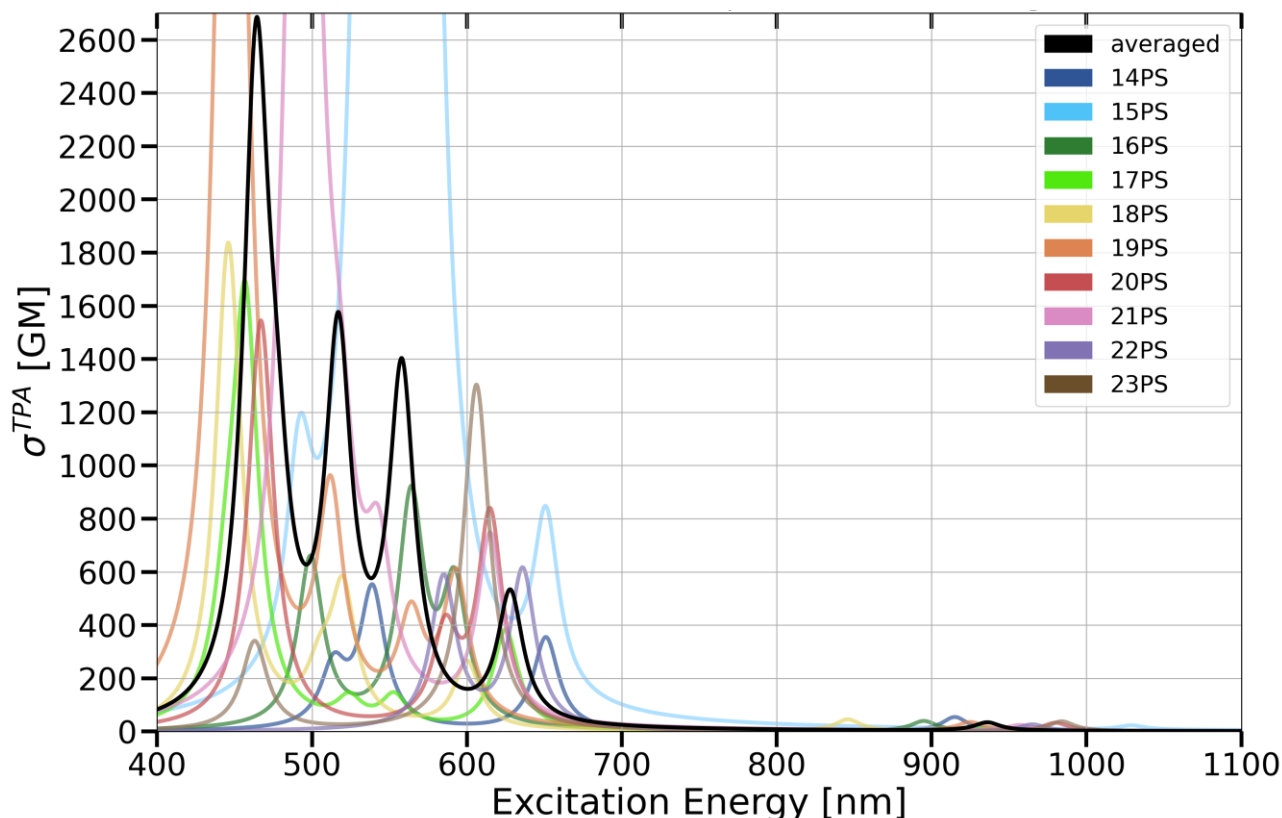


Figure 17. Snapshot resolved and ensemble-averaged two-photon absorption spectra through the S_6 state for the M1 ensemble in polarizable embedding, calculated at the BHandHLYP/aug-cc-pVDZ level of theory. The ten individual QM/MM MD frames are shown by the colored curves, and the ensemble average by the black curve. Two-photon absorption cross-sections, σ^{2PA} , are reported in GM and plotted against the degenerate two-photon excitation wavelength in nm. Lorentzian broadening was applied directly on the wavelength scale, resulting in a fixed width in nm.

The M1 averaged spectrum reproduces the three main spectral features observed in the reference calculations by Gholami *et al.* [97] at the XMCQDPT2, with peaks at ~936 nm, ~628 nm, and ~558 nm, compared to the reference 950 nm, 678 nm, and 644 nm (Fig. 18 below). The near-IR S_1 band is in good agreement in terms of quantitative position but is significantly underestimated in intensity (31 GM vs 472 GM). In contrast, the S_2 and S_3 regions are enhanced relative to the reference, with S_2 increasing from 231 GM to ~486 GM and S_3 from 771 GM to ~1281 GM. This difference is also reflected in the spectral shape, where the M1 averaged ensemble resolves three distinct features in the $S_1 - S_3$ region, while the reference spectrum partially merges the S_2 and S_3 contributions due to spectral broadening (Fig. 18). The reduced S_1 intensity and enhanced $S_2 - S_3$ responses are qualitatively consistent with the BHandHLYP behavior observed for the reduced PSB models in Chapter 4, where part of the 2PA intensity was redistributed from S_1 toward higher computed states. The higher computed states, particularly $S_4 - S_6$, display substantially greater variability across the ensemble and reduced statistical reliability. This suggests that these states are more sensitive to the sampled protein configurations and may be affected by configuration-dependent state mixing or

resonance enhancement. However, the present results do not isolate the underlying origin of this increased variability.

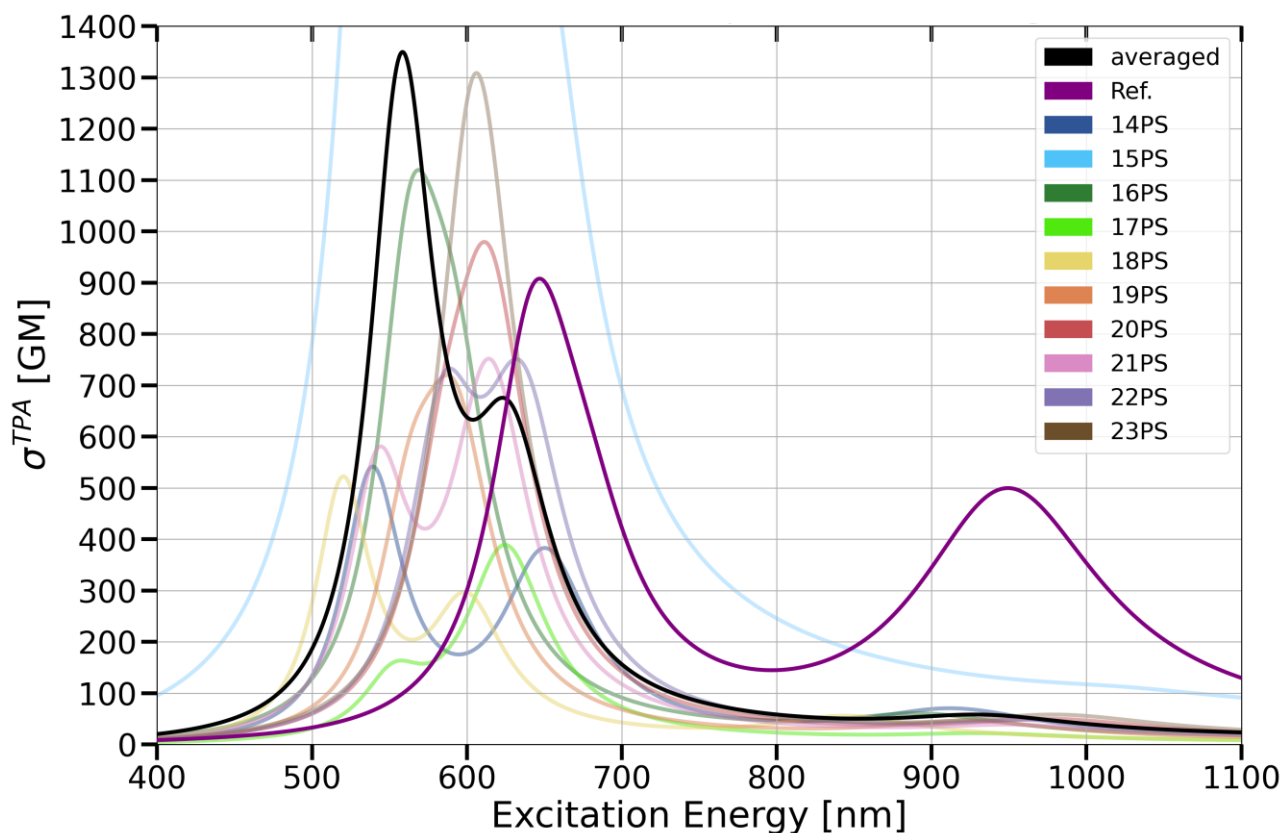


Figure 18. Snapshot resolved and ensemble averaged 2PA spectra through the S_3 state for the M1 ensemble in polarizable embedding, compared with the reference XMCQDPT2 spectrum of Gholami et al. [97]. The ten individual QM/MM MD frames are shown together with the ensemble average and reference spectrum. 2PA cross-sections, σ^{2PA} , are reported in GM and plotted against the degenerate two-photon excitation wavelength in nm. Lorentzian broadening was applied in the energy domain using a half-width at half-maximum of 0.1 eV.³

In summary, the M1 ensemble retains a structured 2PA spectral profile despite strong configuration dependence, with three consistent features across the $S_1 - S_3$ region. While the S_1 contribution is well reproduced in energy, its intensity is underestimated, with spectral weight shifted toward S_2 and S_3 , consistent with TD-DFT behavior observed in Chapter 4. Unlike the reduced PSB models, no instabilities such as negative or unrecoverable σ^{2PA} values are observed in the lower excited-states. Instead, these effects are shifted to higher states ($S_4 - S_6$), where increased variability reflects configuration-dependent state mixing and possible resonance contributions. σ^{2PA} is governed

⁴ The previous spectra within this work were generated by applying Lorentzian/Gaussian broadening directly along the wavelength (nm) axis using a width parameter corresponding to a HWHM of 0.1 eV, resulting in fixed peak widths across the spectrum. In contrast, the current spectra here was broadened in energy (eV) using a Lorentzian function with the same HWHM of 0.1 eV. When represented on a wavelength scale, this leads to a nonlinear variation in peak widths, with features becoming increasingly broadened at longer wavelengths.

primarily by energy spacing, transition character, and configuration-dependent state mixing, making it the most configuration-dependent property among those studied.

5.2.2 The M2 Ensemble and its Averages in Polarizable Embedding

The excitation energies obtained for the M2 QM-region similarly exhibit only modest variation across all ten frames (14 – 23PS). Within the lowest excited-state, S_1 , excitation energy spans from 2.51 eV to 2.94 eV, or a total spread of 0.43 eV across the ensemble as seen in Table S4. Similarly to M1, the higher excited-states are not uniformly more variable, with spreads ranging from 0.32 eV to 0.62 eV across the $S_2 - S_6$ manifold. Standard deviations remain confined to only $\sim 0.11 - 0.20$ eV, again, below the typical TD-DFT excitation energy errors [142, 215, 216]. Consequently, the relatively low configurational fluctuations and preserved state ordering observed in M1 are also retained in the larger M2 QM-region, indicating that thermal fluctuations and QM-region expansion affect the low-lying electronic structure only moderately. The averaged S_1 state records a ΔE of 2.68 eV (463 nm for single-photon excitation and 929 nm for degenerate two-photon excitation), corresponding to a moderate blue-shift of roughly 25 – 35 nm relative to the experimental absorption maximum. In fact, the M2 expansion is on average only 0.03 eV larger than M1 in the S_1 state. The relatively small shift between M1 and M2 further suggests that the fundamental low-lying excitation is already reasonably captured within the smaller M1 QM-region.

Moving to the higher excited-states, S_2 averages out at 4.04 eV, giving rise to a large energetic separation of 1.36 eV from the isolated S_1 state. In contrast, the remaining $S_2 - S_6$ manifold remains quite compressed, with neighboring state separations ranging only between 0.16 – 0.44 eV. Compared to M1, the higher-lying manifold in M2 appears less separated into persistent $S_3 - S_4$ and $S_5 - S_6$ pairings. While some frames still show local pairings, such as $S_3 - S_4$ in 14PS and 20PS, or $S_5 - S_6$ in 20PS and 23PS, other frames show a more compressed manifold, most clearly in 18PS and 21PS where $S_3 - S_6$ lie within relatively small energetic separations. This reduced spacing suggests that the higher excited-states remain considerably more sensitive to small structural perturbations and potential state mixing than the well-isolated S_1 state. Expanding the QM-region from M1 to M2 therefore appears to preserve the fundamental low-lying electronic structure while simultaneously increasing the compression of the higher excited-state manifold, and no state reordering is observed within the present (M2) ensemble neither.

Similarly to M1 ensemble, the one-photon transition properties within M2 preserve the qualitative low-lying bright peak while exhibiting much larger relative fluctuations across the ensemble than the

excitation energies themselves, as seen from the larger standard deviations relative to their averages in Table S5. Nevertheless, the overall qualitative behavior remains highly consistent with the M1 ensemble. The S_1 state remains the dominant transition across all sampled configurations, with an averaged oscillator strength of 1.03 and corresponding transition dipole moment of 10.09 D, both exhibiting only modest variation throughout the ensemble. Just like in M1, the bright one-photon response remains strongly localized within the low-lying π - π^* excitation despite the expanded QM-region. On the other hand, higher excited-states remain substantially weaker and exhibit increasingly larger relative fluctuations in oscillator strengths and corresponding μ_{01} . S_2 and S_3 continue to retain moderate intensities throughout the ensemble, however their relative ordering fluctuates considerably depending on the sampled configuration, similarly to M1. In several frames S_2 remains the dominant brighter contribution (not withstanding S_1), whereas in others, S_3 becomes comparably bright or even exceeds S_2 . Compared to M1, however, the variability of S_3 also increases moderately, further reflecting the increased redistribution of intensity in these higher states. The distinction between S_4 and S_5 also becomes less well-defined throughout the ensemble, with both states remaining weak and exhibiting standard deviations comparable to their averaged intensities, indicating very limited statistical reliability, while S_6 remains predominantly dark across most configurations. These fluctuations are once again consistent with the reduced energetic spacing observed within the $S_2 - S_6$ manifold in ΔE . As the higher excited-states become more collectively compressed in M2, intensity redistribution between neighboring states becomes increasingly sensitive to local structural perturbations. Nevertheless, despite this increased variability within the higher manifold, the qualitative low-lying one-photon structure remains preserved, with S_1 consistently dominating the spectral response throughout the ensemble. Expanding the QM-region from M1 to M2 therefore preserves the fundamental low-lying transition character while simultaneously increasing the collective mixing susceptibility of the higher excited-states.

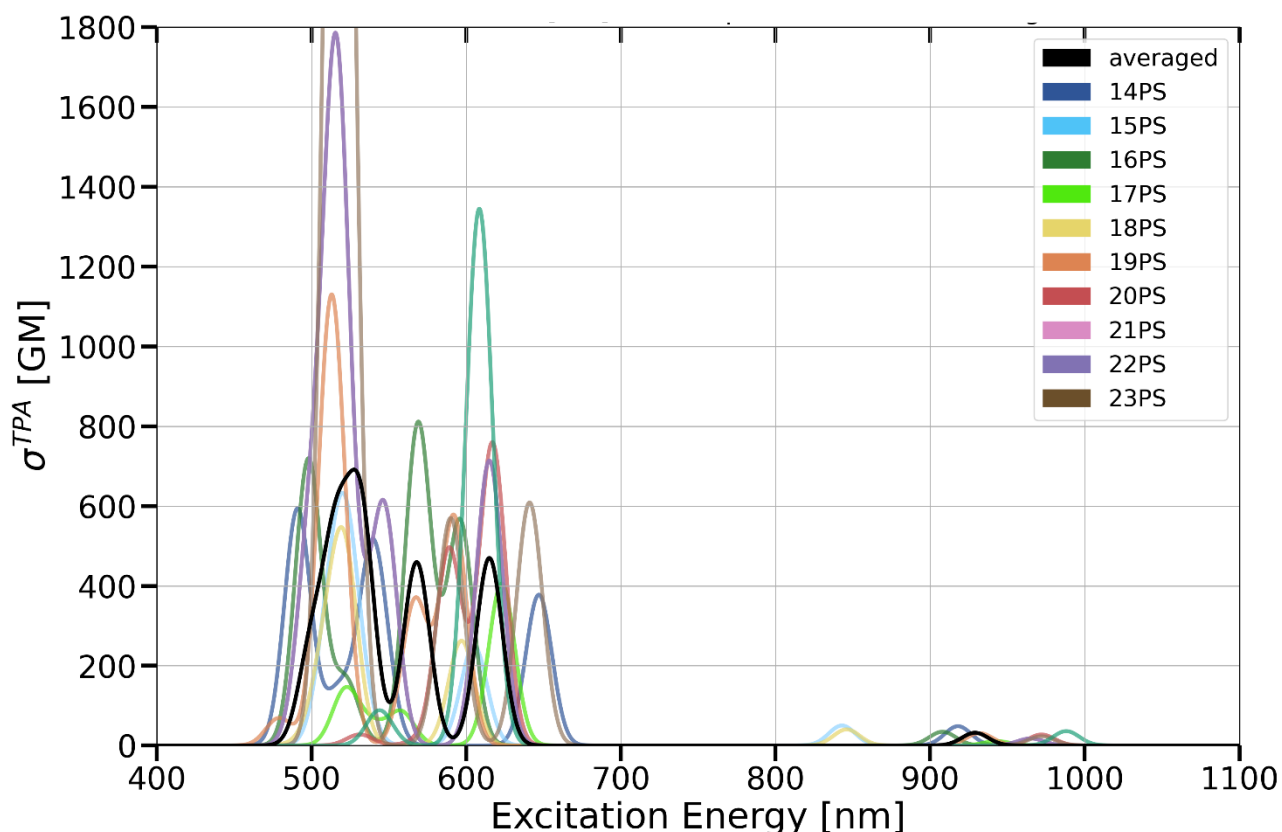


Figure 19. Snapshot resolved and ensemble averaged 2PA spectra through to the S_6 state for the M2 ensemble in polarizable embedding, calculated at the BHandHLYP/aug-cc-pVDZ level of theory. The ten individual QM/MM MD frames are shown by the colored curves, and the ensemble average by the black curve. Two-photon absorption cross-sections, σ^{2PA} , are reported in GM and plotted against the degenerate two-photon excitation wavelength in nm. Lorentzian broadening was applied directly on the wavelength scale, resulting in a fixed width in nm.

