

# Computational Study of Two Photon Absorption in Fluorescent Protein Chromophores

by

Mohammad Alsayed Alaraby Abdelsalam Salem

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Chemistry

University of Alberta

©Mohammad Alsayed Alaraby Abdelsalam Salem, 2016

# Abstract

Two-photon absorption (TPA) microscopy of fluorescent proteins (FPs) is a powerful bio-imaging tool. Advantages of TPA microscopy include better focus and less out-of-focus bleaching, together with absorption at longer wavelengths than in one-photon absorption (OPA), which leads to deeper penetration in tissues. However, TPA probes are usually associated with less sensitivity than OPA alternatives and thus designing fluorophores with large TPA probability (cross section) is an important area of research. A great variety of FPs have been synthesized from canonical amino acids and characterized for both their OPA and TPA properties. Although a small number of non-canonical amino acids (ncAAs) have been utilized in designing new FPs, they were not characterized for their TPA properties. In addition, incorporating ncAAs in FPs is a demanding task and thus preceding the experiment with a computational rationalization that guides the choice of ncAAs is prudent. The unique light-absorbing and fluorescence ability of FPs is due to the formation of a chromophore by a post-translational modification of three precursory amino acids within the protein shell. While the protein environment can strongly affect the photophysical properties of the chromophore, the goal of this work is to highlight ncAA-modified chromophores that have computationally large intrinsic TPA cross sections. Given the size of the chromophores, time-dependent density functional theory (TD-DFT) is the method of choice to scan their TPA properties. The TD-DFT results (using four functionals) on a given set of natural chromophores were compared to a wave-function-based method (CC2) and to averaged experimental data from the FPs; unlike OPA data, TPA measurements on isolated chromophore analogues have not been made. The comparison shows that TD-DFT with B3LYP underestimates the

absolute TPA cross sections, but can be used in a semi-quantitative fashion to study the trends across various structures or the effect of conformational change on the TPA of a given structure. TD-DFT at the B3LYP/6-31+G(d,p) level of theory was further used to screen twenty-two possible chromophores that can be formed upon replacing a precursory amino acid (Tyr66) from those that form the green FP (GFP) chromophore with a ncAA. A proposed chromophore with a nitro substituent was found to have a large TPA cross section (29 GM) that is more than 7 times that of the native GFP chromophore as determined at the same level of theory. Classical molecular dynamics performed on a nitro-modified FP confirmed its stability and the large TPA cross section of the chromophore at various conformations it assumed within the protein pocket. Intrigued by the recent interest in designing ncAA-derived red FPs (RFPs), the same set of GFP-based chromophores, but with an extended structure (an acylimine moiety) that is characteristic for many RFPs, were screened for their TPA properties. In the screening of these RFP-derived chromophores, both B3LYP and CAM-B3LYP functionals were used together with re-screening the GFP-derived ones with CAM-B3LYP for completeness. Computing TPA cross-sections with B3LYP and CAM-B3LYP yield similar overall trends. Results using both functionals agree that the RFP-derived model of the gold FP (GdFP) chromophore has the largest intrinsic TPA cross section (50 GM according to B3LYP). TPA was further computed for selected chromophores following conformational changes: variation of the dihedral angle of the acylimine moiety and the tilt and twist angles between the rings of the chromophore. The TPA cross-section assumed an oscillatory trend with the rotation of the acylimine dihedral, and the TPA is maximized in the planar conformation for almost all models. One chromophore bearing a hydroxyquinoline ring is also shown to be comparable to that of the GdFP-like chromophore in terms of TPA cross section. The conformational study on the hydroxyquinoline-modified chromophore shows that the acylimine angle has a much stronger effect on the TPA than its tilt and twist angles. Having an intrinsic TPA ability that is more than 7 times that of the native RFP chromophore, the GdFP- and the hydroxyquinoline- modified chromophores are very promising for experimental investigation. The GFP-derived

chromophore with a nitro substituent is likewise interesting. Overall, a number of new FP chromophores built from ncAAs, with large intrinsic TPA, have been proposed and the strong effect of conformation on TPA explored. The present work will hopefully spur complementary experimental tests in this burgeoning field.

# Preface

Chapter 1 is adapted from the book review that has been published as: M. Alaraby Salem, Melis Gedik and Alex Brown, *Handbook of Computational Chemistry*, Ed. Jerzy Leszczynski. Springer Netherlands, 2016, 1-19. I wrote the introduction, discussion of theory and computation, the section on fluorescent proteins, and the conclusion. Melis Gedik wrote the section on nucleic acid bases.

Chapter 2 is adapted from the published paper: M. Alaraby Salem and Alex Brown, *J. Chem. Theory Comput.*, 2014, **10**, 3260-3269. In the published paper, the reported TPA cross sections were too large by a factor of 4 when given in macroscopic units (GM) and thus rescaled when reported in Chapter 2.

A version of Chapter 3 is published as M. Alaraby Salem and Alex Brown, *Phys. Chem. Chem. Phys.*, 2015, **17**, 25563-25570.

A version of Chapter 4 is published as M. Alaraby Salem, Isaac Twelves and Alex Brown, *Phys. Chem. Chem. Phys.*, 2016, Accepted. Isaac Twelves is an undergraduate researcher who helped in running GAMESS jobs, extracting information from crystal structures, and writing the introduction of the manuscript.

To my mother, Ebtisam I. Hussein and my motherland, Egypt.

# Acknowledgements

I would like to express my gratitude to my PhD supervisor, Professor Alex Brown. In short, he is a too-good-to-be-true supervisor. I would also like to thank my supervisory committee: Professor Mariusz Klobukowski, the first one to teach me computational chemistry, and Professor Robert Campbell who was very supportive with discussions on fluorescent proteins. I extend my thanks to all members of the Brown Group: Ryan, Stephanie, Ekadashi, Shuai, Mohammad and Melis. I also thank all the undergraduate members for the excellent atmosphere they helped create in the Brown Group, especially the ones I mentored: Andrew, Isaac, and Wenyu.

To my loving mother, Ebtesam I. Hussein, who really paved all avenues for me to do my graduate studies: I can not thank you enough. You have fulfilled your role via pushing me to where you always wished I could reach. I acknowledge I would have never been where I am standing now without your support. I wish to thank my beloved wife who accompanied me in this journey. We came to Edmonton as a couple and now we are a family of five. Thank you for taking care of me and our children through this period and thank you for enduring my shortcomings.

I deeply thank all my friends who lived with me in Edmonton, especially Ahmed T. Ayoub and Ahmad S. Abdelfattah. This PhD journey would have lacked lots of its flavour without you. You were always by my side and you will never be forgotten. Last, but not least, I wholeheartedly thank my high school friend, Mahmoud Hassan, who has always provided me with genuine and unconditional support.

# Table of Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Introduction . . . . .	1
1.2	Theory . . . . .	2
1.2.1	The Two-level Model . . . . .	4
1.2.2	Computational Implementation . . . . .	5
1.3	Applications . . . . .	7
1.3.1	Fluorescent Proteins . . . . .	7
1.3.2	Nucleic Acid Bases . . . . .	14
1.4	Conclusions . . . . .	18
1.5	Scope of the thesis . . . . .	18
<b>2</b>	<b>Chapter2</b>	<b>20</b>
2.1	Introduction . . . . .	20
2.2	Computational Methods . . . . .	23
2.2.1	Chromophore Structures . . . . .	23
2.2.2	Theory and Computations . . . . .	24
2.3	Results and Discussion . . . . .	27
2.3.1	Lowest-Energy Transition to $S_1$ . . . . .	28
2.3.2	High-Energy Transitions to $S_n$ . . . . .	33
2.4	Conclusions . . . . .	42
<b>3</b>	<b>Chapter3</b>	<b>45</b>
3.1	Introduction . . . . .	45
3.2	Computational Methods . . . . .	49
3.2.1	Chromophores . . . . .	49
3.2.2	Ab initio computations . . . . .	51
3.2.3	Molecular Dynamics Simulation . . . . .	52
3.3	Results and Discussion . . . . .	53
3.3.1	TPA cross sections . . . . .	54
3.3.2	Protein Stability . . . . .	57
3.3.3	Conformational Analysis of The Chromophore . . . . .	57
3.4	Conclusion . . . . .	61

<b>4</b>	<b>Chapter4</b>	<b>64</b>
4.1	Introduction . . . . .	64
4.2	Computational Methods . . . . .	67
4.3	Results and Discussion . . . . .	70
	4.3.1 TPA cross sections . . . . .	71
	4.3.2 TPA change with acylimine rotation . . . . .	76
	4.3.3 TPA change with twist and tilt . . . . .	78
4.4	Conclusion . . . . .	80
<b>5</b>	<b>Conclusions</b>	<b>82</b>
5.1	Summary of Thesis Research . . . . .	82
5.2	Future Perspective . . . . .	84
<b>A</b>	<b>Appendix to Chapter 2</b>	<b>99</b>
<b>B</b>	<b>Appendix to Chapter 3</b>	<b>128</b>
<b>C</b>	<b>Appendix to Chapter 4</b>	<b>157</b>

# List of Tables

1.1	Wavelengths and cross sections of the modified uridine analogues . . .	17
2.1	Energies and cross sections for $S_0$ to $S_1$ transitions . . . . .	29
2.2	Cross sections with different methods for RFP chromophore . . . . .	32
2.3	Energies and cross sections for $S_0$ to $S_1$ transitions of RFP chromophore	35
2.4	Energies and cross sections for $S_0$ to $S_1$ transitions of other chromophores	36
2.5	Dipole-moment elements contributing to higher transitions of RFP chromophore . . . . .	38
3.1	Excited-state properties for $S_0$ to $S_1$ transition of GFP-based chromophores . . . . .	55
3.2	Dipole elements of the $S_0$ to $S_1$ transition of GFP-based chromophores	56
3.3	Main values for the angles studied in the nitro-modified chromophore	59
4.1	One-photon excitation energies [in eV], OPA oscillator strengths and TPA cross sections [in GM] for the transition to $S_1$ as determined at the B3LYP/6-31G+(d,p) level of theory and PCM with parameters for $H_2O$ . Between brackets is the difference between the property computed for the RFP-derived chromophore and that previously computed for the corresponding GFP-derived chromophore. <sup>75</sup> . . . . .	74
A1	Coordinates of the optimized (H-capped) natural chromophore models	99
A2	Coordinates of the optimized (methyl-capped) natural RFP chromophore model . . . . .	104
A3	Excited-state properties for 8 transitions of natural chromophores (PCM)	106
A4	Lambda diagnostic and OPA oscillator strength for 8 transitions of natural chromophores . . . . .	108
A5	Excited-state properties for 8 transitions of natural chromophores (gas phase) . . . . .	110
A6	Summary of experimental data used for comparison . . . . .	112
A7	Orbitals involved in the studied excitations for the RFP <sub>A</sub> chromophore	114
A8	Orbitals involved in the studied excitations for the RFP <sub>A</sub> chromophore	115
A9	Orbitals involved in the studied excitations for the BFP1 chromophore	116
A10	Orbitals involved in the studied excitations for the CFP chromophore	117