The 2PA response within the M2 ensemble remains the most configuration-dependent property investigated in this sub-chapter, similarly to the 2PA behavior previously observed in M1. Nevertheless, despite the large variability observed across individual configurations, the lower excited-states continue to produce consistent 2PA signals throughout the ensemble. As a result, the averaged spectrum retains a structured multi-peak profile despite the large variability between individual configurations. As seen in figure 19 and Table S6 of the supporting information, the S_1 state produces a weak but consistently resolved contribution in the near-IR region, averaging at about 32 GM and 929 nm with a relatively modest standard deviation of only 12.66 GM. Similarly to M1, this state remains both the weakest yet most statistically stable transition throughout the entire ensemble. The higher excited-states exhibit substantially stronger 2PA responses although with significantly larger configurational fluctuations reflected in their standard deviations. The S_2 state averages at approximately 470 GM at 615 nm with a standard deviation of ~ 202 GM, it is among the more statistically stable higher-lying states within the ensemble and consistently contributing to the dominant averaged spectral region. In contrast, the neighboring S_3 and S_4 states display considerably larger variability. While S_3 averages at approximately 460 GM at 568 nm, its standard deviation

reaches ~ 351 GM, reflecting strong configuration dependence while simultaneously partially overlapping with the neighboring S_2 contribution within the averaged spectrum..

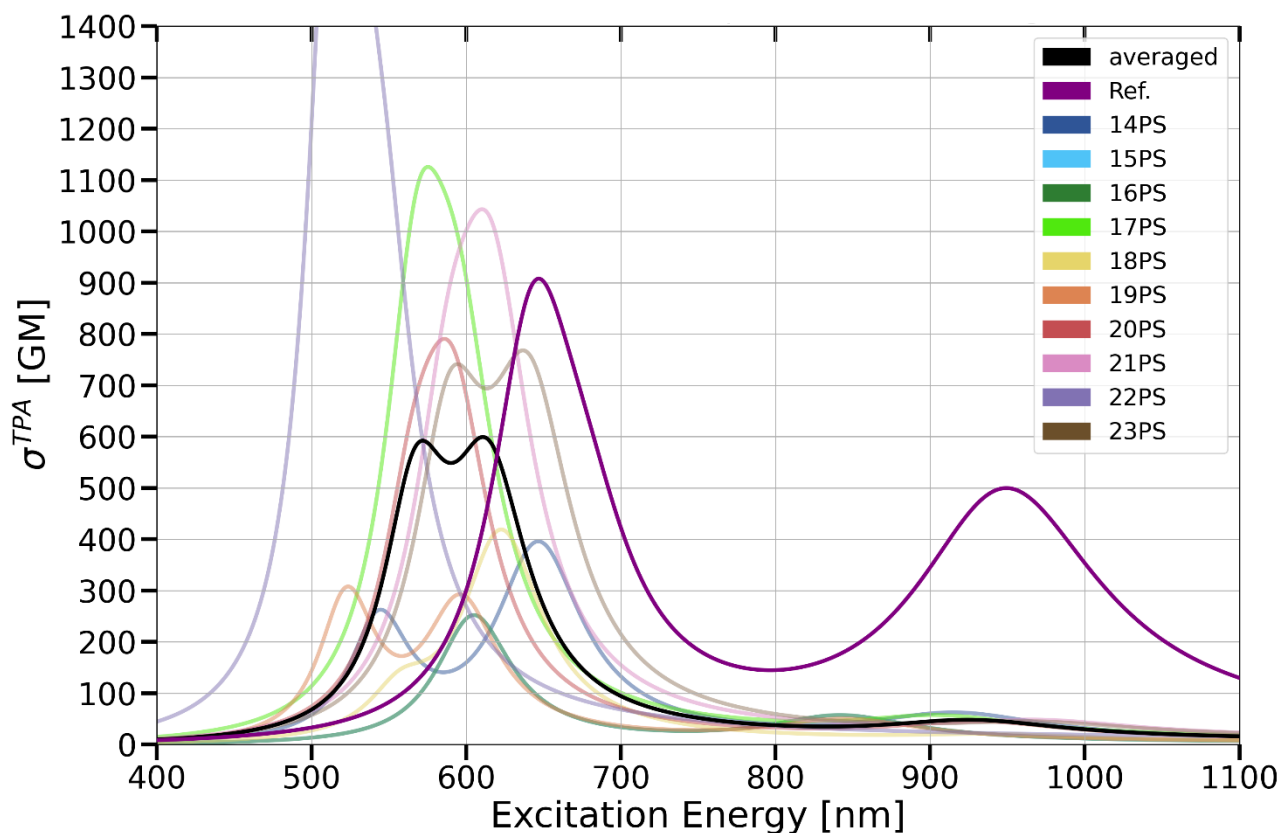


Figure 20. Snapshot resolved and ensemble averaged 2PA spectra through the S_3 state for the M2 ensemble in polarizable embedding, compared with the reference XMCQDPT2 spectrum of Gholami et al. [97]. The ten individual QM/MM MD frames are shown together with the ensemble average and reference spectrum. 2PA cross-sections, σ^{2PA} , are reported in GM and plotted against the degenerate two-photon excitation wavelength in nm. Lorentzian broadening was applied in the energy domain using a half-width at half-maximum of 0.1 eV, producing wavelength-dependent peak widths when represented on the wavelength scale.

This variability becomes even more pronounced for S_4 , which forms the dominant averaged feature within the M2 model at approximately 572 GM and 532 nm, while simultaneously exhibiting an exceptionally large standard deviation of ~ 1232 GM, exceeding twice its averaged intensity. A closer inspection of both individual spectral profiles in figure 19 and Table S6 of the SI reveals that this behavior is primarily driven by several strongly enhanced frames, most notably 22PS and, to a lesser extent, 19PS, which exhibit substantially amplified σ^{2PA} responses relative to the remainder of the ensemble. Similar behavior becomes increasingly severe for S_5 and S_6 , where several configurations produced extremely enhanced responses that were excluded from the averaged analysis to avoid complete statistical domination by a small subset of frames. In particular, the excluded S_5 frames (22PS and 23PS) and S_6 frames (16PS, 17PS, 22PS, and 23PS) all occur within highly compressed regions of the higher excited-state manifold, where energetic separations between neighboring states

frequently approach only $\sim 0.10 - 0.20$ eV. Furthermore, several of these amplified higher state contributions occur within the $\sim 4.8 - 5.30$ eV region, approximately twice the energy of the isolated S_1 excitation ($2.51 - 2.73$ eV), consistent with conditions favorable for near-resonant enhancement in degenerate 2PA. Compared to M1, where the lower $S_1 - S_3$ region remained comparatively structured and consistent despite the strongly amplified S_3 contribution observed in 15PS, the M2 ensemble instead exhibits a broader and more collectively distributed higher-energy region throughout the $S_4 - S_6$ manifold.

In M1, the enhanced S_3 response in 15PS likely reflects the unusually reduced S_1 excitation energy within that configuration, increasing proximity to near-resonant conditions for higher excited-state transitions. Interestingly, several problematic configurations recur across both ensembles, most notably 22PS and 23PS, which repeatedly exhibit amplified higher state responses and outlier behavior across multiple excited-states. Likewise, 16PS already exhibits an extremely amplified S_6 response in M1 and remains strongly enhanced in M2, while 17PS shifts from a more pronounced S_5 enhancement in M1 toward stronger S_6 amplification in M2. The increased instability within the higher excited-states is also reflected statistically in the averaged responses. While S_1 and S_2 remain comparatively stable with coefficients of variation of only $\sim 40\%$ and $\sim 43\%$ respectively, S_3 already approaches $\sim 76\%$, whereas S_4 exceeds 200% , indicating that the variability within the higher state manifold becomes comparable to or substantially larger than the averaged intensities themselves. This behavior remains consistent with the compressed energetic structure established in the ΔE analysis, where higher excited-state separations frequently approach only $\sim 0.10 - 0.20$ eV. Furthermore, several of the amplified S_5 and S_6 contributions occur within the $\sim 4.9 - 5.1$ eV region, approximately twice the energy of the isolated S_1 excitation (~ 2.68 eV), consistent with conditions favorable for near-resonant enhancement in degenerate 2PA. Consequently, unlike M1 where the higher excited-states still retained partial $S_3 - S_4$ and $S_5 - S_6$ subgroup separation throughout portions of the ensemble, the M2 manifold appears more collectively compressed across $S_4 - S_6$. This suggests that amplified higher state responses are more strongly associated with compressed energetic structure and possible near-resonant enhancement effects rather than isolated statistical noise.

5.2.3 Single-Frame QM Region Expansion and Spectral Convergence

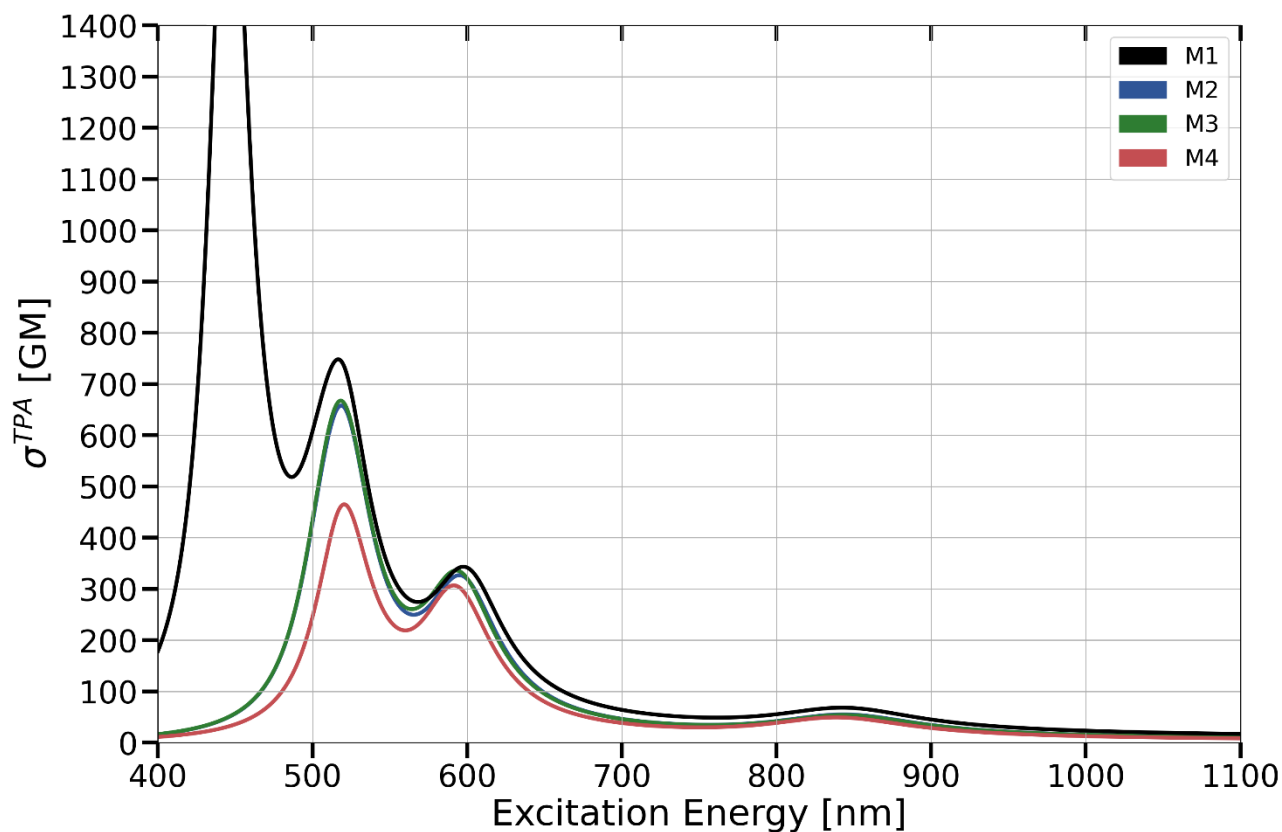


Figure 21. 2PA spectra through the S_6 state for the 18 ps Rhodopsin snapshot calculated using the M1 – M4 QM-region definitions in polarizable embedding at the BHandHLYP/aug-cc-pVDZ level of theory. The models represent the successive expansion of the QM subsystem from the smallest M1 region to the largest M4 region. 2PA cross-sections, σ^{2PA} , are reported in GM and plotted against the degenerate two-photon excitation wavelength in nm. Lorentzian broadening was applied in the energy domain using a half-width at half-maximum of 0.1 eV.

Expansion of the QM region beyond the original M1 model reveals that the dominant low-lying excited-state structure is already largely converged within the smaller QM regions. Across the polarizable embedding models, the S_1 excitation changes only modestly from 2.93 eV in M1 to approximately 2.95 eV in M4, while the corresponding low-lying one-photon properties remain similarly stable. This behavior remains consistent with the conclusions established previously for the configurational ensembles, where the dominant retinal excitation remained comparatively robust despite thermal fluctuations and QM-region expansion. Therefore, increasing the QM subsystem beyond M1 does not fundamentally alter the dominant low-energy electronic structure, suggesting that the smaller M1 model already captures much of the essential low-lying physics governing Rhodopsin excitation as seen in Fig. 21 above and quantified in table S7.

Unsurprisingly, substantially larger differences emerge within the higher excited-state manifold, particularly when examining two-photon response. While expansion from M1 toward M4 modifies excitation energies only moderately, the associated δ^{2PA} responses undergo substantially stronger redistribution between neighboring states. Within the PE calculations, higher excited-states remain comparatively compressed relative to the isolated S_1 state throughout the M2 – M4 models, where neighboring states frequently remain separated by only a few tenths of an eV while simultaneously exhibiting large changes in response magnitude. Unlike M1, where the strongest $S_1 - S_6$ response is dominated by the enhanced S_6 contribution, expansion toward larger QM regions does not produce progressively more stable higher state behavior. Instead, the dominant response becomes redistributed between neighboring states. M2 and M3 exhibit a more compact group of moderate peaks within the 500 – 600 nm region, whereas M4 shifts the strongest contribution back toward S_6 near 520 nm. Therefore, increasing the QM region primarily redistributes higher state responses rather than eliminating the strong higher state sensitivity observed throughout Chapter 5. This behavior remains consistent with the interpretation developed throughout Chapter 5, where amplified higher state response appears more strongly associated with compressed manifolds and possible near-resonant conditions rather than QM-region size alone. Consequently, expanding the QM region beyond M1 primarily changes where higher state response appears within the manifold rather than eliminating the strong configurational sensitivity observed throughout Chapter 5.

Finally, to separate the effects of environmental treatment from those of QM-region size, ensemble-averaged PE, EE, and NE results were compared for both M1 and M2 using the same ten snapshots. Not all excited states were recovered for every snapshot, and states showing strong near-resonant enhancement were excluded from the averages. The ensemble-averaged results, including missing and excluded values, are provided in Table S3. The low-energy excitation profile remains broadly recognizable across the embedding schemes, although individual states respond differently to the environmental treatment. In both M1 and M2, the S_1 excitation energy changes only modestly, whereas S_2 and S_3 are more sensitive. NE produces the largest red-shifts, particularly for S_2 . The 2PA response is considerably more sensitive and does not follow a uniform trend across the states. In both models, the S_1 cross-section increases from PE to EE and NE. For S_2 , PE and EE give comparable responses, whereas NE reduces the intensity, most strongly in M2. S_3 behaves differently, with PE producing the largest response in both QM regions and EE and NE giving substantially lower values. Above S_3 , the distribution of 2PA intensity varies strongly with both the embedding treatment and QM-region size. The particularly large S_5 and S_6 responses obtained with NE in M1 should be interpreted cautiously, since this energy region was affected by frequent near-resonant enhancement in multiple frames, even after the most strongly affected values were excluded. Overall, the

embedding treatment has a comparatively modest effect on the lowest excitation energies and general one-photon profile, but it can substantially redistribute 2PA intensity among the excited states. This sensitivity becomes more pronounced for the higher states, where the response also depends strongly on the selected QM-region size.

6. Summary and Conclusions

Across both rPSB models and full QM/MM Rhodopsin ensembles, the present work demonstrates that two-photon absorption remains among the most method-sensitive and configuration-sensitive spectral properties investigated. While excitation energies and one-photon properties remain comparatively robust across electronic-structure methods and configurational sampling, nonlinear response properties exhibit substantially stronger sensitivity to energetic compression, state interactions, and near-resonant enhancement effects. By combining coupled-cluster benchmarking, TD-DFT assessment, de/methylation studies, and configurationally sampled QM/MM ensembles, the current work establishes consistent relationships between molecular environment, excitation energies, dipole-based properties, and two-photon response.

Particularly in PSB3 and dm-PSB4, CC2 excitation energies typically deviated from CC3 by less than 1%, while CCSD systematically overestimated the same transitions by approximately 3 – 4%. For the lower-lying two-photon responses, both CC2 and CCSD generally remained close to CC3, although their relative agreement varied across states and systems. This overall behavior was partly retained upon methylation to PSB4, where both methods reproduced several of the low-lying δ^{2PA} contributions reasonably well despite the increasing compression of the higher excited-state manifold. Although part of this agreement may arise from favorable error cancellation within dipole-based quantities, the results still show that the approximate CC methods capture the main low-lying excitation-energy and two-photon trends across the smaller retinal analogues. However, as the conjugated double-bond system expanded, the higher excited-state manifold became progressively more compressed. This eventually exposed the limitations of CC2 in the higher excited-states, where state reordering, negative σ^{2PA} values, and unstable δ^{2PA} behavior appeared relative to the more physically consistent CC3 reference. Importantly, CCSD did not consistently resolve these issues either, as negative or unstable σ^{2PA} values remained in several problematic higher excited-states. In contrast, CC3 retained substantially more stable and internally consistent nonlinear response behavior throughout the investigated systems, particularly within the compressed $S_2 - S_4$ manifolds. Therefore, these results support CC2 as a practical baseline for excitation energies and qualitative low-lying spectral trends, while the quantitative accuracy of the two-photon properties remains dependent on

the particular state and system. At the same time, this also highlights the necessity of higher-order correlation treatment for reliable recovery of the compressed higher excited-states.

Within the TD-DFT investigations, the CAM-B3LYP and BHandHLYP functionals consistently provided the best overall qualitative agreement with CC3 for the low-lying one- and two-photon spectral trends. However, accurate recovery of one-photon properties did not translate into reliable nonlinear response behavior. Although CAM-B3LYP and BHandHLYP reproduced excitation energies, oscillator strengths, and transition dipoles comparatively well, both functionals systematically redistributed 2PA intensity away from the lowest excited-state toward compressed higher excited-states, with BHandHLYP exhibiting slightly stronger S_1 suppression together with somewhat improved higher state redistribution compared to CAM-B3LYP. Lower exact-exchange functionals produced substantially less reliable higher excited-state behavior, exhibiting stronger state compression, increased delocalization, and different placement of bright and dark responses among the computed roots. Notably, the comparatively stable δ^{2PA} trends across the TD-DFT models appear to benefit partly from compensating deviations among μ_0 , μ_1 , excitation energies, and μ_{01} , despite significant inaccuracies in the individual electronic properties. Furthermore, increasing the exact amount of exchange energy improves nonlinear response not only through improved dipole recovery, but also through more reliable separation and redistribution of intermediate states within quadratic-response itself.

Across both CC and TD-DFT calculations within the reduced size PSB models, increasing basis-set size mostly affected quantitative properties rather than changing the overall spectral trends. Diffuse functions became increasingly important for recovering higher excited-state response within compressed and more delocalized manifolds, while excitation energies and dominant low-energy transitions remained comparatively robust across moderate basis sets. Nevertheless, it appears that the choice of method itself plays a much larger role than basis-set size when it comes to higher excited-states, particularly regarding state ordering, mixing, and nonlinear response stability. In other words, simply increasing the basis-set size alone is generally not enough to resolve these higher state issues.

The structural modification studies further demonstrated that relatively small changes to the retinal polyene chain produce systematic but highly non-uniform changes across both one- and two-photon properties. Previous studies [79, 95] have already shown that methylation patterns, bond distortions, and torsional deformations can predictably shift excitation energies through modifications to conjugation and charge distribution, while Walczak *et al.* [92] further demonstrated that these trends

extend beyond excitation energies into oscillator strengths and permanent dipole moments. The present work demonstrates that these structure-property relationships extend further into nonlinear response behavior. Specifically, modifications involving C9, C10, and C13 methylation patterns systematically redistributed δ^{2PA} intensities despite comparatively modest energetic shifts and comparatively smaller changes in one-photon properties. Stronger nonlinear responses were generally associated with planar analogues containing the C13 methyl group, with the 13-methyl and 9,10,13-trimethyl structures producing the largest δ^{2PA} enhancements, whereas structures lacking the C13 methyl and/or containing torsional distortions consistently produced weaker response. The results consistently showed that favorable nonlinear response cannot be predicted from any singular electronic property alone. While larger $\Delta\mu$ generally correlated with enhanced δ^{2PA} , excitation energies, transition dipoles, and excited-state redistribution collectively govern the final response. Higher excited-states further exhibited substantially greater sensitivity to structural modification, displaying stronger redistribution, increased state mixing, and less predictable nonlinear behavior compared to the dominant lower excited-states. Furthermore, structural perturbations which reduce planarity, such as the twisted analogues, consistently produced reductions in $\Delta\mu$, transition dipoles, and δ^{2PA} , reinforcing that nonlinear response depends more strongly on how these excited-state properties interact rather than on any individual electronic property alone.

The QM/MM Rhodopsin ensembles investigated in Chapter 5 demonstrated that configurational averaging is essential for reliable interpretation of nonlinear response properties. Previous QM/MM investigations have already shown that accurate recovery of retinal excitation energies and spectral features requires configurational sampling rather than reliance on individual structures [97, 158, 218]. The present work demonstrates that this requirement extends even further into nonlinear response behavior, where individual configurations frequently exhibited strongly amplified higher state responses and substantial variability throughout the excited-state manifold. Nevertheless, ensemble averaging consistently recovered a structured multi-peak spectrum whose low-lying peak placement was qualitatively closer to the XMCQDPT2/PE results of Gholami et al. [97] than to the substantially weaker and more red-shifted CAM-B3LYP/EE spectrum reported by Palczewska et al. [16], particularly within the $S_1 - S_3$ region. Given the considerable differences in electronic-structure method, QM-region size, embedding treatment, and configurational sampling, this comparison is intended to place the present results within the range of previous predictions rather than provide a direct validation against either study. Interestingly, the lower excited-states remained comparatively robust throughout both M1 and M2, with excitation energies, one-photon properties, and bright-state ordering exhibiting only modest configurational sensitivity despite the highly heterogeneous protein environment. Expansion of the QM region from M1 to M2 altered the low-lying excited-state

structure only moderately, with the averaged S_1 excitation shifting by merely ~ 0.03 eV, indicating that the dominant low-energy excitation is already reasonably captured within the smaller M1 model. In contrast, QM-region expansion primarily influenced the higher excited-state manifold, leading to a more compressed $S_4 - S_6$ region together with stronger redistribution of excited-state contributions and substantially larger variability. Unlike the lower excited-states, where coefficients of variation generally remained below $\sim 50\%$, higher excited-states frequently approached or exceeded 100% , indicating variability comparable to or even larger than their averaged intensities. This behavior was particularly evident for recurring problematic configurations such as 22PS and 23PS, together with repeated contributions from 16PS and 17PS, which consistently produced amplified higher state responses across both ensembles. These amplified responses repeatedly emerged within compressed higher excited-state regions where energetic separations frequently approached only $\sim 0.10 - 0.20$ eV, while several higher excited-states also occurred at approximately twice the isolated S_1 excitation energy, producing conditions favorable for near-resonant enhancement of higher state response.

Extension towards even larger QM regions further reinforced this conclusion. Expansion from M1 toward M4 produced only comparatively modest changes within the low-lying excited-state manifold, with dominant excitation energies, one-photon properties, and bright-state ordering remaining largely preserved despite progressively larger QM descriptions. Similarly, comparison between PE, EE, and NE calculations showed that the S_1 excitation energy remains comparatively stable across embedding treatments, whereas S_2 and S_3 are more sensitive, particularly under NE. The associated 2PA intensities vary more strongly than the excitation energies and are redistributed among the excited states depending on both embedding treatment and QM-region size. Therefore, QM-region expansion and embedding have a comparatively modest effect on the general low-energy one-photon profile, but a stronger and more state-dependent effect on the nonlinear response.

Overall, while excitation energies remained comparatively robust throughout Chapter 5, increasing sensitivity emerged when moving toward one-photon properties, dipole descriptors, and ultimately nonlinear response itself, with 2PA remaining the most configuration-dependent property investigated. Furthermore, permanent dipoles and $\Delta\mu$ alone remained insufficient for describing the observed response behavior, particularly within Rhodopsin where these quantities became strongly dominated by the large ground-state polarization contribution. Instead, the present results suggest that higher excited-state nonlinear response depends more strongly on energetic spacing and interactions between neighboring excited-states than on isolated electronic descriptors alone.

Taken together, these findings suggest a practical strategy for future computational studies of retinal and Rhodopsin multiphoton absorption. CC2 can provide an efficient starting point for screening excitation energies and broad spectral trends in smaller retinal models, while quantitative 2PA values and closely spaced higher excited states require validation using higher-level correlation treatments. The two-state model can similarly support the initial ranking of retinal analogues, although it should remain a trend descriptor rather than a replacement for full quadratic-response calculations. The de/methylation results further identify substitution pattern and polyene-chain torsion as structural factors through which transition dipoles, changes in permanent dipole moment, and 2PA response may be modified. For Rhodopsin, the results show that configurational sampling and an appropriate treatment of the protein environment become increasingly important for nonlinear properties, particularly within the more compressed and variable higher excited-state manifold. Together, these observations provide methodological and molecular guidance for future screening of retinal analogues and Rhodopsin variants for multiphoton activation and imaging applications.

The same framework provides a starting point for extending the analysis toward three-photon absorption (3PA). Although the corresponding 3PA datasets have already been generated, their detailed analysis remains ongoing. This data will allow the configurational sensitivity, QM-region dependence, and embedding effects identified for 2PA to be examined in the context of 3PA. They will also help determine whether the manifold-compression and near-resonance effects observed throughout the present work become more important at the next order of nonlinear response.

7. Supporting Information

CC3							BHandHLYP						
states	ΔE	δ^{2PA}	σ^{2PA}	f	μ	$\Delta\mu$	states	ΔE	δ^{2PA}	σ^{2PA}	f	μ	$\Delta\mu$
ground	-	-	-	-	4.333	-	ground	-	-	-	-	3.298	-
1	4.06	2098	12.69	0.614	-	-	1	4.29	266	1.79	0.675	2.075	1.332
-	-	-	-	-	-	-	2	6.47	2	0.03	0.001	2.813	0.703
-	-	-	-	-	-	-	3	6.49	1390	21.45	0.165	0.707	2.714
-	-	-	-	-	-	-	4	6.68	201	3.29	0.125	0.390	2.911
-	-	-	-	-	-	-	5	7.02	4	0.08	0.001	10.484	7.291
-	-	-	-	-	-	-	6	7.32	62	1.21	0.000	5.767	3.709
-	-	-	-	-	-	-	7	7.37	1	0.01	0.000	2.143	1.224
CAM-B3LYP							B3LYP						
states	ΔE	δ^{2PA}	σ^{2PA}	f	μ	$\Delta\mu$	states	ΔE	δ^{2PA}	σ^{2PA}	f	μ	$\Delta\mu$
ground	-	-	-	-	3.463	-	ground	-	-	-	-	3.398	-
1	4.18	323	2.07	0.637	2.150	1.417	1	4.09	195	1.20	0.564	1.778	1.705
2	6.03	1	0.01	0.000	2.863	0.793	2	5.67	1	0.00	0.000	3.194	0.774
3	6.39	968	14.48	0.147	0.850	2.715	3	5.78	186	2.30	0.204	0.565	3.299
4	6.61	174	2.79	0.142	0.128	3.347	4	6.49	10	0.20	0.001	11.756	8.388
5	6.98	0	0.00	0.000	2.305	1.205	5	6.51	109	1.70	0.113	0.250	3.170
6	7.03	5	0.09	0.002	10.050	6.762	6	6.66	0	0.00	0.000	2.595	0.965
7	7.28	9	0.18	0.000	2.503	1.069	7	6.89	64	1.10	0.001	5.956	3.092
B1LYP							BLYP						
states	ΔE	δ^{2PA}	σ^{2PA}	f	μ	$\Delta\mu$	states	ΔE	δ^{2PA}	σ^{2PA}	f	μ	$\Delta\mu$
ground	-	-	-	-	3.383	-	ground	-	-	-	-	3.439	-
1	4.12	210	1.30	0.584	1.841	1.630	1	3.94	132	0.80	0.468	1.504	1.997
2	5.83	1	0.00	0.000	3.107	0.753	2	5.09	1	0.00	0.000	3.590	0.937
3	5.90	274	3.50	0.198	0.563	3.237	3	5.27	41	0.40	0.244	0.584	3.526
4	6.53	122	1.90	0.114	0.316	3.088	4	5.99	15	0.20	0.001	12.361	8.928
5	6.53	9	0.10	0.001	11.626	8.286	5	6.09	0	0.00	0.000	2.932	0.852
6	6.80	0	0.00	0.000	2.504	1.010	6	6.31	0	0.00	0.001	1.986	1.583
7	6.91	62	1.10	0.001	5.964	3.194	7	6.34	94	1.40	0.101	0.189	3.628
-	-	-	-	-	-	-	8	6.48	45	0.70	0.001	6.319	3.114
-	-	-	-	-	-	-	9	6.96	9734	172.70	0.008	0.698	3.039

Table S1. Ground-state and available excited-state electronic, one-photon absorption, and two-photon absorption properties of PSB3 calculated using CC3, CAM-B3LYP, BHandHLYP, B1LYP, B3LYP, and BLYP with the aug-cc-pVDZ basis set. Vertical excitation energies, ΔE , are reported in eV; two-photon transition strengths, δ^{2PA} , in a.u.; and two-photon absorption cross-sections, σ^{2PA} , in GM. Oscillator strengths, f , are dimensionless and reported in the length gauge. Permanent dipole-moment magnitudes, μ , and changes in permanent dipole moment, $\Delta\mu$, are reported in Debye. Ground-state rows report μ_0 , while excited-state rows report μ_n . Higher excited-state CC3 results were unrecoverable; therefore, only the available CC3 S_1 result is reported, and unavailable values are indicated by dashes.

General notes for Tables S2a–S2d. Vertical excitation energies, ΔE , are reported in eV; two-photon transition strengths, δ^{2PA} , in a.u.; two-photon absorption cross-sections, σ^{2PA} , in GM; and transition dipole moments, μ_0 , in a.u. Oscillator strengths, f , are dimensionless. Permanent dipole-moment magnitudes, μ , and their differences, $\Delta\mu$ and $\Delta\mu(\text{ST})$, are reported in Debye. State 0 reports the ground-state permanent dipole-moment magnitude, μ_0 , while states S_1 – S_5 report μ_n . $\Delta\mu = \|\mu_n - \mu_0\|$ is the magnitude of the Cartesian vector difference between the ground- and excited-state dipole moments, whereas $\Delta\mu(\text{ST}) = \|\mu_n\| - \|\mu_0\|$ is the absolute difference between their total magnitudes. The two-photon quantities correspond to linearly polarized, parallel, equal-energy photons. For MN15 and M11, the one-photon properties were obtained using the gauge-invariant current-dependent formalism, whereas the two-photon properties were obtained using the available gauge-variant formulation. For BHandHLYP and CAM-B3LYP, the one-photon transition properties are reported in the length representation and the two-photon properties were obtained using quadratic response. Values affected by strong near-resonant enhancement are retained for diagnostic completeness and should not be interpreted quantitatively.

MN15																	
Wild-Type – MN15 (comb)									9-methyl – MN15 (comb)								
State	ΔE	δ ^{2PA}	σ ^{2PA}	f	μ ₀₁ (au)	μ	Δμ	Δμ(ST)	State	ΔE	δ ^{2PA}	σ ^{2PA}	f	μ ₀₁ (au)	μ	Δμ	Δμ(ST)
0						6.497			0						6.690		
1	2.92	1655	5.1	1.348	11.035	4.771	1.739	1.726	1	2.86	1135	3.4	1.210	10.560	5.289	1.492	1.401
2	4.38	1361	9.6	0.089	2.313	4.544	2.057	1.953	2	4.43	5881	42.2	0.132	2.801	5.854	0.893	0.836
3	5.03	202635	1867.4	0.149	2.794	6.069	0.463	0.428	3	5.00	156285	1425.1	0.248	3.619	6.373	0.620	0.316
4	5.20	1895	18.7	0.006	0.572	4.494	2.010	2.003	4	5.34	78	0.8	0.037	1.355	2.044	4.752	4.646
5	5.28	133	1.3	0.000	0.018	2.680	3.824	3.817	5	5.38	151555	1579.1	0.000	0.066	3.157	3.615	3.532
13-methyl – MN15 (comb)									9-10-di-methyl(CS) – MN15 (comb)								
State	ΔE	δ ^{2PA}	σ ^{2PA}	f	μ ₀₁ (au)	μ	Δμ	Δμ(ST)	State	ΔE	δ ^{2PA}	σ ^{2PA}	f	μ ₀₁ (au)	μ	Δμ	Δμ(ST)
0						5.771			0						6.993		
1	2.99	2504	8.1	1.388	11.068	3.651	2.133	2.120	1	2.92	920	2.8	1.377	11.160	5.728	1.327	1.265
2	4.45	1390	10.1	0.091	2.320	4.463	1.314	1.309	2	4.36	1007	7.0	0.090	2.329	5.379	1.693	1.614
3	5.11	95557	907.9	0.120	2.489	6.145	0.377	0.373	3	4.99	0	0.0	0.089	2.174	2.575	4.636	4.418
4	5.45	41861	455.0	0.052	1.582	4.777	1.097	0.994	4	5.01	169519	1538.1	0.000	0.028	5.421	1.644	1.572
5	5.55	25	0.3	0.000	0.020	3.654	2.135	2.117	5	5.12	11809	112.9	0.022	1.069	4.884	2.126	2.109
9-10-di-methyl(twist) – MN15 (comb)									9-10-13-tri-methyl(CS) – MN15 (comb)								
State	ΔE	δ ^{2PA}	σ ^{2PA}	f	μ ₀₁ (au)	μ	Δμ	Δμ(ST)	State	ΔE	δ ^{2PA}	σ ^{2PA}	f	μ ₀₁ (au)	μ	Δμ	Δμ(ST)
0						6.663			0						7.259		
1	2.86	573	1.7	1.274	10.836	5.639	1.088	1.024	1	2.79	2986	8.4	1.274	10.974	5.034	2.224	2.224
2	4.34	1691	11.7	0.082	2.225	5.984	0.707	0.679	2	4.24	811	5.3	0.143	2.976	5.810	1.459	1.449
3	4.93	136527	1208.6	0.122	2.554	6.629	0.107	0.034	3	4.87	2	0.0	0.116	2.505	2.768	4.493	4.491
4	5.05	15560	143.2	0.047	1.559	2.421	4.562	4.242	4	4.91	250665	2167.7	0.000	0.048	4.705	2.553	2.553
5	5.09	5194	48.4	0.002	0.335	3.410	3.484	3.253	5	4.93	79998	708.8	0.009	0.696	5.042	2.218	2.216
9-10-13-tri-methyl(twist) – MN15 (comb)																	
State		ΔE	δ ^{2PA}	σ ^{2PA}	f	μ ₀₁ (au)	μ	Δμ	Δμ(ST)								
0							6.483										
1		2.68	1423	3.8	1.031	10.065	4.899	1.642	1.584								
2		4.07	717	4.3	0.101	2.557	5.991	0.621	0.491								
3		4.81	519936	4377.5	0.035	1.388	6.311	0.291	0.172								
4		4.85	376045	3231.0	0.108	2.417	2.382	4.576	4.100								
5		5.06	18277	166.3	0.005	0.500	1.892	4.844	4.591								

Table S2a. Ground-state and S_1 - S_5 excited-state electronic, one-photon absorption, and two-photon absorption properties of all seven RPSB5 models calculated using MN15/aug-cc-pVDZ.