A11	Orbitals involved in the studied excitations for the GFP <sub>A</sub> chromophore	118
A12	Orbitals involved in the studied excitations for the KO <sub>A</sub> chromophore	119
A13	Orbitals involved in the studied excitations for the OR <sub>A</sub> chromophore	120
A14	Orbitals involved in the studied excitations for the BLB chromophore	121
A15	Grouped excitations for a subset of natural chromophores	122
A16	Grouped excitations for a subset of natural chromophores	124
A17	Grouped excitations for the non-planar model of RFP <sub>A</sub> chromophore	127
B1	Coordinates of the optimized (H-capped) GFP-based chromophore models	128
B2	Atom Types and charges for the <i>control</i> model	142
B3	Atom Types and charges for the <i>nitro</i> model	143
B4	The $\delta^{TPA}$ values for the GFP-based chromophores	149
B5	Comparison of the average bond lengths of the <i>control</i> chromophore	150
B6	Comparison of the average bond lengths of the <i>nitro</i> chromophore	151
B7	PCM and Gas phase computations for selected chromophores	152
B8	The z-matrix used in the conformational analysis of the nitro-chromophore. The entries: meth, tilt and twist refer to the methyl bridge, tilting and twisting angles referred to in the main text.	153
B9	difference between permanent moments for the nitro-modified chromophore	153
C1	Tilt, twist and acylimine ( $\theta_{\text{acylimine}}$ ) dihedral angles for red fluorescent chromophores in their native protein crystal structures. Also included are the two other dihedral angles ( $\theta_1$ and $\theta_2$ ) defining the geometry of the extended conjugation chain.	159
C2	Coordinates of the optimized (mainly methyl-capped) RFP-based chromophore models	160
C3	One-photon excitation energies [in eV] and the corresponding wavelengths [in nm], OPA oscillator strengths, TPA transition moments ( $\delta^{TPA}$ ) and TPA cross-sections [in GM] for the transition to S <sub>1</sub> as determined at the B3LYP/6-31G+(d,p) Level of Theory and PCM with parameters for H <sub>2</sub> O. The properties for RFP-derived chromophores (Figure C1) are computed in Chapter 4, while the GFP-derived ones were previously reported (see Chapter 3). <sup>75</sup>	183
C4	Differences between the properties computed for the RFP-derived chromophore and those previously computed for the corresponding GFP-derived chromophore for the transition to S <sub>1</sub> as in Table C3.	183
C5	One-photon excitation energies [in eV] and the corresponding wavelengths [in nm], OPA oscillator strengths, TPA transition moments ( $\delta^{TPA}$ ) and TPA cross-sections [in GM] for the transition to S <sub>1</sub> as determined at the B3LYP and CAM-B3LYP/6-31G+(d,p) Level of Theory in the gas phase for the RFP-derived chromophores given in Figure C1.	184

C6	TPA cross sections via the truncated SOS approach using 2 states; i.e. the 2-level model (2LM), and via quadratic response (QR), and the contributing components to the 2LM expression ((Equation 5 in the main text)) determined at the B3LYP and CAM-B3LYP/6-31+G(d,p) levels of theory in the gas phase. . . . .	186
C7	One-photon excitation energies [in eV] and the corresponding wavelengths [in nm], OPA oscillator strengths, TPA transition moments ( $\delta^{\text{TPA}}$ ) and TPA cross-sections [in GM] for the transition to $S_1$ as determined at the CAM-B3LYP/6-31G+(d,p) Level of Theory and PCM with parameters for $H_2O$ . Models for the RFP-derived chromophores are given in Figure C1. Equivalent GFP-derived chromophores are given in Figure 2 of Chapter 3. <sup>75</sup> . . . . .	187
C8	Differences between the properties computed for the RFP-derived chromophore and those previously computed for the corresponding GFP-derived chromophore for the transition to $S_1$ as in Table C7. . . . .	187
C9	The $\Lambda$ -diagnostic for the first excited state computed at the B3LYP and CAM-B3LYP/6-31+G(d,p) level of theory in PCM ( $H_2O$ ) for the RFP-derived models (see Figure C1). . . . .	188
C10	One-photon excitation energies [in eV] and the corresponding wavelengths [in nm], OPA oscillator strengths, TPA transition moments ( $\delta^{\text{TPA}}$ ) and TPA cross-sections [in GM] for the transition to $S_1$ as determined at the B3LYP/6-31G+(d,p) Level of Theory (except where noted) and PCM with parameters for $H_2O$ . The properties are computed for the RFP-derived chromophores (Figure C1) at varying Acylimine dihedral angle ( $\theta_{\text{acylimine}}$ ). . . . .	189
C11	TPA Cross-sections [in GM] extracted from Tables C10 and C12 as determined at the B3LYP/6-31G+(d,p) Level of Theory (except where noted) with PCM ( $H_2O$ ). The Cross-sections correspond to the $S_0$ to $S_1$ transition except for Models 22 and BLB at the near-planar conformations ( $\theta_{\text{acylimine}} = 0, 10, 20, 340$ and $350$ ) where the transition is to the second excited state. . . . .	201
C12	TD-DFT data for the first 3 states of Models 22 and BLB at their near-planar conformations ( $\theta_{\text{acylimine}} = 0, 10, 20, 340$ and $350$ ). Level of Theory is B3LYP/6-31G+(d,p) with PCM ( $H_2O$ ). . . . .	202
C13	Components contributing to the 2LM expression (Equation 5 in the main text) determined at the B3LYP and CAM-B3LYP/6-31+G(d,p) levels of theory in the gas phase for Model 21 at $\theta_{\text{acylimine}}$ varying from $0^\circ$ to $90^\circ$ . All data is for the $S_0$ to $S_1$ transition except for the CAM-B3LYP data for conformers with $\theta_{\text{acylimine}} = 0, 10$ in which data for the $S_0$ to $S_2$ is given because the first excited-state was dark. . . . .	203

C14	Variation of TD-DFT properties with tilt and twist angles (see Table C1) for Model 21 at fixed $\theta_{\text{acylimine}}$ of $0^\circ$ , $90^\circ$ , $180^\circ$ and $270^\circ$ . The TD-DFT values are for the transition to $S_1$ as determined at the B3LYP/6-31G+(d,p) level of theory and PCM with parameters for $\text{H}_2\text{O}$ . . . .	205
-----	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----

# List of Figures

1.1	GFP chromophore highlighting the head and tail positions that vary between FP chromophores. . . . .	9
1.2	Structures of modified uridine nucleosides . . . . .	16
2.1	Models for chromophores named after common fluorescent proteins they represent. . . . .	25
3.1	Chromophore maturation in GFP. . . . .	46
3.2	GFP-based chromophores made from ncAAs . . . . .	50
3.3	RMSD and RMSF of the reference and modified proteins . . . . .	58
3.4	An overlay of average structures of the reference and modified proteins . . . . .	59
3.5	Angles studied in the nitro-modified chromophore . . . . .	59
3.6	Variation of TPA cross sections and OPA oscillator strengths with angles . . . . .	60
3.7	Variation of permanent dipoles with angles . . . . .	62
4.1	A RFP chromophore model showing various angles that were varied in the TPA conformational study. The dihedral angle of the acylimine moiety ( $\theta_{\text{acylimine}}$ ) is made by atoms $O_1$ , $C_1$ , $N_1$ and $C_2$ . The twist angle is the dihedral between atoms $C_3$ , $C_4$ , $C_5$ and $C_6$ while the tilt angle is the dihedral between atoms $C_4$ , $C_5$ , $C_6$ and $N_2$ . Model chromophores are generated by replacing the highlighted part with moieties derived from nCAAs. A subset of these moieties is given below the parent structure while the full list is given in Figure C1 in Appendix C. . . . .	69
4.2	TPA cross sections of the GFP- and RFP-derived chromophores for the transition to $S_1$ as determined at the B3LYP/6-31G+(d,p) and the CAM-B3LYP/6-31G+(d,p) levels of theory with PCM ( $H_2O$ ). For B3LYP, cross sections of the RFP- and GFP-derived chromophores are given in Table 4.1 and taken from our previous work, <sup>75</sup> respectively. The data for cross sections computed with CAM-B3LYP are given in Table C7 in Appendix C. . . . .	75
4.3	TPA cross sections calculated via the components of the 2-level model (2LM) and via quadratic response at the B3LYP/6-31+G(d,p) level of theory in the gas phase. The contributing components to the 2LM expression are given in Table C6 in Appendix C. . . . .	76

4.4	Variation of TPA with rotation of the dihedral angle of the acylimine moiety ( $\theta_{\text{acylimine}}$ in Figure 4.1). The TPA values are for the transition to $S_1$ (with some exceptions noted in the main text) as determined at the B3LYP/6-31G+(d,p) level of theory and PCM with parameters for $H_2O$ . The dashed curve corresponds to a scan of the RFP model using CAM-B3LYP at the same level of theory. Only the two dotted curves were generated using H-capping while all other computations were done on $CH_3$ -capped chromophores. . . . .	79
4.5	Variation of TPA cross sections with tilt and twist angles (see Figure 4.1) for Model 21 at fixed $\theta_{\text{acylimine}}$ of $0^\circ$ , $90^\circ$ , $180^\circ$ and $270^\circ$ . The TPA values are for the transition to $S_1$ as determined at the B3LYP/6-31G+(d,p) level of theory and PCM with parameters for $H_2O$ . . . . .	80
B1	Variation of Energy for various conformers of the nitro-modified chromophore . . . . .	152
C1	Chromophore models built from the RFP-parent structure. RFP and BLB refer to the DsRed and mBlueberry chromophore models, respectively. The rest are built via the replacement of the phenol ring of Tyr-66 with the corresponding moiety in a non-canonical amino acid. The numbering matches the corresponding previously studies models. <sup>75</sup>	158
C2	A visual representation for the TPA cross sections computed in the gas phase for the RFP-derived chromophores as given in Table C5 with the corresponding PCM values as given in Tables C3 and C7. . . . .	184
C3	Plots for the B3LYP (left) and CAM-B3LYP (right) data given in Table C6. Top figures compare the TPA cross sections computed via quadratic response (QR) to those obtained from the 2LM approximation. Middle figures compare the magnitudes of the transition dipole moment ( $\mu_{01}$ ) and the difference between permanent moments of excited and ground states ( $\mu_{01}$ ). The alignment between the $\mu_{01}$ and $\Delta\mu$ vectors is plotted in the bottom figures in terms of $\cos^2\theta$ (the vectors are perfectly aligned when $\cos^2\theta = 1$ ). . . . .	185
C4	Top: TPA cross sections computed for Model 21 (at $\theta_{\text{acylimine}}$ varying from $0^\circ$ to $90^\circ$ ) via the 2LM in the gas phase (as given in Table C13) versus the corresponding PCM values for B3LYP (given in Table C11). Bottom: The magnitudes of the dipole moments contributing to the 2LM expression (as given in Table C13). . . . .	204
C5	Variation of OPA oscillator strengths with tilt and twist angles (see Figure 1 in the main text) for Model 21 at fixed $\theta_{\text{acylimine}}$ of $0^\circ$ , $90^\circ$ , $180^\circ$ and $270^\circ$ . The OPA values are for the transition to $S_1$ as determined at the B3LYP/6-31G+(d,p) level of theory and PCM with parameters for $H_2O$ . . . . .	215

# List of Abbreviations and Acronyms

2-AP	2-aminopurine
2LM	Two-level model
6-MAP	4-amino-6-methyl- 8-(2'-deoxy- $\beta$ -D-ribofuranosyl)-7(8H)-pteridone
6-MI	6-methylisoxanthopterin
ADC	Algebraic diagrammatic construction
aug-cc-pVDZ	Augmented correlation-consistent valence double-zeta
avGFP	<i>Aequorea victoria</i> green fluorescent protein
B3LYP	Becke, three-parameter, Lee-Yang-Parr
BLB	mBlueberry1 chromophore model
CAM-B3LYP	Coulomb-attenuating method - Becke, three-parameter, Lee-Yang-Parr
CC	Coupled cluster
CC2	Second order approximate coupled cluster singles and doubles
CC3	Approximate coupled cluster singles, doubles and triples
CCS	Coupled cluster singles model
CCSD	Coupled cluster singles and doubles model
CD	Cholesky decomposition
CFP	Cyan fluorescent protein
DFT	Density functional theory
DNA	Deoxyribonucleic acid
DsRED	<i>Discosoma sp.</i> red fluorescent protein
ECFP	Enhanced cyan fluorescent protein
EGFP	Enhanced green fluorescent protein

EOM	Equation of motion
eV	Electron Volt
FP	Fluorescent protein
GdFP	Gold fluorescent protein
GFP	Green fluorescent protein
GM	Göppert-Mayer
HOMO	Highest occupied molecular orbital
ICT	Intermolecular charge-transfer
KO <sub>A</sub>	Kusabira-Orange fluorescent protein
LC-BLYP	Long-range corrected BLYP functional
LUMO	Lowest unoccupied molecular orbital
LYS	Lysine
MAD	Mean absolute difference
MD	Molecular dynamics
ncAA	Non-canonical amino acid
OPA	One-Photon Absorption
OR <sub>A</sub>	mOrange fluorescent protein
OS	Oscillator strength
PBE0	hybrid Perdew, Burke, and Ernzerhof exchange-correlation DFT functional
PCM	Conductor-like polarizable continuum model
PDB	Protein Data Bank
PE	Polarizable embedding
PE-DFT	Polarizable embedding - Density Functional Theory
QM/MM	Quantum mechanics and molecular mechanics
RFP	red fluorescent protein
RI	Resolution of identity
RI-CC2	Resolution of identity - second order approximate coupled clusters singles
RMSD	Root mean square deviation
RMSF	Root mean square fluctuation
RNA	Ribonucleic acid

Ser	Serine
SOS	Sum over states
tC	tricyclic cytosine
TD-	Time dependent
TD-DFT	Time dependent density functional theory
TPA	Two-photon absorption
Tyr	Tyrosine
YFP	Yellow fluorescent protein

# List of Symbols

$\sigma^{\text{TPA}}$	Two-photon absorption cross section
$\delta^{\text{OPA}}$	One-photon absorption transition moment
$\delta^{\text{TPA}}$	Two-photon absorption transition moment
$\omega$	(photon) Energy
$\omega_n$	Energy of the $n^{\text{th}}$ excited state
$\mu$	Dipole moment
$\Delta\mu_{0n}$	Difference between the permanent dipole moments of the $n^{\text{th}}$ excited state and the ground state
$a_0$	Bohr radius
$\alpha$	Fine structure constant
$c$	Speed of light
$\Gamma$	lifetime broadening
$S_{\alpha\beta}$	Element of the two-photon transition matrix
$S_n$	$n^{\text{th}}$ electronic level
$P$	Induced polarization
$E$	Electric field
$\chi^n$	$n^{\text{th}}$ order electric susceptibility
$\hbar$	Reduced Planck constant
$\Lambda$	Overlap diagnostic quantity

# Chapter 1

## Two-photon Absorption: Theory and Applications For Biological Imaging<sup>1</sup>

### 1.1 Introduction

Two photon absorption (TPA) is defined as the simultaneous absorption of two photons, of the same or different energies, leading to excitation to a higher electronic state. Although this phenomenon was predicted theoretically in 1931 by Maria Göppert-Mayer,<sup>1</sup> experimental evidence was not available until 1961<sup>2</sup> after the development of lasers. Since TPA is a third-order nonlinear process, where the absorption is directly proportional to the square of incident light intensity, intense light sources, i.e., lasers, are required to observe significant TPA. The quadratic dependence on light intensity provides better focus and less out-of-focus bleaching and thus deeper penetration in scattering media, such as tissues, as compared to one-photon absorption (OPA).<sup>3,4</sup> However, there is a drawback as TPA probes are usually associated with less sensitivity. This drove the desire to design TPA fluorophores with large TPA probabilities; see, for example, the reviews by Pawlicki et al.<sup>5</sup> and Refs. 6 and 7 for more recent work. In TPA, the transition to the excited electronic state is achieved by two photons and thus the energy of incident photons is approximately half of those used for

---

<sup>1</sup>This chapter is adapted (with some additions) from the published book chapter: M. Alaraby Salem, Melis Gedik and Alex Brown, *Handbook of Computational Chemistry*, Ed. Jerzy Leszczynski. Springer Netherlands, 2016, 1-19.

OPA in the same system. Therefore, the excited state leading to fluorescence can be accessed at longer wavelengths; desirable in biological media where the absorption in the IR or near-IR region is required to overcome the competitive absorption of other naturally-present pigments. Moreover, TPA follows different quantum mechanical selection rules from OPA and therefore both spectra can be used in a complementary fashion.<sup>4</sup>

The focus of this introductory chapter is the TPA of biological molecules which is gaining more interest as exemplified by the increasing number of recent studies cited here. Optimizing the photophysical TPA properties of proteins and other molecules that could be naturally present in biological systems is of paramount advantage. These molecules are produced through hijacking the native transcription mechanism in the cell and thus they are usually more benign to the complex cell environment than synthetically introduced dyes or quantum dots that may have other interactions and/or cell-toxicity. Moreover, no membrane penetration is needed in the case of endogenously-produced biological molecules. For more information on the advantages of using biological molecules, see, for instance, Ref. 8 about fluorescent proteins and the references therein.

In this chapter, we review the theory of TPA highlighting the computational approaches used to study biological molecules. We then discuss the computational studies on the TPA of fluorescent proteins which are of central importance in modern biological imaging and the TPA of nucleic acid base analogues which provide a newer imaging tool that appears to be very promising. We conclude each section with open questions that should drive more computational (and experimental) work. This introductory chapter is concluded with a summary of the research that will be discussed in the thesis.

## 1.2 Theory

The theory of multi photon absorption has been reviewed by Cronstrand et al.<sup>9</sup> Here, we present a brief description of the theory of TPA. The theoretical description of

TPA can be understood by considering the interaction of an external electric field,  $E$ , with a dielectric material. The induced polarization,  $P$ , within the dielectric material is given by

$$P = \varepsilon_0\chi^{(1)}E + \varepsilon_0\chi^{(2)}E^2 + \varepsilon_0\chi^{(3)}E^3 + \dots, \quad (1.1)$$

where  $\varepsilon_0$  is the permittivity of free-space and  $\chi^{(n)}$  is  $n^{\text{th}}$  order electric susceptibility:  $\chi^{(1)}$  represents the linear susceptibility and the quantities  $\chi^{(2)}$ ,  $\chi^{(3)}$ ,  $\dots$ , denote the nonlinear susceptibilities. The two-photon absorption probability (or the cross section,  $\sigma^{\text{TPA}}$ ) is related to the imaginary part of the third order susceptibility, as

$$\sigma^{\text{TPA}} = \frac{24\pi^2\hbar\omega^2}{c^2} \text{Im}(\chi^{(3)}), \quad (1.2)$$

where  $\omega$  is the photon energy,  $c$  is the speed of light and  $\hbar$  is the (reduced) Planck constant.

Susceptibilities are closely related to (hyper)polarizabilities which can, in principle, be computed by quantum mechanical methods.<sup>9</sup> However, a great simplification, can be achieved when the photon energy is such that near-resonant conditions are reached.<sup>9</sup> The two-photon transition matrix can then be defined as

$$S_{\alpha\beta} = \sum_n \left[ \frac{\langle 0|\mu_\alpha|n\rangle\langle n|\mu_\beta|f\rangle}{\omega_n - \omega} + \frac{\langle 0|\mu_\beta|n\rangle\langle n|\mu_\alpha|f\rangle}{\omega_n - \omega} \right] \quad (1.3)$$

where  $\mu_\alpha$  (or  $\mu_\beta$ ) refers to the dipole moment operator in one cartesian direction ( $\alpha, \beta = x, y$  and  $z$ ),  $\omega_n$  is the energy spacing from the ground state,  $|0\rangle$ , to the intermediate state  $|n\rangle$ , and  $|f\rangle$  is the final excited state.

Although this sum-over-states expression (SOS) has been used to compute TPA (see, for example, Ref. 10), it is generally not computationally efficient due to its slow convergence. On the other hand, response theory offers a more computationally efficient framework through the residue analysis of (hyper)polarizabilities. The two-photon transition matrix elements,  $S_{\alpha\beta}$ , and the transition dipole moments between excited states can be deduced from the single and double residues of the quadratic response function, respectively.<sup>11</sup>

For linearly polarized light, the transition moment for TPA is

$$\delta^{\text{TPA}} = \frac{1}{15} \sum_{\alpha\beta} [S_{\alpha\alpha}S_{\beta\beta}^* + 2S_{\alpha\beta}S_{\alpha\beta}^*]. \quad (1.4)$$

It can then be shown that the transition moment is related to the previous definition in equation 1.2 by

$$\sigma^{TPA} = \frac{24\pi^2\hbar\omega^2}{c^2} \text{Im}(\chi^{(3)}) = \frac{4\pi^3 a_0^5 \alpha \omega^2}{c} \delta^{TPA} \quad (1.5)$$

where  $\alpha$  is the fine structure constant and  $a_0$  is the Bohr radius. The two-photon absorption cross section can be related to experimental measurements by including a term for life-time broadening,  $\frac{1}{\pi\Gamma}$ , derived from a Lorentzian function:

$$\sigma^{TPA} = \frac{4\pi^2 a_0^5 \alpha \omega^2}{c} \frac{1}{\Gamma} \delta^{TPA}, \quad (1.6)$$

where  $\omega = \frac{\omega_f}{2}$  is the photon energy and  $\Gamma$  is the broadening factor. In previous computations,  $\Gamma$  has been chosen as 0.1 eV<sup>9,12,13</sup> or 0.5 eV<sup>14,15</sup> for comparison with experiment. The TPA cross section,  $\sigma^{TPA}$ , is conveniently reported in units of 10<sup>-50</sup> cm<sup>4</sup> s molecule<sup>-1</sup> photon<sup>-1</sup> that is referred to as one GM (which stands for ‘‘Göppert-Mayer’’). Care should be taken when values of computed cross sections are compared in microscopic units due to variations caused by different forms of Eq. 1.6. In a recent comprehensive study, Beerepoot et al.<sup>16</sup> discuss the reasons behind these variations and give recommendations for the details that should be provided when TPA cross sections are reported: (i) Excitation energies and TPA strengths in atomic units, (ii) the formulae for the conversion to macroscopic units (Eq. 1.6) and that of the transition moment (Eq. 1.4), (iii) the proper choice of the conversion to match the experiments (here, we use the full definition for  $\delta^{TPA}$  and scale Eq. 1.6 for comparison with single-beam experiments), and (iv) the type of lineshape function (Lorentzian or Gaussian) and the broadening factor. The interested reader is invited to examine the details of the discussion provided by Beerepoot et al.<sup>16</sup>

### 1.2.1 The Two-level Model

The SOS expression (Eq. 1.3) can be truncated to include only the initial and final (two) states; the so-called two-level model (2LM). More generally, the most significant transitions in an excitation scheme can be considered in an  $N$ -level model. Such

truncations can be used to relate TPA cross sections to transition dipole moments and permanent dipole moments of the ground state and excited states.<sup>17-19</sup>

For the first excited state ( $f = 1$ ), and setting  $\omega_0 = 0$ , the SOS expression can be written within the two-level model as

$$S_{\alpha\beta} = \frac{2}{\omega_1} [\mu_{\beta 01} (\mu_{\alpha 11} - \mu_{\alpha 00}) + \mu_{\alpha 01} (\mu_{\beta 11} - \mu_{\beta 00})]. \quad (1.7)$$

For the same Cartesian direction ( $\alpha = \beta$ ), it can be expressed as

$$S_{\alpha\alpha} = \frac{4}{\omega_1} \mu_{\alpha 01} (\mu_{\alpha 11} - \mu_{\alpha 00}). \quad (1.8)$$

From this expression and Eqs. (3) and (4), it is clear that (within the two-level model) the TPA cross section is proportional to the squares of the transition dipole moment and the difference between the permanent dipole moments of the excited and ground states. Hence, structural or environmental effects that increase the extent of charge delocalization upon excitation (while having a large transition dipole moment) would increase the TPA of a system (see Section 1.3.1).