M11																			
Wild-Type – M11 (comb)									9-methyl – M11 (comb)										
State	ΔE	δ ^{2PA}	σ ^{2PA}	f	μ ₀₁ (au)	μ	Δμ	Δμ(ST)	State	ΔE	δ ^{2PA}	σ ^{2PA}	f	μ ₀₁ (au)	μ	Δμ	Δμ(ST)		
0						6.942			0						7.081				
1	2.91	4695	14.5	1.400	11.254	4.053	2.899	2.889	1	2.85	3480	10.3	1.270	10.838	4.641	2.576	2.440		
2	4.81	33599	282.4	0.063	1.864	4.556	2.582	2.386	2	4.90	13275	115.9	0.211	3.369	5.091	2.220	1.991		
3	5.46	2136	23.0	0.000	0.105	3.076	9.058	3.866	3	5.30	2359305	24185.0	0.158	2.799	8.035	0.987	0.953		
4	5.48	184	2.0	0.159	2.768	3.070	3.898	3.872	4	5.39	9	0.1	0.000	0.072	6.421	9.366	0.661		
5	5.51	4487881	49112.4	0.000	0.081	8.143	1.231	1.201	5	5.60	1332	15.1	0.000	0.028	2.614	4.473	4.468		
13-methyl – M11 (comb)									9-10-di-methyl(CS) – M11 (comb)										
State	ΔE	δ ^{2PA}	σ ^{2PA}	f	μ ₀₁ (au)	μ	Δμ	Δμ(ST)	State	ΔE	δ ^{2PA}	σ ^{2PA}	f	μ ₀₁ (au)	μ	Δμ	Δμ(ST)		
0						6.336			0						7.357				
1	3.01	6351	20.8	1.446	11.265	2.939	3.407	3.397	1	2.89	3041	9.2	1.423	11.400	5.070	2.344	2.287		
2	4.87	35811	308.5	0.032	1.319	3.931	2.479	2.406	2	4.79	26577	221.7	0.086	2.182	4.950	2.548	2.406		
3	5.51	403817	4466.6	0.195	3.053				3	5.27	4	0.0	0.000	0.037	2.775	4.680	4.582		
4	5.56	779	8.7	0.000	0.047	4.285	9.866	2.051	4	5.42	5589266	59948.1	0.080	1.974	8.380	1.058	1.023		
5	5.73	53	0.6	0.000	0.042	3.961	2.428	2.375	5	5.47	557	6.0	0.001	0.265	2.744	9.781	4.613		
9-10-di-methyl(twist) – M11 (comb)									9-10-13-tri-methyl(CS) – M11 (comb)										
State	ΔE	δ ^{2PA}	σ ^{2PA}	f	μ ₀₁ (au)	μ	Δμ	Δμ(ST)	State	ΔE	δ ^{2PA}	σ ^{2PA}	f	μ ₀₁ (au)	μ	Δμ	Δμ(ST)		
0						6.991			0						7.804				
1	2.83	2146	6.2	1.317	11.078	5.045	2.030	1.946	1	2.80	7934	22.6	1.359	11.304	4.231	3.583	3.574		
2	4.77	18931	156.3	0.127	2.650	5.273	1.822	1.718	2	4.63	37495	292.3	0.037	1.445	5.946	1.899	1.859		
3	5.28	1636578	16557.8	0.044	1.485	5.871	1.123	1.120	3	5.16	20	0.2	0.000	0.070	3.412	4.400	4.392		
4	5.35	2677598	27701.3	0.038	1.364	4.946	2.091	2.046	4	5.27	3674802	37167.8	0.203	3.191	8.371	0.602	0.566		
5	5.41	78086	824.9	0.002	0.288	2.621	8.728	4.370	5	5.35	4799	49.6	0.002	0.319	3.720	8.608	4.085		
									9-10-13-tri-methyl(twist) – M11 (comb)										
State	ΔE	δ ^{2PA}	σ ^{2PA}	f	μ ₀₁ (au)	μ	Δμ	Δμ(ST)											
0						7.008													
1	2.69	4423	11.6	1.079	10.288	4.192	2.860	2.816											
2	4.41	31603	223.5	0.091	2.337	5.876	1.188	1.132											
3	5.21	78235134	771463.1	0.118	2.448	7.648	0.720	0.640											
4	5.33	133334517	1363819.6	0.028	1.181	2.620	4.466	4.388											
5	5.40	8397999890	88446150.1	0.002	0.312	3.336	9.463	3.672											

Table S2b. Ground-state and S_1 - S_5 excited-state electronic, one-photon absorption, and two-photon absorption properties of all seven RPSB5 models calculated using M11/aug-cc-pVDZ.

BHandHLYP															
Wild-Type – BhandHLYP (aug-cc-pVDZ)								9-methyl – BhandHLYP (aug-cc-pVDZ)							
State	ΔE	δ^{2PA}	σ^{2PA}	f	$\mu_{01}(\text{au})$	μ	Δμ	State	ΔE	δ^{2PA}	σ^{2PA}	f	$\mu_{01}(\text{au})$	μ	Δμ
0						6.334		0						6.500	
1	3.02	1660	5.5	1.455	11.263	4.495	1.849	1	2.96	1143	3.7	1.308	10.785	4.908	1.612
2	4.64	4433	34.9	0.084	2.178	4.386	2.067	2	4.68	1828	14.7	0.138	2.788	5.594	0.909
3	5.22	119000	1187.9	0.144	2.698	6.658	0.337	3	5.15	78000	757.9	0.258	3.634	6.359	0.305
4	5.53	55100	617.3	0.039	1.359	3.813	2.598	4	5.7	1020000	12140.9	0.046	1.453	2.883	3.741
5	5.73	1403	16.8	0.000	0.087	6.447	12.767	5	5.73	105	1.3	0.000	0.068	6.449	11.403
13-methyl – BhandHLYP (aug-cc-pVDZ)								9-10-di-methyl(CS) – BhandHLYP (aug-cc-pVDZ)							
State	ΔE	δ^{2PA}	σ^{2PA}	f	$\mu_{01}(\text{au})$	μ	Δμ	State	ΔE	δ^{2PA}	σ^{2PA}	f	$\mu_{01}(\text{au})$	μ	Δμ
0						5.606		0						6.812	
1	3.10	2597	9.2	1.503	11.313	3.331	2.284	1	3.01	939	3.1	1.481	11.384	5.383	1.430
2	4.71	6037	49.1	0.075	2.047	3.973	1.644	2	4.61	2896	22.6	0.089	2.260	5.098	1.718
3	5.26	46200	468.3	0.132	2.574	5.162	0.446	3	5.19	126489	1243.4	0.095	2.195	6.592	0.220
4	5.78	91600	1121.1	0.072	1.813	4.305	1.302	4	5.47	47511	520.7	0.027	1.141	2.702	4.117
5	5.81	449	5.6	0.000	0.041	8.115	13.712	5	5.76	20	0.2	0.001	0.212	8.117	14.705
9-10-di-methyl(twist) – BhandHLYP (aug-cc-pVDZ)								9-10-13-tri-methyl(CS) – BhandHLYP (aug-cc-pVDZ)							
State	ΔE	δ^{2PA}	σ^{2PA}	f	$\mu_{01}(\text{au})$	μ	Δμ	State	ΔE	δ^{2PA}	σ^{2PA}	f	$\mu_{01}(\text{au})$	μ	Δμ
0						6.460		0						7.115	
1	2.96	547	1.8	1.374	11.063	5.315	1.147	1	2.91	3239	10.1	1.404	11.283	4.680	2.435
2	4.60	1281	9.9	0.087	2.234	5.637	0.824	2	4.50	3848	28.6	0.116	2.612	5.513	1.610
3	5.10	87083	830.0	0.123	2.523	5.884	0.581	3	5.06	152555	1435.1	0.076	1.996	6.590	0.569
4	5.44	113167	1225.1	0.057	1.660	2.290	4.260	4	5.31	13008	134.3	0.074	1.920	1.935	5.216
5	5.72	17895	214.6	0.002	0.275	7.362	13.500	5	5.59	118	1.4	0.000	0.070	2.752	4.365
9-10-13-tri-methyl(twist) – BhandHLYP (aug-cc-pVDZ)															
State	ΔE	δ^{2PA}	σ^{2PA}	f	$\mu_{01}(\text{au})$	μ	Δμ	State	ΔE	δ^{2PA}	σ^{2PA}	f	$\mu_{01}(\text{au})$	μ	Δμ
0						6.360		0						6.360	
1	2.8	1595	4.6	1.129	10.315	4.524	1.848	1	2.8	1595	4.6	1.129	10.315	4.524	1.848
2	4.31	3672	25.0	0.101	2.485	5.692	0.800	2	4.31	3672	25.0	0.101	2.485	5.692	0.800
3	5.00	405638	3718.5	0.092	2.206	5.495	0.902	3	5.00	405638	3718.5	0.092	2.206	5.495	0.902
4	5.24	131136	1317.8	0.089	2.112	1.116	5.388	4	5.24	131136	1317.8	0.089	2.112	1.116	5.388
5	5.66	1062848	12440.6	0.001	0.237	5.104	11.441	5	5.66	1062848	12440.6	0.001	0.237	5.104	11.441

Table S2c. Ground-state and S₁ - S₅ excited-state electronic, one-photon absorption, and two-photon absorption properties of all seven RPSB5 models calculated using BHandHLYP/aug-cc-pVDZ.

CAM-B3LYP															
Wild-Type – CAMB3LYP (aug-cc-pVDZ)								9-methyl – CAMB3LYP (aug-cc-pVDZ)							
State	ΔE	δ^{2PA}	σ^{2PA}	f	$\mu_{01}(\text{au})$	μ	Δμ	State	ΔE	δ^{2PA}	σ^{2PA}	f	$\mu_{01}(\text{au})$	μ	Δμ
0						6.533		0						6.704	
1	2.95	2187	7.0	1.396	11.179	4.537	2.007	1	2.88	1541	4.7	1.25	10.710	4.975	1.749
2	4.61	4213	32.8	0.078	2.112	4.300	2.249	2	4.66	6055	48.2	0.14	2.851	5.613	1.095
3	5.17	279000	2732.0	0.142	2.694	5.908	0.625	3	5.08	32900	1872.0	0.24	3.514	6.557	0.322
4	5.41	60900	653.0	0.039	1.373	3.614	2.921	4	5.57	336000	22732.2	0.05	1.516	2.793	4.121
5	5.56	465	5.3	0.000	0.013	2.832	3.719	5	5.66	1160	13.6	0.00	0.059	2.294	4.442
13-methyl – CAMB3LYP (aug-cc-pVDZ)								9-10-di-methyl(CS) – CAMB3LYP (aug-cc-pVDZ)							
State	ΔE	δ^{2PA}	σ^{2PA}	f	$\mu_{01}(\text{au})$	μ	Δμ	State	ΔE	δ^{2PA}	σ^{2PA}	f	$\mu_{01}(\text{au})$	μ	Δμ
0						5.857		0						7.000	
1	3.02	3267	10.9	1.441	11.212	3.417	2.449	1	2.93	1266	4.0	1.422	11.307	5.451	1.549
2	4.68	5863	47.0	0.063	1.891	4.142	1.734	2	4.58	2429	18.7	0.088	2.255	5.053	1.950
3	5.21	100000	994.4	0.139	2.654	5.184	0.673	3	5.13	291836	2815.3	0.091	2.158	6.557	0.446
4	5.67	122000	1436.9	0.074	1.855	4.501	1.358	4	5.29	1	0.0	0.000	0.029	2.518	4.483
5	5.81	226	2.8	0.000	0.032	3.706	2.170	5	5.33	49332	513.1	0.027	1.147	5.451	1.549
9-10-di-methyl(twist) – CAMB3LYP (aug-cc-pVDZ)								9-10-13-tri-methyl(CS) – CAMB3LYP (aug-cc-pVDZ)							
State	ΔE	δ^{2PA}	σ^{2PA}	f	$\mu_{01}(\text{au})$	μ	Δμ	State	ΔE	δ^{2PA}	σ^{2PA}	f	$\mu_{01}(\text{au})$	μ	Δμ
0						6.652		0						7.332	
1	2.88	810	2.5	1.315	10.979	5.370	1.285	1	2.83	3987	11.7	1.337	11.173	4.753	2.580
2	4.57	1608	12.3	0.095	2.340	5.617	1.035	2	4.46	3337	24.3	0.107	2.511	5.754	1.600
3	5.03	203000	1881.6	0.106	2.363	5.811	0.848	3	5.00	383869	3517.0	0.097	2.260	6.668	0.690
4	5.30	156000	1605.4	0.063	1.765	2.133	4.639	4	5.18	6453	63.4	0.069	1.877	1.737	5.617
5	5.37	527	5.6	0.000	0.120	2.495	4.204	5	5.19	9	0.1	0.000	0.053	2.994	4.342
9-10-13-tri-methyl(twist) – CAMB3LYP (aug-cc-pVDZ)															
State	ΔE	δ^{2PA}	σ^{2PA}	f	$\mu_{01}(\text{au})$	μ	Δμ	State	ΔE	δ^{2PA}	σ^{2PA}	f	$\mu_{01}(\text{au})$	μ	Δμ
0						6.567		0						6.567	
1	2.72	2037	5.5	1.078	10.222	4.599	1.977	1	2.72	2037	5.5	1.078	10.222	4.599	1.977
2	4.26	3024	20.1	0.101	2.498	5.866	0.806	2	4.26	3024	20.1	0.101	2.498	5.866	0.806
3	4.94	1044735	9298.0	0.080	2.065	5.597	0.996	3	4.94	1044735	9298.0	0.080	2.065	5.597	0.996
4	5.13	349679	3374.5	0.092	2.171	0.959	5.804	4	5.13	349679	3374.5	0.092	2.171	0.959	5.804
5	5.35	483193	5064.7	0.005	0.493	1.893	4.697	5	5.35	483193	5064.7	0.005	0.493	1.893	4.697

Table S2d. Ground-state and S₁ - S₅ excited-state electronic, one-photon absorption, and two-photon absorption properties of all seven RPSB5 models calculated using CAM-B3LYP/aug-cc-pVDZ.