## 1.2.2 Computational Implementation

In this section, key computational implementations of TPA theory within electronic structure programs that can be used for studying biological molecules are highlighted. The interested reader is directed to see the original references for a more in-depth presentation of the computational details. In principle, any method that can provide excited-state properties, i.e., transition dipole moments between and permanent dipole moments for excited states, can be used to compute TPA cross sections within the SOS formalism (with the concomitant slow convergence mentioned previously). In practice, however, TD-DFT<sup>20</sup> with quadratic response<sup>21-23</sup> has been used extensively to compute TPA cross sections for large molecules because of its relatively low computational cost and, within known limitations, good accuracy.<sup>12,24-31</sup> In standard response theory, each excited state within the frequency range of interest has to be addressed separately to obtain an absorption spectrum. This complicates the problem in large molecules, as they have high excited-state densities. Damped response

theory has been formulated to compute TPA spectra in such cases.<sup>32</sup> Until recently, higher-level *ab initio* methods were limited to small molecules. A benchmark in 2006 evaluated the effect of electron correlation in TPA computation (for small molecules) using a hierarchy of coupled cluster (CC) models ranging from the CC singles (CCS) to the approximate CC singles, doubles, and triples (CC3) models<sup>33,34</sup> and the results were compared to those from TD-DFT.<sup>35</sup> Both Pople- and Dunning-style basis sets were tested. The results highlight the importance of diffuse basis functions especially in larger molecules with the Dunning series performing significantly better than Pople basis sets. These early results showed that TPA cross sections computed via TD-DFT with the CAM-B3LYP functional<sup>36</sup> (and a modest basis set) were in good agreement with the CC3 results (to within 15%). Another study<sup>14</sup> on the yellow fluorescent protein (YFP)<sup>37</sup> showed that TPA results of CAM-B3LYP are in qualitative agreement with those from resolution-of-identity (RI) CC2.<sup>38,39</sup> A recent benchmark compared the TPA cross sections of fluorescent protein (FP) chromophores as computed by TD-DFT for four functionals (B3LYP,<sup>40</sup> PBE0,<sup>41,42</sup> CAM-B3LYP and LC-BLYP<sup>43-46</sup>) and full CC2 to averaged experimental values (see Chapter 2).<sup>13,47</sup> Results showed that the B3LYP<sup>40</sup> functional with the modest 6-31+G(d,p) basis set can be used to semi-quantitatively compare the TPA for the lowest energy excitation of fluorescent protein chromophores. Generally in the response theory formulation, time-dependent perturbation theory is applied to an approximate state and properties are derived from the residues of the relevant response functions (*vide supra*). An alternative formulation employing the expectation value approach has been developed<sup>48</sup> for computing TPA with the equation-of-motion coupled-cluster singles and doubles (EOM-CCSD) approach.<sup>49,50</sup> The method utilizes the RI and Cholesky Decomposition (CD) techniques (For further information regarding these approximations, please see Ref. 48 and the references therein). In this approach, expressions for exact states are first derived and then approximate wave functions are used for evaluation of matrix elements. Other methods that have been applied to large biological molecules, e.g., proteins, are reviewed in Section 1.3.1.

A number of the computational approaches developed for TPA evaluation are

now available in distributed *ab initio* electronic structure software. The DALTON<sup>51</sup> program has flexibility in TPA computation by having both TD-DFT and the CC hierarchy of methods with quadratic response available. The GAMESS-US<sup>52</sup> program can also be utilized for TPA calculations based on TD-DFT. Other software includes Q-Chem with the recent TPA implementation for the algebraic diagrammatic construction (ADC) method<sup>53,54</sup> and EOM-CCSD.<sup>48</sup> The implementation of RI-CC2 in Turbomole<sup>55</sup> (not yet in the latest publicly distributed version, i.e., Version 7.0) also opens the field for the computation of TPA for relatively large biological systems.<sup>37</sup>

## 1.3 Applications

### 1.3.1 Fluorescent Proteins

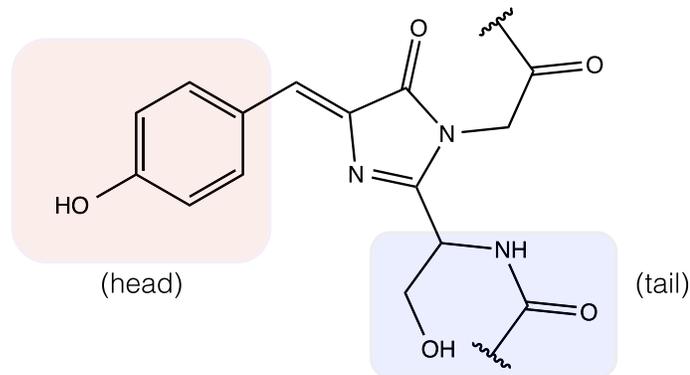
Fluorescent proteins (FPs) are a family of homologues of the *Aequorea victoria* green fluorescent protein (*avGFP*) that was discovered in the 1960s.<sup>56</sup> Later, the cloning of the GFP gene<sup>57</sup> and its expression in other organisms while maintaining its fluorescent properties<sup>58,59</sup> paved the avenue to a new era of microscopy and cell biology. FPs are capable of forming a fluorescent chromophore by a post-translational modification of three amino acids.<sup>60,61</sup> Changing these precursory amino acids yields different chromophores (see Fig 1.1). The chromophore structure provides a coarse tuning of the colour of the fluorescence associated with a given FP. Fine tuning is provided by the rest of the protein residues which form a barrel-shape encapsulating the chromophore while changes in the close-by residues of the chromophore can also strongly influence its spectral properties, as further discussed below. Among the full palette of FPs that is now available spanning red to blue FPs,<sup>62,63</sup> red-shifted FPs have been of great interest due to their lower cell-toxicity and better tissue penetration for light absorption and emission.<sup>64</sup> For more information about FPs, the interested reader is directed to any of the available reviews (see for instance the book chapter by Campbell and Davidson<sup>8</sup> and the references therein). The majority of photophysical studies on FPs focused on their OPA with a recent growing interest in TPA properties. Both OPA and TPA properties of FPs have been reviewed by Nifosi and Tozzini.<sup>62</sup>

Experimentally, there is a known general difficulty in measuring absolute TPA cross sections.<sup>65</sup> Measurements in biological systems like FPs are more challenging and, therefore, the reported measurements of TPA cross sections for a single FP may differ by up to two orders of magnitude, e.g., the reported measurements of the low-energy peak of the enhanced GFP (EGFP)<sup>66</sup> vary from 1.5 GM<sup>67</sup> to 180 GM.<sup>68</sup> In 2011, Drobizhev et al. comprehensively cited the available experimental measurements of TPA cross sections of FPs and discussed the sources of experimental error that could lead to such discrepancies.<sup>4</sup> Nevertheless, while there are discrepancies in the quantitative determination of the TPA cross sections, the experimentally-determined TPA spectra of FPs exhibit common qualitative features: (i) The TPA spectrum generally shows two regions of absorption; one at (nearly) double the wavelength of the OPA peak and an additional (usually stronger) absorption band at a shorter wavelength. (ii) The long-wavelength (low-energy) TPA peak is blue-shifted with respect to the corresponding OPA peak in the FPs with anionic chromophores.<sup>4</sup> In the rest of this section, we discuss the levels of complexity and the various computational approaches and benchmarks associated with TPA studies of FPs highlighting how computations improved our understanding of TPA in FPs and some of the unanswered questions.

### **Levels of Complexity and Computational Approaches**

In general, three levels of complexity influence the TPA of a given FP: (i) The intrinsic nature of the chromophore which is directly related to the extent of  $\pi$ -conjugation and the involved transition and permanent dipole moments (see below).<sup>12,13</sup> (ii) The effect of the protein shell on the geometry of the chromophore.<sup>13,69</sup> This includes consideration of both the change of chromophore conformation due to nearby residues (steric hindrance) and the conformational flexibility permitted by the hydrogen-bond network of interactions (including with water in the protein cavity). (iii) The electric field of the protein surrounding the chromophore, which can greatly influence its TPA.<sup>4,69</sup>

The first level of complexity is addressed by studying the isolated chromophore.



**Figure 1.1:** *GFP chromophore highlighting the head and tail positions that vary between FP chromophores.*

The GFP chromophore is shown in Figure 1.1, where we highlight the two sites that vary across FP chromophores. The “head” arises from the R-group of the amino acid Tyr66 which is the second amino acid in the tripeptide that matures into the chromophore. Point mutation of Tyr66 changes the head of the resulting chromophore. The “tail” of the chromophore arises from the amino acid Ser65. An additional step occurs in the maturation of the red FP (RFP) family resulting in an acylimine moiety which is thought to increase the  $\pi$ -conjugation in these systems.<sup>70</sup> Further cyclization in the tail region can also take place in some chromophores, such as in mKo<sup>71</sup> and mOrange.<sup>72,73</sup> When studying the bare chromophore, the connections to the protein are broken and capped with hydrogens or methyl groups. In proteins with no additional maturation step (i.e., GFP-like), the tail is usually removed and the imidazolinone ring is capped. Methyl capping of the GFP chromophore yields the thoroughly studied HBDI model.<sup>62</sup> Salem and Brown showed (see Chapter 2) that methyl-capping of FP chromophores with modified tails, e.g., RFP or mBlueberry,<sup>74</sup> can be necessary to obtain an optimized structure resembling the one found in the protein (see Chapter 4).<sup>13</sup> Even when only the isolated chromophore is considered, its size renders TD-DFT the method of choice for the majority of TPA studies (see Section 1.2.2). For TD-DFT studies of FP chromophores, it has been shown that B3LYP and PBE0 functionals yield  $S_0$  to  $S_1$  excitation energies that are very close to the experimental values in the FP while both energies and TPA cross sections are

systematically larger with CAM-B3LYP and LC-BLYP (see Chapter 2).<sup>13</sup> TPA cross sections computed with the four functionals (with the 6-31+G(d,p) basis set) were significantly lower than the corresponding CC2 values using the same basis set.<sup>13</sup> The most sophisticated *ab initio* methods considered at this level of complexity, i.e., for isolated chromophores, were RICC2,<sup>14</sup> full CC2<sup>48</sup> and EOM-CCSD.<sup>13</sup> At this level of complexity, the absolute values of the TPA cross sections for a given protein determined experimentally can not be reproduced quantitatively by any available computational method. Therefore, computational studies on the isolated chromophores are useful for initial screening of modified chromophores (i.e., to rank their intrinsic TPA cross section) or to study qualitatively the TPA photophysics for a given chromophore.

To address the second level of complexity, the chromophore needs to be studied in its native conformation dictated by the surrounding protein shell. This can be readily accomplished if the experimental three-dimensional structure is available for the studied FP, i.e., a PDB file is available. To go beyond the single-structure static picture, molecular dynamics (MD) simulations can be utilized to sample accessible chromophore conformations in its protein environment. To date, only one study systematically looked at the variations of TPA with chromophore conformations (see Chapter 3).<sup>75</sup> Only a small number of studies have accounted for the third level of complexity by including the electronic (non-mechanical) effect of the protein environment on its TPA. The polarizable embedding<sup>76</sup> (PE)-DFT/MM scheme allows for mutual polarization between the chromophore in its ground state and the protein (ground-state polarizable embedding). More importantly, PE-DFT accounts for the response of the protein electronic degrees of freedom to the excitation of the chromophore through a response formalism.<sup>69</sup> The two available studies that utilized PE-DFT are discussed in the following section.

### **Understanding TPA of FPs via Computational Studies**

Although studying a bare chromophore is the simplest approximation, it provides a very good starting point for understanding the TPA properties of a given protein.