Average M1/PE 10 frames						
State	ΔE	λ	δ^{2PA}	σ^{2PA}	f	Removed States
1	2.64	943	12003	31.0	1.048	
2	4.01	619	83090	486.1	0.175	
3	4.41	563	184610	1280.7	0.147	
4	4.81	516	173971	1385.6	0.094	20/22/23PS
5	5.14	484	53049	561.2	0.000	17/18/21PS*
6	5.34	465	235400	2408.1	0.024	14/16/22PS
Average M1/EE 10 frames						
State	ΔE	λ	δ^{2PA}	σ^{2PA}	f	Removed States
1	2.73	909	18736	52.2	1.559	
2	4.03	616	87354	523.1	0.201	
3	4.27	582	14864	105.2	0.110	
4	4.49	554	56824	395.6	0.067	
5	4.69	530	18032	144.8	0.049	
6	4.86	510	70986	631.5	0.033	
Average M1/NE 10 frames						
State	ΔE	λ	δ^{2PA}	σ^{2PA}	f	Removed States
1	2.47	1004	36127	81.6	1.329	
2	3.48	713	67312	302.4	0.345	
3	4.10	605	38945	247.6	0.084	
4	4.35	571	928785	272.0	0.055	22PS
5	4.53	549	991495	7976.1	0.048	23PS
6	4.58	542	1061238	7643.7	0.110	16/20/22/23PS
Average M2/PE 10 frames						
State	ΔE	λ	δ^{2PA}	σ^{2PA}	f	Removed States
E.S.	E.E.	$\lambda(\text{nm})$	dTPA	sTPA	f	
1	2.68	929	11945	31.9	1.030	
2	4.04	615	79020	469.9	0.172	
3	4.38	568	67948	459.6	0.124	
4	4.66	532	68638	572.3	0.042	
5	4.80	516	56192	481.7	0.058	22/23PS
6	4.96	500	26234	238.8	0.027	16/17/22/23PS
Average M2/EE 10 frames						
State	ΔE	λ	δ^{2PA}	σ^{2PA}	f	Removed States
1	2.77	897	14592	41.8	1.428	
2	4.05	613	74121	453.6	0.109	
3	4.18	594	23164	149.8	0.094	
4	4.31	575	22232	141.6	0.036	
5	4.42	562	5573	38.1	0.045	
6	4.55	546	8892	65.9	0.057	
7	4.66	532	26530	219.0	0.042	
8	4.74	524	22863	192.8	0.043	
9	4.79	518	50943	439.2	0.024	
Average M2/NE 10 frames						
State	ΔE	λ	δ^{2PA}	σ^{2PA}	f	Removed States
1	2.73	911	18402	50.8	1.382	
2	3.46	721	3447	17.1	0.011	
3	3.71	672	20392	115.9	0.052	
4	3.86	645	16529	99.2	0.028	
5	4.04	614	38863	233.0	0.086	
6	4.24	587	32366	200.0	0.091	
7	4.42	561	32347	241.6	0.158	
8	4.63	536	9703	76.4	0.008	
9	4.76	521	122828	1021.9	0.052	

Table S3. Ensemble averaged excited-state properties of the M1 model (upper block) and M2 model (lower block) under polarizable embedding, electrostatic embedding, and no-embedding treatments, calculated for the same ten QM/MM MD snapshots at the BHandHLYP/aug-cc-pVDZ level of theory. Excitation energies are reported in eV, wavelengths in nm, two-photon transition strengths, δ^{2PA} , in a.u., two-photon absorption cross-sections, σ^{2PA} , in GM, and Oscillator strengths, f , are dimensionless.

State	14PS	15PS	16PS	17PS	18PS	19PS	20PS	21PS	22PS	23PS
1	2.70 (918)	2.94 (843)	2.73 (908)	2.63 (943)	2.93 (846)	2.66 (932)	2.55 (972)	2.57 (965)	2.55 (972)	2.51 (988)
2	3.83 (647)	4.10 (605)	4.16 (596)	3.98 (623)	4.15 (597)	4.19 (592)	4.02 (617)	4.03 (615)	3.87 (641)	4.05 (612)
3	4.56 (544)	4.25 (583)	4.36 (569)	4.44 (558)	4.74 (523)	4.37 (567)	4.21 (589)	4.53 (547)	4.20 (590)	4.08 (608)
4	4.62 (537)	4.66 (532)	4.75 (522)	4.60 (539)	4.79 (518)	4.82 (514)	4.36 (569)	4.66 (532)	4.79 (518)	4.56 (544)
5	4.80 (517)	4.75 (522)	4.98 (498)	4.75 (522)	4.89 (507)	4.84 (512)	4.61 (538)	4.81 (516)	4.90 (506)	4.92 (504)
6	5.05 (491)	4.85 (511)	5.30 (468)	5.18 (479)	5.05 (491)	5.18 (479)	4.67 (531)	4.97 (499)	5.00 (496)	5.14 (482)

State	Mean	Stdev
1	2.68 (929)	0.15
2	4.04 (615)	0.12
3	4.38 (568)	0.20
4	4.66 (532)	0.14
5	4.82 (514)	0.11
6	5.04 (493)	0.18

Table S4. Snapshot resolved vertical excitation energies for the $S_1 - S_6$ states of the ten frames in the M2 ensemble under polarizable embedding, calculated at the BHandHLYP/aug-cc-pVDZ level of theory. Each entry reports the excitation energy, ΔE , in eV, followed in parentheses by the corresponding degenerate two-photon excitation wavelength in nm. The lower block reports the ensemble mean excitation energies, corresponding mean wavelengths, and standard deviations of ΔE .

State	14PS	15PS	16PS	17PS	18PS	19PS	20PS	21PS	22PS	23PS
1	0.985/9.803	0.968/9.313	0.994/9.793	1.082/10.413	0.936/9.182	0.969/9.809	1.175/11.034	1.077/10.508	1.083/10.590	1.034/10.413
2	0.151/3.221	0.152/3.124	0.187/3.443	0.104/2.629	0.148/3.069	0.158/3.152	0.223/3.825	0.124/2.853	0.209/3.778	0.263/4.136
3	0.093/2.324	0.000/0.121	0.221/3.657	0.025/1.212	0.102/2.383	0.245/3.843	0.199/3.528	0.093/2.329	0.145/3.017	0.114/2.708
4	0.129/2.713	0.048/1.648	0.004/0.475	0.084/2.198	0.047/1.606	0.025/1.176	0.002/0.363	0.071/1.999	0.009/0.684	0.004/0.490
5	0.045/1.569	0.083/2.145	0.022/1.073	0.119/2.567	0.140/2.747	0.016/0.922	0.000/0.074	0.042/1.516	0.009/0.704	0.047/1.595
6	0.012/0.789	0.143/2.790	0.002/0.314	0.004/0.474	0.006/0.538	0.000/0.111	0.002/0.317	0.000/0.050	0.007/0.587	0.000/0.136

State	Mean	Stdev
1	1.030/10.086	0.073/0.596
2	0.172/3.323	0.048/0.469
3	0.124/2.512	0.080/1.150
4	0.042/1.335	0.042/0.823
5	0.052/1.491	0.047/0.837
6	0.018/0.611	0.044/0.801

Table S5. Snapshot resolved length-gauge oscillator strengths, f , and transition dipole moments, μ_{0n} , for the $S_1 - S_6$ states of the ten frames in the M2 ensemble under polarizable embedding, calculated at the BHandHLYP/aug-cc-pVDZ level of theory. Each entry is reported as f/μ_{0n} , where f is dimensionless and μ_{0n} is given in Debye. The lower block reports the ensemble means and standard deviations of both properties.

State	14PS		15PS		16PS		17PS		18PS	
	δ^{2PA}	σ^{2PA}	δ^{2PA}	σ^{2PA}	δ^{2PA}	σ^{2PA}	δ^{2PA}	σ^{2PA}	δ^{2PA}	σ^{2PA}
1	18100	48.3	15800	50.0	12600	34.4	4280	10.9	12800	40.3
2	70300	377.8	40600	250.0	88800	563.0	69900	405.6	41600	262.5
3	30900	235.4	177	1.2	116000	807.9	11600	83.8	34000	279.9
4	41300	323.0	10200	81.2	19800	163.7	6080	47.1	28500	239.6
5	15600	131.7	57600	476.1	78900	716.9	16900	139.7	18200	159.4
6	63500	593.3	28500	245.6	469000	4826.4	787000	7736.4	641	6.0

State	19PS		20PS		21PS		22PS		23PS	
	δ^{2PA}	σ^{2PA}	δ^{2PA}	σ^{2PA}	δ^{2PA}	σ^{2PA}	δ^{2PA}	σ^{2PA}	δ^{2PA}	σ^{2PA}
1	11900	30.9	11500	27.4	7270	17.6	9600	22.9	15600	36.0
2	89200	573.7	128000	757.8	120000	714.0	111000	609.0	30800	185.1
3	51900	363.1	75900	492.8	77800	584.9	88200	570.0	193000	1177.0
4	71000	604.3	1500	10.5	15400	122.5	481000	4043.1	11600	88.4
5	62200	533.8	133	1.0	200000	1695.2	2970000	26124.6	78900000	699695.2
6	7000	68.8	3360	26.9	54400	492.3	33600000	307737.9	1670000	16163.8

State	ΔE	δ^{2PA}	λ		σ^{2PA}		Removed States
	Mean(e.V)	Mean(a.u.)	Stdev	Mean(nm)	Mean(GM)	Stdev	
1	2.677	11945	4131	929	31.86	12.66	
2	4.038	79020	34418	615	469.86	201.76	
3	4.377	67947.7	56831	568	459.59	350.91	
4	4.662	68638	146331	532	572.32	1231.84	
5	4.804	56191.625	64378	516	481.73	548.37	22/23PS
6	4.960	26233.5	27333	500	238.80	252.18	16/17/22/23PS

Table S6. Snapshot resolved two-photon absorption properties for the S₁ - S₆ states of the ten frames of the M2 ensemble under polarizable embedding, calculated at the BHandHLYP/aug-cc-pVDZ level of theory. Two-photon transition strengths, δ^{2PA} , are reported in a.u., and two-photon absorption cross-sections, σ^{2PA} , in GM. The lower block reports the mean excitation energies in eV, the corresponding mean degenerate two-photon excitation wavelengths in nm, and the ensemble means and standard deviations of δ^{2PA} and σ^{2PA} . All frame-resolved values are retained in the upper blocks, while frames listed in the “Removed States” column were excluded from the corresponding means and standard deviations because of strong near-resonant enhancement.

18PS-frame Polarizable Embedding BhandHLYP/aug-cc-pVDZ with D3-BJ dampening												
M1							M2					
State.	ΔE	λ (nm)	δ^{2PA}	σ^{2PA}	f	μ_{01} (D)	State.	ΔE	λ (nm)	δ^{2PA}	σ^{2PA}	μ_{01} (D)
1	2.93	846	1.38E+04	43.4	0.949	9.236	1	2.93	846	12800	40.3	0.936
2	4.13	600	4.01E+04	250.58	0.156	3.159	2	4.15	597	41600	262.5	0.148
3	4.77	520	6.00E+04	500.14	0.118	2.554	3	4.74	523	34000	279.9	0.102
4	4.91	505	1.79E+04	158.09	0.152	2.861	4	4.79	518	28500	239.6	0.047
5	5.12	484	1.49E+02	1.43	0.000	0.071	5	4.89	507	18200	159.4	0.140
6	5.56	446	1.62E+05	1834.71	0.039	1.366	6	5.05	491	643	6.0	0.006

M3							M4					
State.	ΔE	λ (nm)	δ^{2PA}	σ^{2PA}	f	μ_{01} (D)	State.	ΔE	λ (nm)	δ^{2PA}	σ^{2PA}	μ_{01} (D)
1	2.93	846	12400	39.0	0.939	9.192	1	2.95	841	11700	37.3	0.924
2	4.17	595	42200	268.8	0.148	3.061	2	3.77	657	3	0.0	0.000
3	4.73	524	24200	198.4	0.086	2.190	3	4.18	593	40400	258.6	0.152
4	4.79	518	40200	337.9	0.074	2.014	4	4.53	548	181	1.4	0.002
5	4.89	507	17000	148.9	0.126	2.607	5	4.55	545	546	4.1	0.002
6	5.04	492	527	4.9	0.006	0.577	6	4.77	520	52200	435.1	0.137

Table S7. State resolved electronic, one-photon absorption, and two-photon absorption properties of the 18PS Rhodopsin snapshot calculated using the M1 - M4 QM-region size under polarizable embedding at the BHandHLYP/aug-cc-pVDZ level of theory with D3(BJ) damping. Results are reported through S₆ for all QM-region sizes. Excitation energies, ΔE , are reported in eV; degenerate two-photon excitation wavelengths, λ , in nm; two-photon transition strengths, δ^{2PA} , in a.u.; two-photon absorption cross-sections, σ^{2PA} , in GM; and transition dipole moments, μ_{01} , in Debye. Oscillator strengths, f , are dimensionless and reported in the length gauge.

Citation:

- [1] Kandori, H. History and Perspectives of Ion-Transporting Rhodopsins in Optogenetics: Light-Sensing Proteins and Their Applications (eds Yawo, H., Kandori, H. & Koizumi, A.) 3–19 (Springer Japan, 2015).
- [2] Y. Shichida and T. Matsuyama, “Evolution of opsins and phototransduction,” *Philos. Trans. R. Soc. B Biol. Sci.*, vol. 364, no. 1531, pp. 2881–2895, 2009.
- [3] S. Costanzi, J. Siegel, I. G. Tikhonova, and K. A. Jacobson, “Rhodopsin and the others: A historical perspective on structural studies of G protein-coupled receptors,” *Curr. Pharm. Des.*, vol. 15, no. 35, pp. 3994–4002, 2009.
- [4] Boll F. Zur Anatomie und Physiologie der Retina. Monatsberichte der Koniglichen Preussischen Akademie der Wissenschaften zu Berlin. 1876:783–7.
- [5] Boll F. Zur Anatomie und Physiologie der Retina. Archiv der Anatomie und Physiologie. 1877:4–36.
- [6] Boll F. On the anatomy and physiology of the retina. Vision Res, vol. 17, pp. 1249–65, 1977.
- [7] Kühne W. Zur Photochimie der Netzhaut. Untersuchungen aus dem Physiologischen Institut der Universitaet Heidelberg. 1877:1–138.
- [8] Kühne W. Über den Sehpurpur. Untersuchungen aus dem Physiologischen Institut der Universitaet Heidelberg. 1878:15–113.
- [9] W. Kühne, “Chemische Vorgänge in der Netzhaut,” in L. Hermann, *Handbuch der Physiologie*, Vol. 3, Part 1 (Vogel, Leipzig, 1879), p. 235.
- [10] Kühne W. Chemical processes in the retina. Vision Res, vol. 17, pp. 1269–316, 1977.
- [11] G. Wald, “The photoreceptor process in vision,” *Am. J. Ophthalmol.*, vol. 40, no. 5 Pt 2, pp. 18–41, 1955.
- [12] G. Wald, “The molecular basis of visual excitation,” *Nature*, vol. 219, no. 5156, pp. 800–807, 1968.
- [13] J. E. Dowling and F. Ratliff, “Nobel prize: 3 named for medicine, physiology award (George Wald, Ragnar Granit and Haldan Keffer Hartline),” *Science*, vol. 158, no. 3800, pp. 468–473, 1967.
- [14] Björn, L. O. Photoenzymes and Light-Activated Enzymes. In Photobiology: The Science of Light and Life; Björn, L. O., Ed.; 3rd ed.; Springer Science+Business Media: New York, 2015; Chapter 11, pp 139–152.
- [15] Petrusevich, L. Computer-aided design of two-photon-absorbing organic chromophores. Doctoral dissertation, Wrocław University of Science and Technology (2023).
- [16] G. Palczewska, F. Vinberg, P. Stremplewski, M. P. Bircher, D. Salom, K. Komar, et al., “Human infrared vision is triggered by two-photon chromophore isomerization,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 111, no. 50, pp. E5445–E5454, 2014.
- [17] D. R. Griffin, R. Hubbard, and G. Wald, “The sensitivity of the human eye to infra-red radiation,” *J. Opt. Soc. Am.*, vol. 37, no. 7, pp. 546–554, 1947.
- [18] W. J. de Grip and S. Ganapathy, “Rhodopsins: An excitingly versatile protein species for research, development and creative engineering,” *Front. Chem.*, vol. 10, p. 879609, 2022.
- [19] Matsuno-Yagi A, Mukohata Y (1977) Two possible roles of bacteriorhodopsin; a comparative study of strains of Halobacterium halobium differing in pigmentation. Biochem Biophys Res Commun 78:237–243.
- [20] J. L. Spudich and R. A. Bogomolni, “Mechanism of colour discrimination by a bacterial sensory rhodopsin,” *Nature*, vol. 312, pp. 509–513, 1984.
- [21] K.-H. Jung, V. D. Trivedi, and J. L. Spudich, “Demonstration of a sensory rhodopsin in eubacteria,” *Mol. Microbiol.*, vol. 47, no. 6, pp. 1513–1522, 2003.
- [22] Nagel, G. et al. Channelrhodopsin-1: a light-gated proton channel in green algae. *Science* 296, 2395–2398 (2002).
- [23] Nagel, G. et al. Channelrhodopsin-2, a directly light-gated cation-selective membrane channel. *Proc. Natl. Acad. Sci. U. S. A.* 100, 13940–13945 (2003).
- [24] A. Terakita, “The opsins,” *Genome Biol.*, vol. 6, p. 213, 2005.
- [25] G. Palczewska, J. Boguslawski, P. Stremplewski, L. Kornaszewski, J. Zhang, Z. Dong, et al., “Noninvasive two-photon optical biopsy of retinal fluorophores,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 117, no. 36, pp. 22532–22543, 2020.
- [26] M. Broser, A. Spreen, P. E. Konold, E. Schiewer, S. Adam, V. Borin, et al., “NeoR, a near-infrared absorbing rhodopsin,” *Nat. Commun.*, vol. 11, p. 5682, 2020.
- [27] R. Palombo, L. Barneschi, L. Pedraza-González, D. Padula, I. Schapiro, M. Olivucci, et al., “Retinal chromophore charge delocalization and confinement explain the extreme photophysics of Neorhodopsin,” *Nat. Commun.*, vol. 13, p. 6652, 2022.
- [28] K. Deisseroth, “Controlling the brain with light,” *Sci. Am.*, vol. 303, no. 5, pp. 48–55, 2010.
- [29] K. Deisseroth, “Optogenetics,” *Nat. Methods*, vol. 8, no. 1, pp. 26–29, 2011.