Through this first level of complexity, Nifosi and Luo explained the origin of the high-energy absorption peak in the TPA spectrum that has no corresponding OPA peak.<sup>27,77</sup> They showed that this peak can be attributed to excitations to higher electronic states ( $S_0$  to  $S_n$ ). Further, they used TD-DFT at the B3LYP/6-31G+(d) level of theory in the gas phase, to compute TPA for a series of FP chromophore models and predicted the presence of such high-energy peaks in other chromophores as well.<sup>12</sup> These predicted absorptions were later validated via experiments.<sup>4</sup> Computations at this first level of complexity also helped to understand the blue shift of the low-energy TPA peak in anionic FP chromophores. It was first suggested that this shift is due to a bright TPA state that is dark under OPA.<sup>67</sup> No computation, however, found evidence for such a state.<sup>78,79</sup> Kamarchik and Krylov computed the non-Condon effects in the TPA spectrum of an anionic HBDI model (see Fig 1.1 and the related discussion). Their results showed that an increase in the transition moment associated with certain vibrational modes of the  $S_1$  electronic state causes a preferential vibronic transition. Ai et al. reached a similar conclusion via a CASSCF computation.<sup>80</sup>

More recently, Salem and Brown screened 22 hypothetical chromophores derived from non-canonical amino acids (ncAAs) for their  $S_0$  to  $S_1$  TPA cross sections (see Chapter 3).<sup>75</sup> A chromophore with a nitro group (that replaces the hydroxyl group in Fig 1.1) exhibited an exceptionally large TPA cross section relative to other ncAA-derived and natural chromophores as computed at the B3LYP/6-31+G(d,p) level of theory in PCM. Using the 2LM (see Sec 1.2.1), they attributed this large intrinsic TPA to both a large transition dipole moment ( $\mu_{01}$ ) and a large difference between permanent dipoles of the excited and ground states ( $\Delta\mu$ ). They accounted for the conformational dependence of the TPA (the second level of complexity) through a MD simulation of a modified EGFP. Results showed that the TPA dependence on conformation is directly related to the change in  $\Delta\mu$  in the direction of the C-C=C bond (i.e., the bridge between the chromophore rings) in the plane of the chromophore.

Chakarbarti and Ruud showed that systems with supermolecular  $\pi$ -stacking have enhanced TPA cross sections associated with intermolecular charge-transfer (ICT)

transitions.<sup>81,82</sup> Interestingly, YFP, a mutant of GFP, has  $\pi$ -stacking between the chromophore and a close-by Tyr residue. Beerepoot et al. studied the TPA of the YFP  $\pi$ -stacking systems (chromophore and close-by residue) via RI-CC2/aug-cc-pVDZ and CAMB3LYP with both aug-cc-pVDZ and a smaller 6-31+G(d) basis set.<sup>26</sup> Their results highlight ICT transitions with significantly enhanced TPA. They further determined the dependence of (part of) the TPA spectrum on the separation between the  $\pi$ -aromatic systems (across which ICT occurs). This work partially accounts for the third level of complexity; that is, the local environment of the chromophore. In a very recent study, Drobizhev et al. evaluated local electric fields in a series of 26 GFP homologues that share the same anionic structure. Their study elegantly compared experimental measurements of OPA and TPA spectra and the deduced total and long-range fields to findings from MD simulations to justify their method of measurements.<sup>83</sup>

To date, only two QM/MM studies fully accounted for the third level of complexity mentioned in Section 1.3.1. The first study by Steindal et al. (at the PE-CAM-B3LYP/6-31+G(d) level of theory) showed that embedding the GFP chromophore in the FP matrix significantly enhances its TPA as compared to the isolated chromophore model.<sup>30</sup> On the technical side, they mention the importance of including PE to reproduce experimental spectra. Although they theoretically accounted for (nearly) all the factors that should affect the TPA of GFP, they failed to reproduce quantitatively the experimentally measured intensities of the TPA spectrum. The need to include PE to correctly account for the protein matrix and the TPA enhancement it produces are confirmed in the second study by List et al. that examined DsRED, a red FP, also at the PE-CAM-B3LYP/6-31+G(d) level of theory. Further, List et al. studied the physical mechanisms leading to the enhancement of TPA cross sections by the electrostatic effects of the protein.<sup>69</sup> To do so, they determined the TPA cross sections for the isolated chromophore at the geometry optimized in vacuum (17 GM) and at its conformation in the native protein (47 GM). By comparing the results of these two computations to that of the full protein via PE (106 GM), they distinguished the effect at different levels of complexity. To account for

the importance of PE, they computed the TPA and other parameters with the PE turned off (which they refer to as NPE and is the same as electrostatic embedding; 51 GM) and with PE frozen at the ground state value only (FPE; 59 GM). Through a plane-projected 2LM that accounts for the angle between dipoles, they noticed that the increased TPA cross section is related to an enlarged  $\Delta\mu$  (the difference between permanent dipoles of the ground and excited states) and its closer alignment with  $\mu_{01}$  (the transition dipole moment). According to their results, the ground state dipole is actually larger than the excited state one, so increasing the former and decreasing the latter increases  $\Delta\mu$ . Computation of the ground and excited state permanent dipoles using different levels of theory showed that the ground state polarization (FPE) causes enhancement of both permanent dipoles with the ground state being slightly more enhanced. Further inclusion of full PE cancels the enhancement of the excited state dipole which increases  $\Delta\mu$  and also decreases the angle between  $\Delta\mu$  and  $\mu_{01}$  leading to the TPA enhancement. Evaluation of the individual contribution of each amino acid showed that the largest contribution to TPA enhancement comes from the positively charged Lys163 which is also mainly responsible of the blue shift of the OPA spectrum. They finally suggested a mutation that should improve the TPA intensity and at the same time red-shift the OPA spectrum.

Although the absolute TPA cross sections computed by List et al.<sup>69</sup> for the whole protein via PE (106 GM) is in good agreement with the experimental result (96 GM),<sup>84</sup> the method clearly needs more validation before there is confidence that experimental TPA cross sections could be computationally reproduced. Including effective external field effects (due to the external electromagnetic field) in a follow-up study nullified the enhancement due to the protein environment and caused the computed cross section (30 GM) to fall even below that of the isolated chromophore (at its native protein conformation).<sup>85</sup> The conclusions of List et al. regarding the direction of  $\Delta\mu$  mediation by the protein<sup>69</sup> contradict those by Drobizhev et al.<sup>84</sup> As pointed out by Beerepoot et al.,<sup>16</sup> there is a need to evaluate the inherent error associated with CAM-B3LYP through a benchmark against a higher-level, more accurate method. More experiments need to be undertaken to test the tuning suggestions

presented by some of the computational studies mentioned above. Finally, FPs for the promising chromophores made from non-canonical amino acids (see Chapter 3)<sup>75</sup> need to be experimentally synthesized and tested.

### 1.3.2 Nucleic Acid Bases

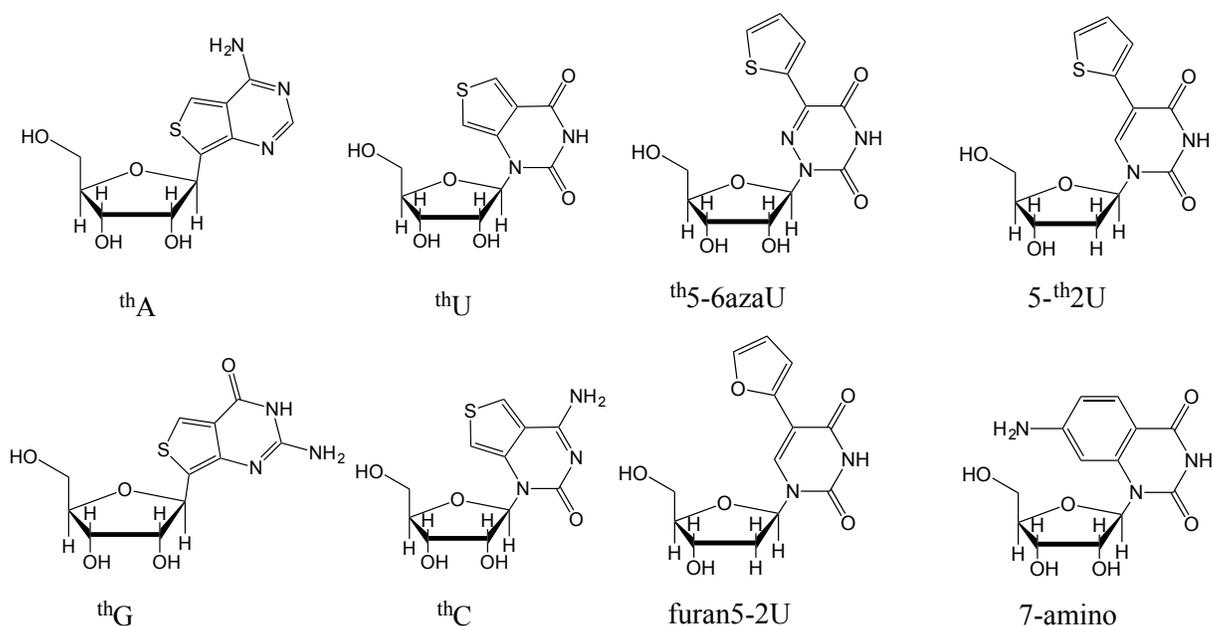
The basic building blocks of the nucleic acids DNA and RNA are the five nitrogenous bases: adenine, cytosine, guanine, thymine, and uracil. The ultra short excited state lifetimes of these bases grant them photostability which is crucial for the protection of the nucleic acid to photo-initiated processes. As a result of their non-emissive nature, the natural nucleobases cannot be utilized as fluorescent probes. Modifications to the structure of the nucleobases can yield molecules with different photophysical characteristics such as enhancement of their fluorescence. In the design of fluorescent base analogues, it is desirable that analogues exhibit sensitivity to their local environment while having minimal impact on the overall nucleic acid structure. The utilization of these “isomorphic” analogues is advantageous as they offer minimal perturbation to the helix, whereas bulky dyes can disrupt interactions necessary for nucleic acid functionality. Experiments utilizing fluorescent DNA/RNA nucleotide analogues have provided an effective way to study conformational dynamics of nucleic acids over a range of timescales.<sup>86–89</sup> For example, studies in the millisecond to second timescale provide information about protein-DNA interactions while fluorescence anisotropy can provide structural characteristics on the picosecond to nanosecond timescale.<sup>90</sup> Fluorescence microscopy at the single molecule level has emerged as a powerful technique to investigate conformational dynamics of DNA.<sup>91</sup> Challenges such as photobleaching arise with excitation in the UV range for many fluorescent base analogues. Investigating the photophysical properties and understanding how they are related to their structure is crucial for the design of more effective analogues. We refer the reader to recent, comprehensive discussions of modified nucleobases by Sinkeldam et al.<sup>92</sup> and Matsika.<sup>93</sup> The most common challenges in the design of fluorescent nucleobase analogues for biological applications are the following (and these echo the general challenges outlined in section 1.1): (i) The absorption cross sections

of nucleobase analogues is much lower compared to the extrinsic dyes and hence their brightness is insufficient for accurate imaging. (ii) Most analogues absorb in the UV region and can be prone to photobleaching. (iii) The penetration of light in biological tissues is poor at these wavelengths due to scattering processes. An alternative approach is to use multi-photon excitation in the near IR range. This enables the analogues to be utilized as probes for applications requiring deep tissue penetration and can overcome the problems arising from photobleaching.

While several experiments have investigated the TPA properties of nucleobase analogues, corresponding computational studies are far more limited; to our knowledge, there is only one.<sup>94</sup> Nearly a decade ago, the TPA cross section of the guanine analogue 6-methylisoxanthopterin (6-MI) and the adenine analogue 4-amino-6-methyl- 8-(2'-deoxy- $\beta$ -D-ribofuranosyl)-7(8H)-pteridone (6-MAP) were measured to be 0.8 GM and 3.4 GM, respectively.<sup>90,95</sup> Recently, Lane and Magennis determined experimentally the TPA cross section of 2-aminopurine (2-AP), as well as tricyclic cytosine (tC), as 0.2 and 1.5 GM, respectively.<sup>96</sup> Although these analogues demonstrate resistance to photobleaching, their brightness is not sufficient for single molecule detection. Recently, a set of isomorphous analogues synthesized by Tor et al. have been investigated for two-photon induced fluorescence.<sup>97</sup> The structures of these modified nucleosides are depicted in Figure 1.2.

### **Computational Investigations of Nucleoside analogues**

As with fluorescent proteins, we can investigate the properties of nucleic acid analogues in varying levels of complexity. The simplest approach (i.e., first level of complexity) is to determine the excitation properties of the isolated nucleobase which is the chromophoric moiety in the molecule. We can build upon that by including the ribose or deoxyribose (nucleoside) and, furthermore, we can investigate the analogues as a nucleotide or sequence of nucleotides (oligonucleotide). Of course in all of these cases, one would ideally want to include the effects of solvent, although often this can be accomplished through the use of polarizable continuum models (PCM). In some cases, the first level of complexity is sufficient for predicting the desired properties,



**Figure 1.2:** Structures of the modified nucleosides as designed by Lane et al.<sup>97</sup>  
 $^{th}A$ :thieno[3,4-*d*]adenosine,  $^{th}U$ :thieno[3,4-*d*]uridine,  $^{th}G$ :thieno[3,4-*d*]guanosine,  
 $^{th}C$ :thieno[3,4-*d*]cytidine,  $^{th}5-6azaU$ :5-(thiophen-2-yl)-6-aza-uridine,  $5-^{th}2U$ :5-(thiophen-2-yl)-2'-deoxyuridine,  $^{furan5-2U}$ :5-(furan-2-yl)-2'-deoxyuridine,  $7-amino$ :7-amino-1-ribose-quinazoline-2,4(1*H*,3*H*)-dione.

and inclusion of the sugar moiety (and phosphate group) does not alter the findings. An extensive analysis of the one-photon excitation properties of the thiophene analogues ( $^{th}A$ ,  $^{th}U$ ,  $^{th}G$  and  $^{th}C$ ) by Gedik and Brown<sup>98</sup> demonstrated this effect. For TPA, it is expected (yet remains to be confirmed) that the sugar group should have a more drastic effect on the cross sections since the orientation/geometry of the ribose will alter the permanent dipole moments (ground and excited state) of the molecule thus affecting  $\Delta\mu$ . Another factor that needs to be considered in the determination of TPA is the possible role of tautomerization. For example, it has been established experimentally that the tC analogue exists as a mixture of amino and imino tautomers.<sup>96</sup> In order to obtain accurate theoretical data, both tautomers would need to be examined.

To our knowledge, the only theoretical investigation of the two-photon excitation properties of nucleic acid analogues was performed by Samanta and Pati.<sup>94</sup> In this

study, the TPA cross sections of the set of analogues depicted in Figure 1.2 were determined at the B3LYP/6-31++G(d,p) level of theory. The computations were performed both in the gas phase and in solvent (implicit solvation with PCM). The OPA parameters have been determined by Samanta et al. in a previous study.<sup>99</sup> A summary of their findings is presented in Table 1.1.

**Table 1.1:** *Two photon absorption wavelengths and cross sections in GM of the modified uridine analogues. Experimental findings from Lane et al.,<sup>97</sup> and computations at the B3LYP/6-31+G(d,p) level of theory in the gas phase and PCM water, by Samanta and Pati.<sup>94</sup> NR = Not Reported.*

Analogue	Gas		Water		Experimental	
	E(nm)	$\sigma^{TPA}$	E(nm)	$\sigma^{TPA}$	E(nm)	$\sigma^{TPA}$
<i>th</i> A	681	7.45	711	3.98	NR	NR
<i>th</i> C	678	4.78	656	11.32	NR	NR
<i>th</i> G	642	5.54	714	16.90	NR	NR
<i>th</i> U	595	3.74	618	12.82	690	0.17
<i>th</i> 5-6azaU	713	7.08	751	31.30	690	3.8
					740	0.81
5- <i>th</i> 2U	670	4.55	681	18.23	690	7.6
					740	0.33
furan5-2U	662	11.36	661	11.70	690	2.1
					740	0.18
7-amino	563	3.72	605	72.65	690	1.8
					740	0.34

The results from Table 1.1 reveal that small alterations in the structure of the nucleosides can greatly impact their photophysical properties. The analogues display a very large variation in their TPA cross sections ranging from approximately 4 – 73 GM (computational determination) or 0.17 – 7.6 GM (experimental measurement) in water. From the computational results, the thiophene and furan analogues display the lowest cross sections while the quinazoline analogue shows the highest  $\sigma^{TPA}$  in water. This trend is not observed in the experimental findings. Both experiment and theory ascertain the  $S_0$  to  $S_1$  transition as both OPA- and TPA-active for all analogues. It is evident that more theoretical studies need to be performed: (i) to

test other functionals and/or higher level *ab initio* methods; (ii) to elucidate the effect of small structural changes on the TPA cross sections; and (iii) to examine the effects of conformational flexibility, e.g., the orientation of the sugar group mentioned previously. Further investigation of two-photon absorption of modified nucleobases by computational approaches is crucial for shedding light on the structure-photophysics interplay, which will ultimately aid in the design of better analogues.

## 1.4 Conclusions

In this chapter, we presented a brief discussion of the theory of TPA. With a focus on TPA in biological molecules, we reviewed the available computational methods and software packages. We then presented the applications of TPA utilizing fluorescent proteins and emissive nucleic acid base analogues. In each case, we gave a brief introduction to the topic followed by a review of the computational studies carried out on these two classes of molecules. We concluded each section with highlights of areas that require further computational investigation. What is clear is that there is a bright future for the computational study of TPA with applications to biological imaging.

## 1.5 Scope of the thesis

The objective of my thesis is to use computational methods to provide a rational basis to the experimental synthesis of FPs that are expected to have improved TPA cross sections. Previous efforts considered only the TPA for proteins made from canonical amino acids and OPA for proteins made from ncAAs. The experimental synthesis of FPs incorporating ncAAs is difficult and thus a priori computational predictions can save effort, time and money. My work was focused on isolated chromophores and only the steric effect of the protein environment was considered for interesting candidates.

In Chapter 2, we study the performance of TD-DFT in computing TPA cross sections (and other properties) of natural chromophores against the more expensive CC2 method and experimental data. We learn that TD-DFT generally underestimates the

TPA cross sections and the use of the B3LYP functional is rationalized for screening purposes. Chapter 3 describes an investigation of 22 GFP-based chromophores that are modified with side-chains from ncAAs. We use molecular dynamics to investigate the stability of a protein having a chromophore with a nitro substituent which has very promising TPA ability. In chapter 4, we investigate an equivalent set of RFP-based chromophores which differ in having an extra acylimine moiety. We see that the chromophore of the previously synthesized Gold FP can have a very large TPA cross section if it is engineered to have this extra acylimine extension. Finally in Chapter 5, I provide a summary of the thesis and mention possible avenues that motivate further potential research.

# Chapter 2

## Two-photon absorption in fluorescent protein chromophores: TD-DFT and CC2 results<sup>1</sup>

### 2.1 Introduction

Fluorescent proteins generally refer to the fluorescent homologues of the green fluorescent protein (GFP) of *Aequorea victoria*. After the discovery of GFP in the 1960s,<sup>56</sup> and later the cloning of the gene,<sup>57</sup> it has been shown that GFP can be expressed in other organisms and yet remain fluorescent.<sup>58,59</sup> Significant investment has been made into improving the photophysical properties of fluorescent proteins (FP) for their nonobtrusive use in real-time bioimaging. A great variety of fluorescent proteins have been synthesized and characterized, spanning a broad spectrum from blue to red fluorescent proteins.<sup>62,63</sup> The color of fluorescence is mainly controlled by the chromophore that is formed by three precursory residues.<sup>60,61</sup> By altering these amino acids, different maturation routes lead to different chromophore structures (many examples are mentioned in this chapter). In addition to the chromophore, the surrounding protein structure can also influence various photophysical properties of the protein.<sup>63</sup> The close-by residues to the chromophore can profoundly change the color of fluorescence, as seen in some variants of GFP that are shifted to produce

---

<sup>1</sup>This chapter is adapted (with minor corrections) from the published paper: M. Alaraby Salem and Alex Brown, *J. Chem. Theory Comput.*, 2014, **10**, 3260-3269.

yellow instead of green light albeit having the same chromophore as in GFP.<sup>37</sup>

Two-photon spectroscopy of cells expressing FPs has received less attention than one-photon spectroscopy but is recently gaining more interest.<sup>62</sup> It is less damaging to the biological system, as it involves the absorption of photons of longer wavelengths (and lower energy). Furthermore, the two-photon absorption (TPA) probability is proportional to the square of the incident light intensity. This provides better focus and less out-of-focus photobleaching, which makes two-photon microscopy preferred for imaging thick tissues with deeper penetration and reduced autofluorescence. Since one-photon absorption (OPA) and TPA follow different quantum-mechanical selection rules, they can be used in a complementary fashion.

In most of the studied FPs, excitation at nearly twice the one-photon wavelength peak results in effective TPA followed by fluorescence. The TPA peak is generally blue-shifted.<sup>100,101</sup> Recent computational studies on the chromophore of GFP showed that this blue-shift is mainly due to the vibronic coupling of states related to the bridge between the chromophore rings causing a non-Condon effect.<sup>29,80</sup> A detailed model of the shift has been developed and applied to a family of red FPs.<sup>102</sup> List et al. have carried out a detailed computational study at the CAM-B3LYP/6-31+G(d) level of theory for the isolated DsRED chromophore and the FP in order to ascertain the role of the protein environment on TPA.<sup>69</sup> Another interesting feature in the TPA profile of FPs is the presence of additional peak(s) in the shorter wavelength (higher energy) region where there is no corresponding significant one-photon absorption (OPA). This was first predicted by time-dependent density functional theory (TD-DFT) computations<sup>12,77</sup> and then experiments showed very strong TPA in this region.<sup>4</sup> Theoretical investigations reveal that excitations to higher excited states of the chromophores yield these bands. These transitions have been shown to be enhanced due to near-resonance conditions,<sup>12,77,103</sup> as further confirmed in this study. An early application for this high-energy band has been introduced using the red fluorescent protein, DsRed.<sup>104</sup> A newer technique was presented employing the simultaneous stimulation of blue and red FPs at the wavelength corresponding to the low-energy absorption and high-energy absorption of each, respectively.<sup>105</sup>

Measuring the probability of TPA (i.e., the TPA cross section) is technically demanding. Several parameters that need experimental evaluation produce experimental errors that contribute to the uncertainty of the absolute cross section. Recently, a thorough measurement of TPA for many of the available FPs was conducted within a wide spectral range.<sup>4</sup> The data shows that the bands that overlap between OPA and TPA profiles do not necessarily show equivalent intensities. The spectral window of these measurements could show the strong TPA in the high-energy regions, but was not enough to fully resolve the spectral peaks in many cases; that is, truncating the spectra before reaching the energy corresponding to the maximal TPA cross section.