- [30] O. P. Ernst, D. T. Lodowski, M. Elstner, P. Hegemann, L. S. Brown, and H. Kandori, "Microbial and animal rhodopsins: Structures, functions, and molecular mechanisms," *Chem. Rev.*, vol. 114, no. 1, pp. 126–163, 2014.
- [31] P. D. Kiser, M. Golczak, and K. Palczewski, "Chemistry of the retinoid (visual) cycle," *Chem. Rev.*, vol. 114, no. 1, pp. 194–232, 2014.
- [32] H. J. Bailes, L.-Y. Zhuang, and R. J. Lucas, "Reproducible and sustained regulation of Gas signalling using a metazoan opsin as an optogenetic tool," *PLoS One*, vol. 7, no. 1, p. e30774, 2012.
- [33] E. Oh, T. Maejima, C. Liu, E. Deneris, and S. Herlitze, "Substitution of 5-HT_{1A} receptor signaling by a light-activated G protein-coupled receptor," *J. Biol. Chem.*, vol. 285, no. 40, pp. 30825–30836, 2010.
- [34] X. Li et al., "Fast noninvasive activation and inhibition of neural and network activity by vertebrate rhodopsin and green algae channelrhodopsin," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 102, pp. 17816–17821, 2005.
- [35] R. D. Airan, K. R. Thompson, L. E. Fenno, H. Bernstein, and K. Deisseroth, "Temporally precise in vivo control of intracellular signalling," *Nature*, vol. 458, no. 7241, pp. 1025–1029, 2009.
- [36] Zhang, F. et al. Channelrhodopsin-2 and optical control of excitable cells. *Nature Meth.* 3, 785 –792 (2006).
- [37] E. S. Boyden, F. Zhang, E. Bamberg, G. Nagel, and K. Deisseroth, "Millisecond-timescale, genetically targeted optical control of neural activity," *Nat. Neurosci.*, vol. 8, no. 9, pp. 1263–1268, 2005.
- [38] F. Zhang, M. Prigge, F. Beyrière, S. P. Tsunoda, J. Mattis, O. Yizhar, et al., "Red-shifted optogenetic excitation: A tool for fast neural control derived from *Volvox carteri*," *Nat. Neurosci.*, vol. 11, pp. 631–633, 2008.
- [39] N. C. Klapoetke, Y. Murata, S. S. Kim, S. R. Pulver, A. Birdsey-Benson, Y. K. Cho, et al., "Independent optical excitation of distinct neural populations," *Nat. Methods*, vol. 11, no. 3, pp. 338–346, 2014.
- [40] J. H. Marshel, Y. S. Kim, T. A. Machado, S. Quirin, B. Benson, J. Kadmon, et al., "Cortical layer-specific critical dynamics triggering perception," *Science*, vol. 365, no. 6453, p. eaaw5202, 2019.
- [41] S. Huang, M. Ding, M. R. G. Roelfsema, I. Dreyer, S. Scherzer, K. A. S. Al-Rasheid, et al., "Optogenetic control of the guard cell membrane potential and stomatal movement by the light-gated anion channel GtACR1," *Sci. Adv.*, vol. 7, no. 28, p. eabg4619, 2021.
- [42] C. Córdova, C. Lozano, B. Rodríguez, I. Marchant, R. Zúñiga, P. Ochova, et al., "Optogenetic control of cancer cell survival in ChR2-transfected HeLa cells," *Int. J. Exp. Pathol.*, vol. 102, no. 6, pp. 242–248, 2021.
- [43] M. Sato, M. Kubo, T. Aizawa, N. Kamo, T. Kikukawa, K. Nitta, and M. Demura, "Role of putative anion-binding sites in cytoplasmic and extracellular channels of *Natronomonas pharaonis* halorhodopsin," *Biochemistry*, vol. 44, pp. 4775–4784, 2005.
- [44] C. Engelhard, I. Chizhov, F. Siebert, and M. Engelhard, "Microbial halorhodopsins: Light-driven chloride pumps," *Chem. Rev.*, vol. 118, no. 21, pp. 10629–10645, 2018, doi: 10.1021/acs.chemrev.7b00715.
- [45] X. Han and E. S. Boyden, "Multiple-color optical activation, silencing, and desynchronization of neural activity, with single-spike temporal resolution," *PLoS One*, vol. 2, no. 3, p. e299, 2007.
- [46] C. Zhang, S. Yang, T. Flossmann, et al., "Optimized photo-stimulation of halorhodopsin for long-term neuronal inhibition," *BMC Biol.*, vol. 17, p. 95, 2019.
- [47] O. A. Masseck, K. Spoida, D. Dalkara, T. Maejima, J. M. Rubelowski, L. Wallhorn, et al., "Vertebrate cone opsins enable sustained and highly sensitive rapid control of Gi/o signaling in anxiety circuitry," *Neuron*, vol. 81, no. 6, pp. 1263–1273, 2014.
- [48] A. Kaneko, K. Inoue, K. Kojima, H. Kandori, and Y. Sudo, "Conversion of microbial rhodopsins: Insights into functionally essential elements and rational protein engineering," *Biophys. Rev.*, vol. 9, no. 6, pp. 861–876, 2017.
- [49] J. Y. Lin et al., "ReaChR: A red-shifted variant of channelrhodopsin enables deep transcranial optogenetic excitation," *Nat. Neurosci.*, vol. 16, pp. 1499–1508, 2013.
- [50] T. Okitsu, Y. Yamano, Y.-C. Shen, T. Sasaki, Y. Kobayashi, S. Morisawa, et al., "Synthesis of one double bond-inserted retinal analogs and their binding experiments with opsins: Preparation of novel red-shifted channelrhodopsin variants," *Chem. Pharm. Bull. (Tokyo)*, vol. 68, no. 3, pp. 265–272, 2020.
- [51] S. Yasuda, T. Akiyama, K. Kojima, T. Ueta, T. Hayashi, S. Ogasawara, S. Nagatoishi, K. Tsumoto, N. Kunishima, Y. Sudo, M. Kinoshita, and T. Murata, "Development of an outward proton pumping rhodopsin

- with a new record in thermostability by means of amino acid mutations,” *J. Phys. Chem. B*, vol. 126, no. 5, pp. 1004–1015, 2022, doi: 10.1021/acs.jpcc.1c08684.
- [52] James, T. Y. et al. A molecular phylogeny of the flagellated fungi (Chytridiomycota) and description of a new phylum (Blastocladiomycota). *Mycologia* 98, 860–871 (2006).
- [53] K. Lehtinen, M. S. Nokia, and H. Takala, “Red light optogenetics in neuroscience,” *Front. Cell. Neurosci.*, vol. 15, p. 778900, 2022.
- [54] M. Broser, “Far-red absorbing rhodopsins, insights from heterodimeric rhodopsin-cyclases,” *Front. Mol. Biosci.*, vol. 8, p. 806922, 2022.
- [55] E. G. Govorunova, O. A. Sineshchekov, H. Li, Y. Wang, L. S. Brown, and J. L. Spudich, “RubyACRs, nonalgal anion channelrhodopsins with highly red-shifted absorption,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 117, no. 37, pp. 22833–22840, 2020.
- [56] M. Broser, T. Andruniów, A. Kraskov, R. Palombo, S. Katz, M. Klotz, et al., “Experimental assessment of the electronic and geometrical structure of a near-infrared absorbing and highly fluorescent microbial rhodopsin,” *J. Phys. Chem. Lett.*, vol. 14, no. 41, pp. 9291–9295, 2023.
- [57] D. M. Nikolaev, A. A. Shtyrov, S. Y. Vyazmin, A. V. Vasin, M. S. Panov, and M. N. Ryazantsev, “Fluorescence of the retinal chromophore in microbial and animal rhodopsins,” *Int. J. Mol. Sci.*, vol. 24, no. 24, p. 17269, 2023.
- [58] Taylor EW, Agarwal A (1993) Sequence homology between bacteriorhodopsin and G-protein coupled receptors: exon shuffling or evolution by duplication? *FEBS Lett* 325:161–166.
- [59] Oppa J (1994) Sequence comparison does not support an evolutionary link between halobacterial retinal proteins including bacteriorhodopsin and eukaryotic G-protein-coupled receptors. *FEBS Lett* 342:7–11.
- [60] T. Hara and R. Hara, “Vision in octopus and squid: Rhodopsin and retinochrome in the squid retina,” *Nature*, vol. 214, no. 5088, pp. 573–575, 1967.
- [61] I. M. Pepe and C. Cugnoli, “Retinal photoisomerase: Role in invertebrate visual cells,” *J. Photochem. Photobiol. B*, vol. 13, no. 1, pp. 5–17, 1992.
- [62] O. Vöcking, L. Leclère, and H. Hausen, “The rhodopsin-retinochrome system for retinal re-isomerization predates the origin of cephalopod eyes,” *BMC Ecol. Evol.*, vol. 21, no. 1, p. 215, 2021.
- [63] J. Zhang, E. H. Choi, A. Tworak, D. Salom, H. Leinonen, C. L. Sander, et al., “Photoc generation of 11-cis-retinal in bovine retinal pigment epithelium,” *J. Biol. Chem.*, vol. 294, no. 50, pp. 19137–19154, 2019.
- [64] Spudich JL, Yang CS, Jung KH et al. (2000) Retinylidene proteins: structures and functions from archaea to humans. *Annu Rev Cell Dev Biol* 16:365–392.
- [65] K. Inoue, T. Tsukamoto, and Y. Sudo, “Molecular and evolutionary aspects of microbial sensory rhodopsins,” *Biochim. Biophys. Acta Bioenerg.*, vol. 1837, no. 5, pp. 562–577, 2014.
- [66] Shichida Y, Imai H: Visual pigment: G-protein-coupled receptor for light signals. *Cell Mol Life Sci.* 1998, 54: 1299-1315. 10.1007/s000180050256. A review of the properties and functions of the various visual opsins.
- [67] P. D. Kiser, M. Golczak, A. Maeda, and K. Palczewski, “Key enzymes of the retinoid (visual) cycle in vertebrate retina,” *Biochim. Biophys. Acta Mol. Cell Biol. Lipids*, vol. 1821, no. 1, pp. 137–151, 2012.
- [68] R. Hara and T. Hara, “Squid m-retinochrome: Two forms of metaretinochrome,” *Vision Res.*, vol. 24, no. 11, pp. 1629–1640, 1984.
- [69] L. H. Andersen, I. B. Nielsen, M. B. Kristensen, M. O. A. El Ghazaly, S. Haacke, M. Brøndsted Nielsen, and M. Åxman Petersen, “Absorption of Schiff-base retinal chromophores in vacuo,” *J. Am. Chem. Soc.*, vol. 127, no. 35, pp. 12347–12350, 2005.
- [70] I. B. Nielsen, L. Lammich, and L. H. Andersen, “S1 and S2 excited states of gas-phase Schiff-base retinal chromophores,” *Phys. Rev. Lett.*, vol. 96, no. 1, p. 018304, 2006.
- [71] Y. Toker, D. B. Rahbek, H. V. Kiefer, J. Rajput, R. Antoine, P. Dugourd, et al., “Photoresponse of the protonated Schiff-base retinal chromophore in the gas phase,” *Phys. Chem. Chem. Phys.*, vol. 15, no. 45, pp. 19566–19569, 2013.
- [72] J. Rajput, D. B. Rahbek, L. H. Andersen, A. Hirshfeld, M. Sheves, P. Altoè, et al., “Probing and modeling the absorption of retinal protein chromophores in vacuo,” *Angew. Chem. Int. Ed.*, vol. 49, no. 10, pp. 1790–1793, 2010.
- [73] N. J. A. Coughlan, K. J. Catani, B. D. Adamson, U. Wille, and E. J. Bieske, “Photoisomerization action spectrum of retinal protonated Schiff base in the gas phase,” *J. Chem. Phys.*, vol. 140, no. 16, p. 164307, 2014.
- [74] Y. Toker and L. H. Andersen, “Gas-phase studies of the retinal protonated Schiff base chromophore,” *Eur. Phys. J. D*, vol. 75, no. 1, p. 7, 2021.

- [75] L. Hofmann and K. Palczewski, "Advances in understanding the molecular basis of the first steps in color vision," *Prog. Retin. Eye Res.*, vol. 49, pp. 46–66, 2015.
- [76] K. A. Freedman and R. S. Becker, "Comparative investigation of the photoisomerization of the protonated and unprotonated n-butylamine Schiff bases of 9-cis-, 11-cis-, 13-cis-, and all-trans-retinals," *J. Am. Chem. Soc.*, vol. 108, no. 6, pp. 1245–1251, 1986.
- [77] D. R. Pepperberg, "Generation of rhodopsin and 'artificial' visual pigments in electrophysiologically active photoreceptors," *Methods Enzymol.*, vol. 81, pp. 452–459, 1982.
- [78] U. M. Ganter, E. D. Schmid, D. Perez-Sala, R. R. Rando, and F. Siebert, "Removal of the 9-methyl group of retinal inhibits signal transduction in the visual process: A Fourier transform infrared and biochemical investigation," *Biochemistry*, vol. 28, no. 14, pp. 5954–5962, 1989.
- [79] Q. Wang, G. G. Kochendoerfer, R. W. Schoenlein, P. J. E. Verdegem, J. Lugtenburg, R. A. Mathies, and C. V. Shank, "Femtosecond spectroscopy of a 13-demethylrhodopsin visual pigment analogue: The role of

- nonbonded interactions in the isomerization process,” *J. Phys. Chem.*, vol. 100, no. 43, pp. 17388–17394, 1996.
- [80] R. Vogel, G. B. Fan, M. Sheves, and F. Siebert, “The molecular origin of the inhibition of transducin activation in rhodopsin lacking the 9-methyl group of the retinal chromophore: A UV–Vis and FTIR spectroscopic study,” *Biochemistry*, vol. 39, no. 30, pp. 8895–8908, 2000.
- [81] R. Vogel, S. Lüdeke, F. Siebert, T. P. Sakmar, A. Hirshfeld, and M. Sheves, “Agonists and partial agonists of rhodopsin: Retinal polyene methylation affects receptor activation,” *Biochemistry*, vol. 45, no. 6, pp. 1640–1652, 2006.
- [82] C. S. Irving, G. W. Byers, and P. A. Leermakers, “Spectroscopic model for the visual pigments: Influence of microenvironmental polarizability,” *Biochemistry*, vol. 9, no. 4, pp. 858–864, 1970.
- [83] P. E. Blatz, M. Lin, P. Balasubramanian, V. Balasubramanian, and P. B. Dewhurst, “A new series of synthetic visual pigments from cattle opsin and homologs of retinal,” *J. Am. Chem. Soc.*, vol. 91, no. 21, pp. 5930–5931, 1969.
- [84] G. G. Kochendoerfer, P. J. E. Verdegem, I. van der Hoef, J. Lugtenburg, and R. A. Mathies, “Retinal analog study of the role of steric interactions in the excited-state isomerization dynamics of rhodopsin,” *Biochemistry*, vol. 35, no. 50, pp. 16230–16240, 1996.
- [85] Y.-C. Shen, T. Sasaki, T. Matsuyama, T. Yamashita, Y. Shichida, T. Okitsu, et al., “Red-tuning of the channelrhodopsin spectrum using long conjugated retinal analogues,” *Biochemistry*, vol. 57, no. 38, pp. 5544–5556, 2018.
- [86] N. AzimiHashemi, K. Erbguth, A. Vogt, T. Riemensperger, E. Rauch, D. Woodmansee, et al., “Synthetic retinal analogues modify the spectral and kinetic characteristics of microbial rhodopsin optogenetic tools,” *Nat. Commun.*, vol. 5, p. 5810, 2014.
- [87] I. Dokukina and O. Weingart, “Spectral properties and isomerisation path of retinal in C1C2 channelrhodopsin,” *Phys. Chem. Chem. Phys.*, vol. 17, no. 38, pp. 25142–25150, 2015, doi: 10.1039/C5CP02650D.
- [88] N. H. List, J. M. H. Olsen, and J. Kongsted, “Excited states in large molecular systems through polarizable embedding,” *Phys. Chem. Chem. Phys.*, vol. 18, no. 30, pp. 20234–20250, 2016, doi: 10.1039/C6CP03834D.
- [89] K. Sneskov, J. M. H. Olsen, T. Schwabe, C. Hättig, O. Christiansen, and J. Kongsted, “Computational screening of one- and two-photon spectrally tuned channelrhodopsin mutants,” *Phys. Chem. Chem. Phys.*, vol. 15, no. 20, pp. 7567–7576, 2013, doi: 10.1039/C3CP44350G.
- [90] S. Gozem, F. Melaccio, R. Lindh, A. I. Krylov, A. A. Granovsky, C. Angeli, et al., “Mapping the excited state potential energy surface of a retinal chromophore model with multireference and equation-of-motion coupled-cluster methods,” *J. Chem. Theory Comput.*, vol. 9, no. 10, pp. 4495–4506, 2013.
- [91] O. Valsson, C. Angeli, and C. Filippi, “Excitation energies of retinal chromophores: Critical role of the structural model,” *Phys. Chem. Chem. Phys.*, vol. 14, pp. 11015–11020, 2012.
- [92] E. Walczak, B. Szeferczyk, and T. Andruniów, “Geometries and vertical excitation energies in retinal analogues resolved at the CASPT2 level of theory: Critical assessment of the performance of CASSCF, CC2, and DFT methods,” *J. Chem. Theory Comput.*, vol. 9, no. 11, pp. 4915–4927, 2013.
- [93] D. Grabarek, E. Walczak, and T. Andruniów, “Assessing the accuracy of various ab initio methods for geometries and excitation energies of retinal chromophore minimal model by comparison with CASPT3 results,” *J. Chem. Theory Comput.*, vol. 12, no. 5, pp. 2346–2356, 2016.
- [94] B. Szeferczyk et al., “Excited-state minima and emission energies of retinal,” *J. Comput. Chem.*, vol. 38, no. 20, pp. 1799–1810, 2017.
- [95] M. Sugihara, J. Hufen, and V. Buss, “Origin and consequences of steric strain in the rhodopsin binding pocket,” *Biochemistry*, vol. 45, no. 3, pp. 801–810, 2006, doi: 10.1021/bi0515624.
- [96] L. De Vico, C. S. Page, M. Garavelli, F. Bernardi, R. Basosi, and M. Olivucci, “Reaction path analysis of the ‘tunable’ photoisomerization selectivity of free and locked retinal chromophores,” *J. Am. Chem. Soc.*, vol. 124, no. 15, pp. 4124–4134, 2002, doi: 10.1021/ja017502c.
- [97] S. Gholami, L. Pedraza-González, X. Yang, A. A. Granovsky, et al., “Multi-state multi-configuration quantum chemical computation of the two-photon absorption spectra of bovine rhodopsin,” *J. Phys. Chem. Lett.*, vol. 10, no. 20, pp. 6293–6300, 2019.