Theoretical studies of the TPA properties of fluorescent proteins are likewise challenging. A combined quantum mechanics/molecular mechanics (QM/MM) approach was used in a recent study to reproduce the OPA and TPA of GFP.<sup>30</sup> The computation succeeded in reproducing most of the experimental structural features of the TPA spectrum, but failed in obtaining the absolute TPA cross sections. Other studies focused on isolated models of the chromophores.<sup>12,77,105</sup> Although the TPA of a chromophore can be largely influenced by the protein environment, studying the isolated chromophore was shown to be a useful starting point to understand the photophysical properties of the FP.<sup>12,77</sup> The study we present in this chapter evaluates the use of four different functionals to compute the TPA of seven FP chromophores in the TD-DFT framework. We compare TPA cross sections computed for different models at various computational levels to experimental data, to CC2 results, and to previous computational work. Several related studies on other systems have been conducted previously.<sup>14,26,106–108</sup> In a pioneering computational study of TPA, Paterson et al.<sup>106</sup> demonstrated that CAM-B3LYP provides the best agreement with benchmark CC3 results; interestingly, the same study shows B3LYP values appear superior when comparing with the TPA cross sections determined using more computationally accessible CC2 method. A study to evaluate TD-DFT methods for determining the TPA of oligophenylvinylenes discouraged the use of the long-range corrected functional (LC-BLYP) and favored B3LYP.<sup>107</sup> It also suggested the exclusion of diffuse functions from the basis sets. Another study on donor- $\pi$ -acceptor molecules favored

PBE0 to B3LYP for stilbene-based and fluorene-based compounds.<sup>108</sup> Our findings concerning the functional performance are generally in accordance with these previous studies. A recent study of the effect of the  $\pi - \pi$  stacking on the TPA of the YFP utilized RI-CC2 and TD-DFT with the CAM-B3LYP functional.<sup>26</sup> The results show very good qualitative agreement between both methods which led to the conclusion that TD-DFT is suitable for examining the TPA of the charge-transfer transitions in YFP. However, scaling had to be employed to match the shifted energies and cross sections. A benchmark of TD-DFT against RI-CC2 favored the use of CAM-B3LYP over other functionals, including B3LYP, for computing the TPA of aromatic compounds with strong charge-transfer character in many of the low-lying electronic states.<sup>14</sup> This may lead to the preference for CAM-B3LYP to B3LYP, since non-long-range corrected functionals are known to have problems with charge-transfer states. In this chapter, we further explore which functional, if any, is the best choice for efficient screening of TPA for FP chromophores in the framework of TD-DFT which is based on response theory.

## 2.2 Computational Methods

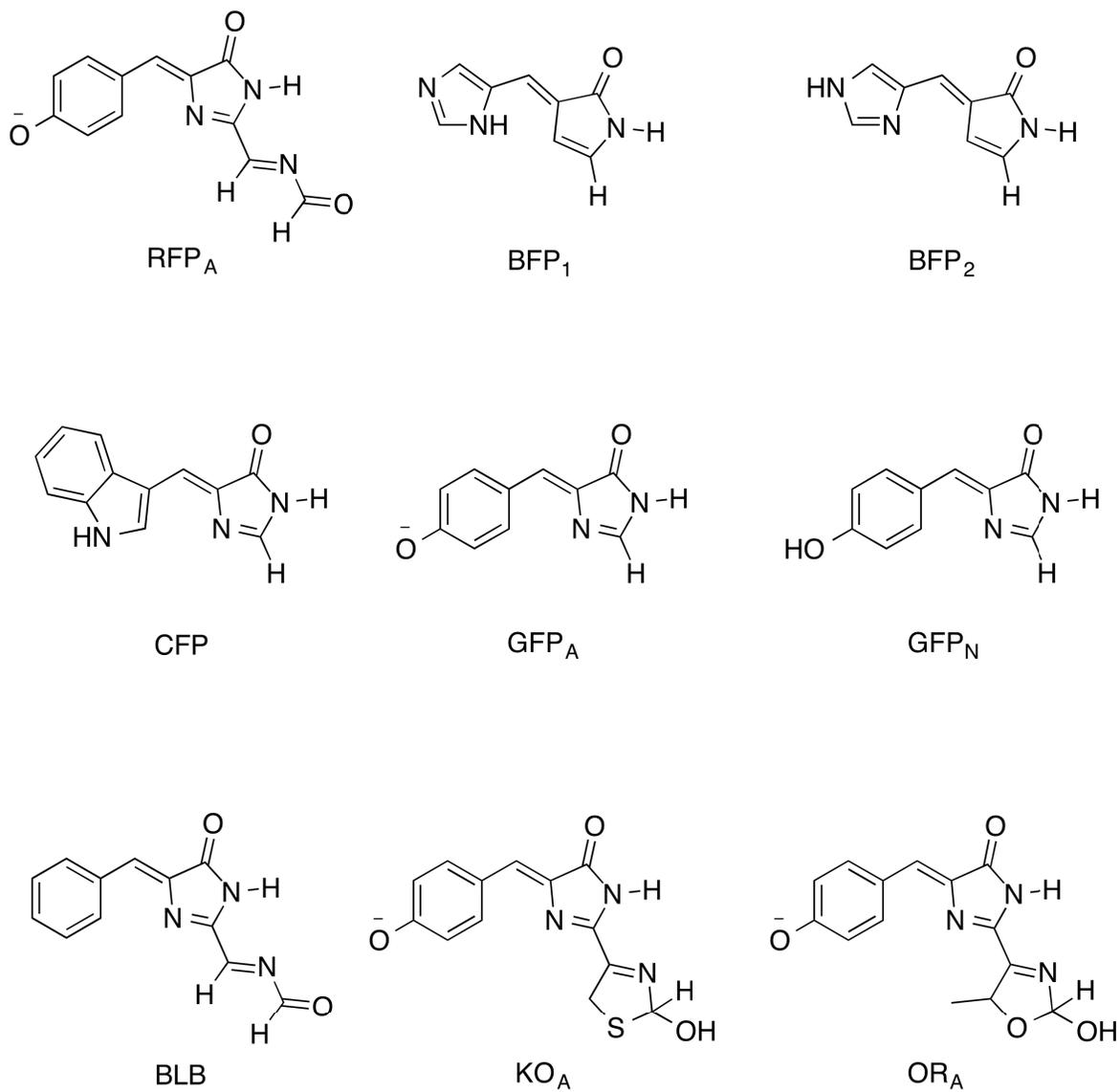
### 2.2.1 Chromophore Structures

The investigated chromophores are shown in Figure 2.1. The chromophore structures were obtained by breaking the connections to the protein backbone and capping with hydrogen atoms. In these models, the  $\pi$ -conjugated system important for TPA is preserved. The chromophores were chosen to represent proteins for which TPA cross sections have been consistently measured by Drobizhev et al.<sup>4</sup> to avoid the variability observed in the experimental measurements of a given FP. RFP represents the chromophore of red fluorescent proteins<sup>109</sup> such as DsRED.<sup>70</sup> Here, we include a carbonyl from the adjacent amino acid to account for the extra acylimine substituent that results from an additional dehydrogenation step during the RFP maturation.<sup>110</sup> This acylimine moiety is also included in BLB that represents the chromophore of mBlueberry1<sup>74</sup> and similar proteins. The blue fluorescent protein BFP<sup>111,112</sup> is rep-

resented by two models, BFP<sub>1</sub> and BFP<sub>2</sub>, to account for its two possible protonation states. CFP represents the chromophore of ECFP<sup>112</sup> and similar proteins where Tyr 66 is replaced by Trp.<sup>112</sup> Two models were used in the case of the green fluorescent protein (GFP)<sup>61</sup> to represent the possible protonation states of the chromophore, as both states are fluorescent in its wild-type.<sup>113</sup> KO<sub>A</sub> and OR<sub>A</sub> represent mKo<sup>71</sup> and mOrange<sup>72,73</sup> proteins, respectively.

## 2.2.2 Theory and Computations

Geometries of the reported model chromophores were optimized in the gas phase using the hybrid Perdew, Burke, and Ernzerhof exchange-correlation density functional PBE0<sup>41,42</sup> and the 6-31+G(d,p) basis set.<sup>114-118</sup> Coordinates for the optimized structures can be found in Table A1 and Table A2 in Appendix A. Excitation energies, OPA oscillator strengths and TPA cross sections were computed with time-dependent density functional theory<sup>20</sup> (TD-DFT) employing linear<sup>119</sup> and quadratic response theory.<sup>21-23</sup> Several functionals have been investigated for determining the photo-physical properties: B3LYP,<sup>40</sup> PBE0, the long range corrected B3LYP adopting the Coulomb-attenuating method (CAM-B3LYP)<sup>36</sup> and the long-range corrected version of the correlation functional of Becke, Lee, Yang, and Parr (LC-BLYP)<sup>43-46</sup> in both the gas phase and using the conductor-like polarizable continuum model (C-PCM, and herein referred to simply as PCM)<sup>120-123</sup> with parameters for water. Transition dipole moments between excited states were determined in the gas phase only, as the computation was very sensitive to cavitation parameters when PCM was used for solvation. The TPA for the smallest chromophores with Cs symmetry (BFP<sub>1</sub>, BFP<sub>2</sub>, GFP<sub>A</sub>, GFP<sub>N</sub>, and CFP) were also computed using the second order approximate coupled cluster singles and doubles model, CC2.<sup>38,39</sup> In all DFT, TD-DFT and CC2 computations, the 6-31G+(d,p) basis set using Cartesian harmonics was employed, except where noted. For one-photon absorption, the oscillator strength (i.e., the transition probability) of a transition from the ground state  $|0\rangle$  to an excited state  $|f\rangle$  is



**Figure 2.1:** Models for chromophores named after common fluorescent proteins they represent.

given by

$$\delta^{\text{OPA}} = \frac{2\omega_f}{3} \sum_{\alpha} |\langle 0 | \mu_{\alpha} | f \rangle|^2 \quad (2.1)$$

where  $\omega_f$  is the excitation energy from ground state to state  $f$  and  $\mu_{\alpha}$  is the dipole moment operator for a certain Cartesian coordinate  $\alpha$ . For two-photon absorption, the two-photon transition matrix has elements

$$S_{\alpha\beta} = \sum_n \left[ \frac{\langle 0 | \mu_{\alpha} | n \rangle \langle n | \mu_{\beta} | f \rangle}{\omega_n - \omega} + \frac{\langle 0 | \mu_{\beta} | n \rangle \langle n | \mu_{\alpha} | f \rangle}{\omega_n - \omega} \right] \quad (2.2)$$

where  $n$  ranges from the ground state 0 to the final excited state  $f$  and  $\omega$  is the photon energy ( $\omega = \frac{\omega_f}{2}$ ). Assuming that the polarization of incident light is linear, the transition moment for TPA is

$$\delta^{\text{TPA}} = \sum_{\alpha\beta} [S_{\alpha\alpha} S_{\beta\beta}^* + 2S_{\alpha\beta} S_{\alpha\beta}^*]. \quad (2.3)$$

The two-photon absorption cross section, which can be related to the experimental measurements, is then<sup>2</sup>

$$\sigma^{\text{TPA}} = \frac{4\pi^2 a_0^5 \alpha \omega^2}{15c \Gamma} \delta^{\text{TPA}} \quad (2.4)$$

where  $a_0$  is the Bohr radius,  $\alpha$  is the fine structure constant,  $c$  is the speed of light, and  $\Gamma$  is the lifetime broadening (derived from a Lorentzian function) which is assumed to be 0.1 eV for comparison to experiment as employed before.<sup>9</sup> All factors should be inserted in eq 2.4 in atomic units (au) and the result is converted from au to GM units.

To investigate for the presence of charge-transfer states, the overlap quantity  $\Lambda$  is reported.<sup>124</sup> It is a nonunique diagnostic value that measures the degree of overlap between virtual and occupied orbitals for a given excitation. The  $\Lambda$  values range from 0 to 1 where small values (<0.4) indicate the evidence for long-range excitations (Rydberg-type or charge-transfer).