- [98] M. G. Vivas, D. L. Silva, L. Misoguti, R. Zalesny, W. Bartkowiak, and C. R. Mendonça, “Degenerate two-photon absorption in all-trans retinal: Nonlinear spectrum and theoretical calculations,” *J. Phys. Chem. A*, vol. 114, no. 10, pp. 3466–3470, 2010, doi: 10.1021/jp910010g.
- [99] T. Ebrey and Y. Koutalos, “Vertebrate photoreceptors,” *Prog. Retin. Eye Res.*, vol. 20, no. 1, pp. 49–94, 2001, doi: 10.1016/S1350-9462(00)00014-8.
- [100] M. Hoffmann, M. Wanko, P. Strodel, P. H. König, T. Frauenheim, K. Schulten, W. Thiel, E. Tajkhorshid, and M. Elstner, “Color tuning in rhodopsins: The mechanism for the spectral shift between bacteriorhodopsin and sensory rhodopsin II,” *J. Am. Chem. Soc.*, vol. 128, no. 33, pp. 10808–10818, 2006, doi: 10.1021/ja062082i.
- [101] F. Collette, T. Renger, F. Müh, and M. Schmidt am Busch, “Red/green color tuning of visual rhodopsins: Electrostatic theory provides a quantitative explanation,” *J. Phys. Chem. B*, vol. 122, no. 18, pp. 4828–4837, 2018, doi: 10.1021/acs.jpcc.8b02702.
- [102] F. Melaccio, M. del C. Marín, A. Valentini, F. Montisci, S. Rinaldi, M. Cherubini, X. Yang, Y. Kato, M. Stenrup, Y. Orozco-Gonzalez, N. Ferré, H. L. Luk, H. Kandori, and M. Olivucci, “Toward automatic rhodopsin modeling as a tool for high-throughput computational photobiology,” *J. Chem. Theory Comput.*, vol. 12, no. 12, pp. 6020–6034, 2016, doi: 10.1021/acs.jctc.6b00367.
- [103] W. Wang, Z. Nossoni, T. Berbasova, C. T. Watson, I. Yapici, K. S. S. Lee, C. Vasileiou, J. H. Geiger, and B. Borhan, “Tuning the electronic absorption of protein-embedded all-trans-retinal,” *Science*, vol. 338, no. 6112, pp. 1340–1343, 2012, doi: 10.1126/science.1226135.
- [104] C. Cheng, M. Kamiya, Y. Uchida, and S. Hayashi, “Molecular mechanism of wide photoabsorption spectral shifts of color variants of human cellular retinol binding protein II,” *J. Am. Chem. Soc.*, vol. 137, no. 41, pp. 13362–13370, 2015, doi: 10.1021/jacs.5b08316.
- [105] M. Karasuyama, K. Inoue, R. Nakamura, H. Kandori, and I. Takeuchi, “Understanding colour tuning rules and predicting absorption wavelengths of microbial rhodopsins by data-driven machine-learning approach,” *Sci. Rep.*, vol. 8, no. 1, Art. no. 15580, 2018, doi: 10.1038/s41598-018-33984-w.
- [106] J. Hufen, M. Sugihara, and V. Buss, “How the counterion affects ground- and excited-state properties of the rhodopsin chromophore,” *J. Phys. Chem. B*, vol. 108, no. 52, pp. 20419–20426, 2004, doi: 10.1021/jp046147k.
- [107] A. Altun, S. Yokoyama, and K. Morokuma, “Spectral tuning in visual pigments: An ONIOM(QM:MM) study on bovine rhodopsin and its mutants,” *J. Phys. Chem. B*, vol. 112, no. 22, pp. 6814–6827, 2008, doi: 10.1021/jp709730b.
- [108] G. Tomasello, G. Olaso-González, P. Altoè, M. Stenta, L. Serrano-Andrés, M. Merchán, G. Orlandi, A. Bottoni, and M. Garavelli, “Electrostatic control of the photoisomerization efficiency and optical properties

- in visual pigments: On the role of counterion quenching,” *J. Am. Chem. Soc.*, vol. 131, no. 14, pp. 5172–5186, 2009, doi: 10.1021/ja808424b.
- [109] N. H. List, J. M. H. Olsen, H. J. Aa. Jensen, A. H. Steindal, and J. Kongsted, “Molecular-level insight into the spectral tuning mechanism of the DsRed chromophore,” *J. Phys. Chem. Lett.*, vol. 3, no. 23, pp. 3513–3521, 2012, doi: 10.1021/jz3014858.
- [110] M. Drobizhev et al., “Color hues in red fluorescent proteins are due to internal quadratic Stark effect,” *J. Phys. Chem. B*, 2010.
- [111] W. R. Zipfel, R. M. Williams, and W. W. Webb, “Nonlinear magic: Multiphoton microscopy in the biosciences,” *Nat. Biotechnol.*, vol. 21, pp. 1369–1377, 2003.
- [112] R. K. P. Benninger and D. W. Piston, “Two-photon excitation microscopy for the study of living cells and tissues,” *Curr. Protoc. Cell Biol.*, vol. 59, pp. 4.11.1–4.11.24, 2013.
- [113] M. Drobizhev, N. S. Makarov, S. E. Tillo, T. E. Hughes, and A. Rebane, “Two-photon absorption properties of fluorescent proteins,” *Nat. Methods*, vol. 8, no. 5, pp. 393–399, 2011.
- [114] Hosoi H, Yamaguchi S, Mizuno H, Miyawaki A, Tahara T. Hidden electronic excited-state of enhanced green fluorescent protein. *J. Phys. Chem. B*. 2008;112:2761–2763. DOI: 10.1021/jp711628u.
- [115] Hashimoto H, et al. Measurement of two-photon excitation spectra of fluorescent proteins with nonlinear Fourier-transform spectroscopy. *Appl. Opt.* 2010;49:3323–3329. DOI: 10.1364/AO.49.003323.
- [116] Blab GA, Lommerse PHM, Cognet L, Harms GS, Schmidt T. Two-photon excitation action cross-sections of the autofluorescent proteins. *Chem. Phys. Lett.*, vol. 350, pp. 71–77, 2001.
- [117] Heikal AA, Hess ST, Webb WW. Multiphoton molecular spectroscopy and excited-state dynamics of enhanced green fluorescent protein (EGFP): Acid-base specificity. *Chem. Phys.*, vol. 274, pp. 37–45, 2001.
- [118] Kawano H, Kogure T, Abe Y, Mizuno H, Miyawaki A. Two-photon dual-color imaging using fluorescent proteins. *Nat. Methods*. 2008;5:373–374. DOI: 10.1038/nmeth0508-373.
- [119] Rizzo MA, Springer G, Segawa K, Zipfel WR, Piston DW. Optimization of pairings and detection conditions for measurements of FRET between cyan and yellow fluorescent proteins. *Microsc. Microanal.* 2006;12:238–254. DOI: 10.1017/S1431927606060235.
- [120] M. Rumi and J. W. Perry, “Two-photon absorption: An overview of measurements and principles,” *Adv. Opt. Photon.*, vol. 2, no. 4, pp. 451–518, 2010.
- [121] Oulianov DA, Tomov IV, Dvornikov AS, Rentzepis PM. Observations on the measurement of two-photon absorption cross-section. *Opt. Comm*, vol. 191, pp. 235–243, 2001.
- [122] Kredel S, et al. Optimized and far-red-emitting variants of fluorescent protein eqFP611. *Chem. Biol.* 2008;15:224–233. DOI: 10.1016/j.chembiol.2008.02.008.
- [123] M. Drobizhev, S. E. Tillo, N. S. Makarov, T. E. Hughes, and A. Rebane, “Absolute two-photon absorption spectra and two-photon brightness of orange and red fluorescent proteins,” *J. Phys. Chem. B*, vol. 113, no. 4, pp. 855–859, 2009.
- [124] H. Lischka, D. Nachtigallová, A. J. A. Aquino, P. G. Szalay, F. Plasser, F. B. C. Machado, et al., “Multireference approaches for excited states of molecules,” *Chem. Rev.*, vol. 118, no. 15, pp. 7293–7361, 2018.
- [125] P. B. Coto, A. Strambi, N. Ferré, and M. Olivucci, “The color of rhodopsins at the ab initio multiconfigurational perturbation theory resolution,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 103, no. 46, pp. 17154–17159, 2006.
- [126] E. Walczak and T. Andruniów, “Impacts of retinal polyene (de)methylation on the photoisomerization mechanism and photon energy storage of rhodopsin,” *Phys. Chem. Chem. Phys.*, vol. 17, pp. 17169–17181, 2015.
- [127] E. D. Hedegård and M. Reiher, “Polarizable embedding density matrix renormalization group,” *J. Chem. Theory Comput.*, vol. 12, no. 9, pp. 4242–4253, 2016.
- [128] D. Manawadu, D. J. Valentine, and W. Barford, “Dynamical simulations of carotenoid photoexcited states using density matrix renormalization group techniques,” *J. Phys. Chem. A*, vol. 127, no. 16, pp. 3714–3727, 2023.
- [129] S. Sirimatayanant and T. Andruniów, “Benchmarking two-photon absorption strengths of rhodopsin chromophore models with CC3 and CCSD methodologies: An assessment of popular density functional approximations,” *J. Chem. Phys.*, vol. 158, no. 9, p. 094106, 2023.
- [130] S. Sirimatayanant and T. Andruniów, “Tuning two-photon absorption in rhodopsin chromophore via backbone modification: The story told by CC2 and TD-DFT,” *J. Chem. Theory Comput.*, vol. 20, no. 18, pp. 8118–8126, 2024.

- [131] M. T. P. Beerepoot, D. H. Friesse, N. H. List, J. Kongsted, and K. Ruud, "Benchmarking two-photon absorption cross sections: Performance of CC2 and CAM-B3LYP," *Phys. Chem. Chem. Phys.*, vol. 17, no. 29, pp. 19306–19314, 2015.
- [132] M. A. Salem and A. Brown, "Two-photon absorption in fluorescent protein chromophores: TDDFT and CC2 results," *J. Chem. Theory Comput.*, vol. 10, no. 8, pp. 3260–3269, 2014.
- [133] D. Grabarek and T. Andruniów, "Assessment of functionals for TDDFT calculations of one- and two-photon absorption properties of neutral and anionic fluorescent proteins chromophores," *J. Chem. Theory Comput.*, vol. 15, no. 1, pp. 490–508, 2019.
- [134] D. Grabarek and T. Andruniów, "Initial excited-state relaxation of locked retinal protonated Schiff base chromophore: An insight from coupled cluster and multireference perturbation theory calculations," *J. Comput. Chem.*, vol. 39, no. 22, 2018.
- [135] O. Valsson and C. Filippi, "Photoisomerization of model retinal chromophores: Insight from quantum Monte Carlo and multiconfigurational perturbation theory," *J. Chem. Theory Comput.*, vol. 6, no. 4, pp. 1275–1292, 2010.
- [136] A. J. A. Aquino, M. Barbatti, and H. Lischka, "Excited-state properties and environmental effects for protonated Schiff bases: A theoretical study," *ChemPhysChem*, vol. 7, no. 10, pp. 2089–2096, 2006.
- [137] Ł. Wolański, D. Grabarek, and T. Andruniów, "Is the choice of a standard zeroth-order Hamiltonian in CASPT2 ansatz optimal in calculations of excitation energies in protonated and unprotonated Schiff bases of retinal?," *J. Comput. Chem.*, vol. 39, no. 20, pp. 1470–1480, 2018.
- [138] M. J. Paterson, O. Christiansen, F. Pawłowski, P. Jørgensen, C. Hättig, T. Helgaker, and P. Salek, "Benchmarking two-photon absorption with CC3 quadratic response theory, and comparison with density-functional response theory," *J. Chem. Phys.*, vol. 124, no. 5, p. 054322, 2006.
- [139] M. Schreiber, M. R. Silva-Junior, S. P. A. Sauer, and W. Thiel, "Benchmarks for electronically excited states: CASPT2, CC2, CCSD, and CC3," *J. Chem. Phys.*, vol. 128, no. 13, p. 134110, 2008.
- [140] K. D. Nanda and A. I. Krylov, "Two-photon absorption cross sections within equation-of-motion coupled-cluster formalism using resolution-of-the-identity and Cholesky decomposition representations: Theory, implementation, and benchmarks," *J. Chem. Phys.*, vol. 142, no. 6, p. 064118, 2015.
- [141] C. Naim, R. Zaleśny, and D. Jacquemin, "Two-photon absorption strengths of small molecules: Reference CC3 values and benchmarks," *J. Chem. Theory Comput.*, vol. 20, no. 20, pp. 9093–9106, 2024.
- [142] A. D. Laurent and D. Jacquemin, "TD-DFT benchmarks: A review," *Int. J. Quantum Chem.*, vol. 113, no. 17, pp. 2019–2039, 2013.
- [143] J. P. Perdew, K. Schmidt, "Jacob's Ladder of Density Functional Approximations for the Exchange-Correlation Energy," *AIP Conference Proceedings* 577, 1 (2001).
- [144] M. J. G. Peach, D. J. Tozer, *J. Phys. Chem. A* 2012, 116, 9783.
- [145] J. S. Sears, T. Koerzdoerfer, C. R. Zhang, J. L. Brédas, *J. Chem. Phys.* 2011, 135, 151103.
- [146] T. Tawada, T. Tsuneda, S. Yanagisawa, T. Yanai, K. Hirao, *J. Chem. Phys.* 2004, 120, 8425.
- [147] M. Miura, Y. Aoki, B. Champagne, *J. Chem. Phys.* 2007, 127, 084103.
- [148] Silva-Junior, M. R.; Schreiber, M.; Sauer, S. P. A.; Thiel, W. *J. Chem. Phys.* 2008, 129, 104103.
- [149] Schreiber, M.; Silva-Junior, M. R.; Sauer, S. P. A.; Thiel, W. *J. Chem. Phys.* 2008, 128, 134110.
- [150] M. Caricato, G. Trucks, M. Frisch, K. Wiberg, *J. Chem. Theory Comput.* 2010, 7, 456.
- [151] F. Furche and R. Ahlrichs, "Adiabatic time-dependent density functional methods for excited state properties," *J. Chem. Phys.*, vol. 117, no. 16, pp. 7433–7447, 2002, doi: 10.1063/1.1508368.
- [152] R. A. King, "On the accuracy of computed excited-state dipole moments," *J. Phys. Chem. A*, vol. 112, no. 25, pp. 5727–5733, 2008, doi: 10.1021/jp800867x.
- [153] E. Tapavicza, I. Tavernelli, U. Rothlisberger, *J. Phys. Chem. A* 2009, 113, 9595.
- [154] A. Masunov, S. Tretiak, *J. Phys. Chem. B* 2004, 108, 899.
- [155] D. H. Friesse, C. Hättig, K. Ruud, *Phys. Chem. Chem. Phys.* 2012, 14, 1175.
- [156] S. Knippenberg, D. R. Rehn, M. Wormit, J. H. Starcke, I. L. Rusakova, A. B. Trofimov, A. Dreuw, *J. Chem. Phys.* 2012, 136, 064107.
- [157] K. Fujimoto, S. Hayashi, J.-Y. Hasegawa, and H. Nakatsuji, "Theoretical studies on the color-tuning mechanism in retinal proteins," *J. Chem. Theory Comput.*, vol. 3, no. 2, pp. 605–618, 2007.
- [158] O. Valsson, P. Campomanes, I. Tavernelli, U. Rothlisberger, and C. Filippi, "Rhodopsin absorption from first principles: Bypassing common pitfalls," *J. Chem. Theory Comput.*, vol. 9, no. 5, pp. 2441–2454, 2013.
- [159] López, C. S.; Faza, O. N.; Esté vez, S. L.; de Lera, A. R. *J. Comput. Chem.* 2006, 27, 116–123.
- [160] Campomanes, P.; Horta, B.; Neri, M.; Vanni, S.; Tavernelli, I.; Rothlisberger, U. Private communication, 2012.

- [161] Silva-Junior, M. R.; Thiel, W. *J. Chem. Theory Comput.* 2010, 6, 1546–1564.
- [162] D. Grabarek and T. Andruniów, “What is the optimal size of the quantum region in embedding calculations of two-photon absorption spectra of fluorescent proteins?” *J. Chem. Theory Comput.*, vol. 16, no. 10, pp. 6439–6455, 2020.
- [163] F. Jensen, *Introduction to Computational Chemistry*, 2nd ed. Chichester, UK: John Wiley & Sons Ltd, 2007.
- [164] P. Hohenberg and W. Kohn, “Inhomogeneous Electron Gas,” *Physical Review*, vol. 136, no. 3B, pp. B864–B871 (November 9, 1964), DOI: 10.1103/PhysRev.136.B864 27,36-40.
- [165] L. E. Ratcliff, S. Mohr, G. Huhs, T. Deutsch, M. Masella, and L. Genovese, “Challenges in large scale quantum mechanical calculations”, *WIREs Comput. Mol. Sci.*, vol. 7, p. e1290, 2017.
- [166] A. M. Khah, P. Reinholdt, J. M. H. Olsen, J. Kongsted, and C. Hättig, “Avoiding electron spill-out in QM/MM calculations on excited-states with simple pseudopotentials,” *J. Chem. Theory Comput.*, vol. 16, pp. 1373–1381, 2020.
- [167] D. Grabarek and T. Andruniów, “Removing artifacts in polarizable embedding calculations of one- and two-photon absorption spectra of fluorescent proteins,” *J. Chem. Phys.*, vol. 153, no. 21, p. 215102, 2020.
- [168] P. Reinholdt, J. Kongsted, and J. M. H. Olsen, “Polarizable density embedding: A solution to the electron spill-out problem in multiscale modeling,” *J. Phys. Chem. Lett.*, vol. 8, pp. 5949–5958, 2017.
- [169] I. Rivalta, A. Nenov, and M. Garavelli, “Modelling retinal chromophores photoisomerization: From minimal models in vacuo to ultimate bidimensional spectroscopy in Rhodopsins,” *Phys. Chem. Chem. Phys.* 16, 16865–16879 (2014) (DOI: 10.1039/c3cp55211j).
- [170] J. J. Szymczak, M. Barbatti, and H. Lischka, “Mechanism of ultrafast photodecay in restricted motions in protonated Schiff bases: The pentadieniminium cation,” *J. Chem. Theory Comput.* 4, 1189–1199 (2008) (DOI: 10.1021/ct800148n).
- [171] P. Zhou, J. Liu, K. Han, and G. He, “The photoisomerization of 11-cis-retinal protonated Schiff base in gas phase: Insight from spin-flip density functional theory,” *J. Comput. Chem.* 35, 109–120 (2013) (DOI: 10.1002/jcc.23463).
- [172] S. Gozem, A. I. Krylov, and M. Olivucci, “Conical intersection and potential energy surface features of a model retinal chromophore: Comparison of EOM-CC and multireference methods,” *J. Chem. Theory Comput.* 9, 284–292 (2012) (DOI: 10.1021/ct300759z).
- [173] J. J. Szymczak, M. Barbatti, and H. Lischka, “Is the photoinduced isomerization in retinal protonated Schiff bases a single or double-torsional process?,” *J. Phys. Chem. A* 113, 11907–11918 (2009) (DOI: 10.1021/jp903329j).
- [174] I. F. Galván, M. Vacher, A. Alavi, C. Angeli, F. Aquilante, J. Autschbach, J. J. Bao, S. I. Bokarev, N. A. Bogdanov, R. K. Carlson et al., “OpenMolcas: From source code to insight,” *J. Chem. Theory Comput.* 15, 5925–5964 (2019).
- [175] F. Aquilante, J. Autschbach, A. Baiardi, S. Battaglia, V. A. Borin, L. F. Chibotaru, I. Conti, L. De Vico, M. Delcey, F. I. Galván et al., “Modern quantum chemistry with [open]MOLCAS,” *J. Chem. Phys.* 152, 214117 (2020).
- [176] T. Yanai, D. P. Tew, and N. C. Handy, “A new hybrid exchange–correlation functional using the coulomb-attenuating method (CAM-B3LYP),” *Chem. Phys. Lett.* 393, 51–57 (2004).
- [177] A. D. Becke, “New mixing of Hartree–Fock and local density-functional theories,” *J. Chem. Phys.* 98, 1372–1377 (1993).
- [178] C. Adamo and V. Barone, “Toward reliable adiabatic connection models free from adjustable parameters,” *Chem. Phys. Lett.* 274, 242–250 (1997).
- [179] A. D. Becke, “Density-functional thermochemistry. III. The role of exact exchange,” *J. Chem. Phys.* 98, 5648–5652 (1993).
- [180] C. Lee, W. Yang, and R. G. Parr, “Development of the Colle–Salvetti correlation-energy formula into a functional of the electron density,” *Phys. Rev. B* 37, 785–789 (1988).
- [181] A. D. Becke, “Density-functional exchange-energy approximation with correct asymptotic behavior,” *Phys. Rev. A* 38, 3098–3100 (1988).
- [182] K. Aidas, C. Angeli, K. L. Bak, V. Bakken, R. Bast, L. Boman, O. Christiansen, R. Cimiraglia, S. Coriani, P. Dahle et al., “The Dalton quantum chemistry program system,” *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 4, 269–284 (2013).
- [183] See <http://daltonprogram.org> for Dalton, a molecular electronic structure program release v2020.0 (2020).
- [184] A. Köhn and C. Hättig, “Analytic gradients for excited-states in the coupled cluster model CC2 employing the resolution-of-the-identity approximation,” *J. Chem. Phys.* 119, 5021–5036 (2003).

- [185] A. Dreuw and M. Head-Gordon, "Failure of time-dependent density functional theory for long-range charge-transfer excited states: The zincbacteriochlorin–bacteriochlorin and bacteriochlorophyll–spheroidene complexes," *J. Am. Chem. Soc.*, vol. 126, pp. 4007–4016, 2004.
- [186] A. J. Cohen, P. Mori-Sánchez, and W. Yang, "Insights into current limitations of density functional theory," *Science*, vol. 321, no. 5890, pp. 792–794, 2008.
- [187] M. Chołuj, M. M. Alam, M. T. P. Beerepoot, S. P. Sitkiewicz, E. Matito, K. Ruud, and R. Zaleśny, "Choosing bad versus worse: Predictions of two-photon-absorption strengths based on popular density functional approximations," *J. Chem. Theory Comput.*, vol. 18, no. 2, pp. 1046–1060, 2022.
- [188] J. J. Eriksen, S. P. A. Sauer, K. V. Mikkelsen, O. Christiansen, H. J. A. Jensen, and J. Kongsted, "Failures of TDDFT in describing the lowest intramolecular charge-transfer excitation in para-nitroaniline," *Mol. Phys.*, vol. 111, nos. 9–11, pp. 1235–1248, 2013.
- [189] J. P. Perdew and A. Zunger, "Self-interaction correction to density-functional approximations for many-electron systems," *Phys. Rev. B*, vol. 23, pp. 5048–5079, 1981.
- [190] P. Mori-Sánchez, A. J. Cohen, and W. Yang, "Localization and delocalization errors in density functional theory and implications for band-gap prediction," *Phys. Rev. Lett.*, vol. 100, p. 146401, 2008.
- [191] M. J. G. Peach, P. Benfield, T. Helgaker, and D. J. Tozer, "Excitation energies in density functional theory: An evaluation and a diagnostic test," *J. Chem. Phys.*, vol. 128, p. 044118, 2008.
- [192] *J. Am. Chem. Soc.* 2009, 131, 8, 2818–2820 DOI: 10.1021/ja8087482 – (Stein, Kronik, Baer) Reliable Prediction of Charge Transfer Excitations in Molecular Complexes Using Time-Dependent Density Functional Theory.
- [193] P. J. E. Verdegem, P. H. M. Bovee-Geurts, W. J. de Grip, J. Lugtenburg, and H. J. M. de Groot, "Retinylidene ligand structure in bovine rhodopsin, metarhodopsin-I, and 10-methylrhodopsin from internuclear distance

- measurements using ^{13}C -labeling and 1-D rotational resonance MAS NMR,” *Biochemistry*, vol. 38, no. 35, pp. 11316–11324, 1999.
- [194] Gartner, W.; Ternieden, S. Influence of a steric hindrance in the chromophore of Rhodopsin on the quantum yield of the primary photochemistry. *J. Photochem. Photobiol., B* 1996, 33 (1), 83–86, DOI: 10.1016/1011-1344(95)07225-X.
- [195] Koch, D.; Gartner, W. Steric hindrance between chromophore substituents as the driving force of Rhodopsin photoisomerization: 10-methyl-13-demethyl retinal containing Rhodopsin. *Photochem. Photobiol.* 1997, 65 (1), 181–186, DOI: 10.1111/j.1751-1097.1997.tb01896.x.
- [196] DeLange, F.; Bovee-Geurts, P. H.; Portier, J. V. M. D.; Verdegem, P. J.; Lugtenburg, J.; de Grip, W. J. An additional methyl group at the 10-position of retinal dramatically slows down the kinetics of the Rhodopsin photocascade. *Biochemistry* 1998, 37 (5), 1411–1420, DOI: 10.1021/bi972397y.
- [197] Grotjahn, R.; Furche, F. Gauge-invariant excited-state linear and quadratic response properties within the meta-generalized gradient approximation. *J. Chem. Theory Comput.* 2023, 19 (15), 4897–4911, DOI: 10.1021/acs.jctc.3c00259.
- [198] Ahmadzadeh, K.; Li, X.; Rinkevicius, Z.; Norman, P.; Zaleśny, R. Toward accurate two-photon absorption spectrum simulations: Exploring the landscape beyond the generalized gradient approximation. *J. Phys. Chem. Lett.* 2024, 15 (4), 969–974, DOI: 10.1021/acs.jpclett.3c03513.
- [199] I. A. Elayan, L. Rib, R. A. Mendes, and A. Brown, “Beyond explored functionals: A computational journey of two-photon absorption,” *J. Chem. Theory Comput.*, vol. 20, no. 9, pp. 3879–3893, 2024.
- [200] S. G. Balasubramani et al., “TURBOMOLE: Modular program suite for ab initio quantum-chemical and condensed-matter simulations,” *J. Chem. Phys.*, vol. 152, no. 18, p. 184107, 2020.
- [201] TURBOMOLE V7.5 2020, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989–2007; TURBOMOLE GmbH, since 2007; available from <https://www.turbomole.org>, 2020.
- [202] Jacquemin, D. Excited-state dipole and quadrupole moments: Td-dft versus cc2. *J. Chem. Theory Comput.* 2016, 12, 3993–4003, DOI: 10.1021/acs.jctc.6b00498.
- [203] L. Gagliardi, R. Lindh, and G. Karlström, “Local properties of quantum chemical systems: The LoProp approach,” *J. Chem. Phys.*, vol. 121, pp. 4494–4500, 2004.
- [204] I. F. Galván et al., “OpenMolcas: From source code to insight,” *J. Chem. Theory Comput.*, vol. 15, pp. 5925–5964, 2019.
- [205] J. M. H. Olsen, “PyFraME: Python tools for Fragment-based Multiscale Embedding,” version 0.2.0, Zenodo, 2018. DOI: 10.5281/zenodo.1443314.
- [206] D. W. Zhang and J. Z. H. Zhang, “Molecular fractionation with conjugate caps for full quantum mechanical calculation of protein-molecule interaction energy,” *J. Chem. Phys.*, vol. 119, pp. 3599–3605, 2003.
- [207] P. Söderhjelm and U. Ryde, “How accurate can a force field become? A polarizable multipole model combined with fragment-wise quantum-mechanical calculations,” *J. Phys. Chem. A*, vol. 113, pp. 617–627, 2009.
- [208] Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M. Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins,

- Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* 1995, 117, 5179–5197. DOI: 10.1021/ja00124a002.
- [209] W. L. Jorgensen, J. Chandrasekhar, and J. D. Madura, “Comparison of simple potential functions for simulating liquid water,” *J. Chem. Phys.*, vol. 79, p. 926, 1983.
- [210] J. Hutter, M. Iannuzzi, F. Schiffmann, and J. Vande-Vondele, “CP2K: Atomistic simulations of condensed matter systems”, *WIREs Comput. Mol. Sci.*, vol. 4, pp. 15–25, 2014.
- [211] Shirts, M. R.; Klein, C.; Swails, J. M.; Yin, J.; Gilson, M. K.; Mobley, D. L.; Case, D. A.; Zhong, E. D. Lessons learned from comparing molecular dynamics engines on the SAMPL5 dataset. *J. Comput.-Aided Mol. Des.* 2017, 31, 147–161. (Published online 2016). DOI: 10.1007/s10822-016-9977-1.
- [212] Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general AMBER force field. *J. Comput. Chem.* 2004, 25, 1157–1174. DOI: 10.1002/jcc.20035.
- [213] G. Lippert, J. Hutter, and M. Parrinello, “The Gaussian and augmented-plane-wave density functional method for ab initio molecular dynamics simulations”, *Theor. Chem. Acc.*, vol. 103, pp. 124–140, 1999.
- [214] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, “A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H- Pu,” *J. Chem. Phys.*, vol. 132, p. 154104, 2010.
- [215] D. Jacquemin, V. Wathelet, E. A. Perpète, and C. Adamo, “Extensive TD-DFT benchmark: Singlet-excited states of organic molecules,” *J. Chem. Theory Comput.*, vol. 5, no. 9, pp. 2420–2435, 2009.
- [216] A. Dreuw and M. Head-Gordon, “Single-reference ab initio methods for the calculation of excited states of large molecules,” *Chem. Rev.*, vol. 105, no. 11, pp. 4009–4037, 2005.
- [217] R. A. Mathies and J. Lugtenburg, “The primary photoreaction of rhodopsin,” in *Molecular Mechanisms in Visual Transduction*, D. G. Stavenga, W. J. de Grip, and E. N. Pugh Jr., Eds., *Handbook of Biological Physics*, vol. 3. Amsterdam: Elsevier Science B.V., pp. 55–90, 2000.
- [218] A. Altun, S. Yokoyama, and K. Morokuma, “Mechanism of spectral tuning going from retinal in vacuo to bovine rhodopsin and its mutants: Multireference ab initio quantum mechanics/molecular mechanics studies,” *J. Phys. Chem. B*, vol. 112, no. 51, pp. 16883–16890, 2008, doi: 10.1021/jp807172h.

Scientific Output

The following publications and conference presentations arose from the research conducted as part of this doctoral thesis.

Peer-Reviewed Journal Articles

1. Saruti Sirimatayanant and Tadeusz Andruniów, “Benchmarking two-photon absorption strengths of rhodopsin chromophore models with CC3 and CCSD methodologies: An assessment of popular density functional approximations,” *The Journal of Chemical Physics*, vol. 158, article 094106, pp. 1–8, 2023. DOI: 10.1063/5.0135594.
2. Saruti Sirimatayanant and Tadeusz Andruniów, “Tuning two-photon absorption in rhodopsin chromophore via backbone modification: The story told by CC2 and TD-DFT,” *Journal of Chemical Theory and Computation*, vol. 20, no. 18, pp. 8118–8126, 2024. DOI: 10.1021/acs.jctc.4c00675.

Conference Publication

1. Saruti Sirimatayanant and Tadeusz Andruniów, “OPA/TPA spectra of the wt-rhodopsin calculated using TD-DFT/MM with polarizable embedding,” in *PANIC Summer School 2022: PhoBiA Annual Nanophotonics International Conference*, Wrocław, Poland, 30 May–3 June 2022. Wrocław: Publishing House of Wrocław University of Science and Technology, 2022, p. 41.

Oral Conference Presentations

1. Saruti Sirimatayanant and Tadeusz Andruniów, “TPA spectra of the retinal chromophore model calculated with TD-DFT and coupled-cluster methods,” Photonics and Bionanotechnology Association Annual Nanophotonics International Conference (PANIC), hosted online by Wrocław University of Science and Technology, 26–28 May 2021.
2. Saruti Sirimatayanant and Tadeusz Andruniów, “TPA spectra of the retinal chromophore model calculated with TD-DFT and coupled-cluster methods,” 8th Doctoral Symposium on Chemistry, University of Łódź, Łódź, Poland, 24 September 2021.
3. Saruti Sirimatayanant and Tadeusz Andruniów, “OPA/TPA spectra of the wt-rhodopsin calculated using QM/MM with polarizable embedding,” PhoBiA Annual Nanophotonics International Conference (PANIC), Wrocław University of Science and Technology, Wrocław, Poland, 30 May–3 June 2022.
4. Saruti Sirimatayanant and Tadeusz Andruniów, “Two-photon absorption spectrum of rhodopsin calculated using TD-DFT/MM with a polarizable embedding scheme,” Modeling and Design of Molecular Materials (MDMM), University of Gdańsk, Gdańsk, Poland, 19–22 September 2022.

5. Saruti Sirimatayanant and Tadeusz Andruniów, "Two-photon absorption spectrum of rhodopsin calculated using TD-DFT/MM with a polarizable embedding scheme," Workshop "Sequence-Regulated Polymers: Towards Functional Systems Mimicking Living Matter," MimicLS Kick-Off Meeting, PORT Polski Ośrodek Rozwoju Technologii, Wrocław, Poland, 30 January 2023.
6. Saruti Sirimatayanant and Tadeusz Andruniów, "Spectral tuning in reduced models of retinal protonated Schiff base via modulation of the methyl groups," Konferencja Naukowa „Chemia we Wrocławsku” WrocChem 2023, Wrocław University of Science and Technology, Wrocław, Poland, 12 June 2023.
7. Saruti Sirimatayanant and Tadeusz Andruniów, "Investigating the effects of the immediate environment of rhodopsin on its 2PA signal with QM/MM," 66. Zjazd Polskiego Towarzystwa Chemicznego and XI Kongres Technologii Chemicznej, Adam Mickiewicz University, Poznań, Poland, 15–20 September 2024.
8. Saruti Sirimatayanant and Tadeusz Andruniów, "Investigating the effects of the immediate environment of rhodopsin on its 2PA signal with QM/MM," Online Seminar of Theoretical Chemists "Hamiltonian," University of Wrocław, Wrocław, Poland, 16 October 2024.
9. Saruti Sirimatayanant and Tadeusz Andruniów, "Investigating the effects of the immediate environment of rhodopsin on its 2PA signal with QM/MM," VIII Conference "Od teorii do eksperymentu: w poszukiwaniu Świętego Graala w nauce 2," University of Wrocław, Wrocław, Poland, 12 February 2025.
10. Saruti Sirimatayanant and Tadeusz Andruniów, "Investigating the effects of the immediate environment of rhodopsin on its 2PA and 3PA signals with QM/MM," Modeling and Design of Molecular Materials 2025, Wrocław, Poland, 7–11 September 2025.