CC2 computations, the computations of permanent dipole moments of excited states, and transition dipole moments between excited states were done using the

---

<sup>2</sup>Note: In the published paper, the reported TPA cross sections were too large by a factor of 4 when given in macroscopic units (GM) due to the use of excitation energy ( $\omega_f$ ) in eq 2.4 rather than photon energy ( $\frac{\omega_f}{2}$ )

DALTON<sup>51</sup> software package. All other computations were carried out with GAMESS<sup>52</sup> (May-2012 version), and all computations with C-PCM have utilized default parameters. MacMolPlt,<sup>125</sup> version 7.5, was used to generate plots of the primary orbitals involved in the excitations.

## 2.3 Results and Discussion

Comparing computed TPA cross sections to experimental measurements is a very challenging task. One reason is the great variation between experimental measurements of TPA cross sections for the same fluorescent protein.<sup>4</sup> For instance, reported cross sections for one peak of EGFP<sup>66</sup> vary from 1.5 GM<sup>67</sup> to 180 GM.<sup>68</sup> To relate our results to a consistent experimental reference, we limit our comparison to the results published by Drobizhev et al.<sup>4</sup> Another problem is that each chromophore is found in different proteins and the protein environment affects its photophysical behavior especially in the case of TPA cross section.<sup>102</sup> TPA depends on the difference between permanent dipole moments of a given excited state and the ground state in addition to its dependence on transition dipole moments, as can be seen in the sum-of-states expression (eq 2.2). Therefore, the electric field within the protein can strongly affect the TPA absorption.<sup>4</sup> In addition, changes in the chromophore conformation itself due to the surrounding protein residues strongly affect the TPA absorption, as discussed in section 2.3.1. Hence, we choose to compare our computed values to the average values of cross sections for the proteins that share the same chromophore when a range of values is available. The conclusions drawn from such comparison will only be qualitative or semiquantitative in nature, but the knowledge of maximum and minimum values of the data range included in an average shed light on possible effects of local environment (including electric field) and conformation on the TPA cross section. We also compare (some of) the TD-DFT results to those determined using the significantly more computationally demanding CC2 method. In the following sections, we first discuss the lowest-energy transition to  $S_1$  (section 2.3.1) and then the higher energy transitions to  $S_n$  (section 2.3.2), where  $n$  is the

excited state of interest which varies according to the system (i.e., chromophore) and the computational method. These should correspond to the band of the longest wavelength and other band(s) of shorter wavelengths in the experimental TPA spectrum of a given protein, respectively. For each chromophore, we computed 8 excitations per functional for TD-DFT; for CC2 only  $S_1$  is reported, as higher states were difficult to converge for many of the systems. In section 2.3.2, we firstly provide a thorough analysis of the data determined for the higher transitions for one chromophore, namely RFP. Further, we only report data for the most significant transitions and compare these to the available experimental data. Full data of the 8 excitations as determined using TD-DFT, including excitation energies,  $\Lambda$  values, OPA oscillator strengths, and TPA cross sections, for all chromophores is provided in Appendix A (Tables A3, A4 and A5).

### 2.3.1 Lowest-Energy Transition to $S_1$

The one-photon energies and TPA cross sections from this chapter and previous computational work by Nifosi and Luo<sup>12,77</sup> are shown in Table 2.1 together with the corresponding experimental measurements. In Table 2.1, the experimental data correspond to average values for the excitation energies and TPA cross sections of different proteins sharing the same chromophore. Throughout this paper, experimental energies are multiplied by a factor of two for ease of comparison to the computational data - no correction is made for the blue-shift observed in the TPA of many FPs.<sup>29,102</sup> The experimental data for specific proteins considered are grouped by chromophore in Table A6 in Appendix A. To obtain the average, proteins with colors significantly shifted from other proteins sharing the same chromophore were excluded: from the proteins sharing the GFP chromophore, the yellow homologues were omitted and from the proteins sharing the RFP<sub>A</sub> chromophore, we excluded tagBFP<sup>126</sup> that is blue-shifted. mKok, mOrange, and mBlueberry1 were the only proteins of their types in the experimental data set, so KO<sub>A</sub>, OR<sub>A</sub> and BLB are compared directly to the values for these proteins, respectively.

As shown in Table 2.1, different functionals under study yield values very close

**Table 2.1:** *Excitation energies and TPA cross sections for the transitions to  $S_1$  (bands of lower energy and longer wavelengths) as determined at the TD-DFT/6-31+(d,p) level of theory in PCM ( $H_2O$ ) and at the CC2/6-31+G(d) level of theory in the gas phase.<sup>a</sup>*

Model	Energy (eV)						Exp	Prev Calc
	B3LYP	PBE0	CAM	LC	CC2			
RFP <sub>A</sub>	2.28	2.31	2.32	2.29	2.26	2.26	2.36	
						(2.18-2.36)		
BFP <sub>1</sub>	3.4	3.46	3.55	3.59	3.78	3.31	3.59	
BFP <sub>2</sub>	3.59	3.66	3.79	3.83	4.09	3.31	3.62	
CFP	3.14	3.23	3.37	3.41	3.58	2.85	3.32	
GFP <sub>A</sub>	2.96	3	3.02	2.97	2.89	2.83	2.97	
						(2.68-2.87)		
GFP <sub>N</sub>	3.34	3.41	3.56	3.62	3.8	3.11	3.46	
						(3.06-3.21)		
KO <sub>A</sub>	2.45	2.5	2.59	2.58	-	2.38	-	
OR <sub>A</sub>	2.48	2.53	2.6	2.59	-	2.3	2.59	
BLB	2.68	2.77	3.04	3.19	3.24	3.05	-	
Model	TPA cross section (GM)						Exp	Prev Calc <sup>b</sup>
	B3LYP	PBE0	CAM	LC	CC2			
RFP <sub>A</sub>	1	1	2	3	-	56	5.7	
						(15-139)		
BFP <sub>1</sub>	1	1	2	3	11	12	0.3	
BFP <sub>2</sub>	1	1	2	3	2	12		
CFP	7	7	7	8	26	23	6.2	
GFP <sub>A</sub>	1	1	2	4	2	46	0.1	
						(12-85)		
GFP <sub>N</sub>	4	4	5	6	20	36	1.4	
						(15-56)		
KO <sub>A</sub>	7	7	8	7	-	41		
OR <sub>A</sub>	6	6	6	6	-	67	3.2	
BLB	23	24	25	23	-	4		

<sup>a</sup> Experimental values are averages for the values measured by Drobizhev et al.<sup>4</sup> in the cases where the model is represented by several proteins (range is given below each average when applicable) and previous calculations refer to those by Nifosi and Luo.<sup>12,77</sup> Experimental energies are doubled to match the OPA energies computed by the different methods. <sup>b</sup>Reported to the same number of decimal places as given by Nifosi and Luo.<sup>12,77</sup> Details of level of theory provided in main text.

to the energetics for the  $S_1$  transition - similar to what has been shown in previous OPA and TPA studies.<sup>12,77,127</sup> Our results indicate that the energetics are weakly functional dependent with CAM-B3LYP and LC-BLYP giving in most cases slightly higher energies than B3LYP and PBE0. Similar to what has been shown before,<sup>127</sup> comparison to the previous computations also indicate that energetics are not greatly affected by change of medium and basis, e.g., the excitation energies using PCM (H<sub>2</sub>O) are approximately 0.1 eV smaller than the gas-phase results, see Table A5. The energies for the B3LYP and PBE0 functionals are closer to the experimental values (with the exception of BLB). We see that the first model for BFP, namely BFP<sub>1</sub>, behaves better than BFP<sub>2</sub> in terms of reproducing the experimental energy. The CC2 energies for the neutral models are larger than those from any of the four functionals (even if one accounted for a small gas-phase to PCM shift). The size of the systems precluded using a larger basis than the modest 6-31+G(d,p) and, in fact, CC2 results were not obtained for the two largest chromophore models.

Previous calculations by Nifosi and Luo<sup>12,77</sup> used similar chromophore models but with methyl-group capping (c.f. Hydrogen capping in the present study). In the previous study, the B3LYP functional was used in the gas phase employing the following basis sets: 6-31G(d) for all chromophores and 6-31+G(d) for the anionic and some neutral ones. The values reported in Table 2.1 are for those determined with the diffuse basis functions; that is, 6-31+G(d), except for BFP models. These values deviate 0.19 eV (BFP<sub>1</sub>) to 0.01 eV (GFP<sub>A</sub>) from the present results at the B3LYP/6-31+G(d,p) in PCM (water) level of theory.

In general, the TPA cross sections are (slightly) more functional dependent, see Table 2.1. In most cases, CAM-B3LYP and LC-BLYP produce larger TPA cross sections than B3LYP and PBE0. Computations using CAM-B3LYP and LC-BLYP functionals are closer to the experimental measurements of the TPA cross sections with mean absolute difference (MAD) of 30 GM for both. This MAD is nearly the same as that for B3LYP and PBE0 (31 GM for both). In general, the TD-DFT computed values are lower than the experimental ones except for BLB. In this case, LC-BLYP and CAM-B3LYP still produce higher values than B3LYP and PBE0 sug-

gesting that these long-range corrected functionals simply yield higher values than their standard counterparts. This comparison is further investigated in section 2.3.2. Comparing the TD-DFT results using the B3LYP, PBE0, and CAM-B3LYP functionals in the gas-phase (see Table A5 in Appendix A) to the CC2 values for the five chromophores, one obtains MAD of approximately 9 GM for all. The CC2 computations were outside the experimental ranges for energetics but closer to experiment for TPA cross sections. Given that CC2 is computationally demanding compared to TD-DFT and deviates from the experimental data, we do not encourage implementing CC2 in a screening procedure. It can, however, be used for more in-depth analysis - which is beyond the scope of this chapter.

Comparison to the previous computational results leads to further insight into the parameters affecting the determination of TPA properties. The RFP chromophore was previously computed at the B3LYP/6-31+G(d) level of theory in the gas phase to have a cross section of 6 GM.<sup>12,77</sup> Our relevant value using B3LYP/6-31+G(d,p) in PCM (water) is 1 GM. Factors that might affect such change include the different basis set, different medium, different model (methyl vs H-capping) or molecular geometry. In order to investigate these factors, the TPA was determined for different models of the RFP chromophore, see Table 2.2. Using the same geometry (optimized at the PBE0/6-31+G(d,p) level of theory), we repeated the TPA computation in the gas phase at the B3LYP/6-31+G(d,p) and the B3LYP/6-31+G(d) levels of theory and the results were approximately 2 GM for both. Looking at the available gas phase values in Appendix A (see Tables A3, A4 and A5), one can draw the conclusion that they are generally larger than PCM values. The TPA cross section of 2 GM is approximately three times smaller than the previously calculated TPA cross section by Nifosi and Luo.<sup>12,77</sup> Methyl capping was then tried. The methyl-capped model was optimized at the PBE0/6-31+G(d,p) level of theory in the gas phase and the TPA was computed at the same level of theory as in the previous study (B3LYP/6-31+G(d) in the gas phase). Our result was in accordance with that of the previous work: 9 GM and 6 GM, respectively (the difference might be attributed to slightly different geometries and energies). We found that the carbonyl group in the model

**Table 2.2:** Comparison of different computations performed on the RFP chromophore using the B3LYP functional.

Medium	Basis	Capping	Carbonyl Moiety <sup>a</sup>	cross section (GM)
PCM	6-31+G(d,p)	H	coplanar	1
gas	6-31+G(d,p)	H	coplanar	2
gas	6-31+G(d)	H	coplanar	2
gas	6-31+G(d)	CH <sub>3</sub>	noncoplanar	9
PCM	6-31+G(d,p)	H	noncoplanar <sup>b</sup>	5

<sup>a</sup> The planar structures with hydrogen-capping and the nonplanar structure with methyl-capping are true minima optimized in the gas phase at the PBE0/6-31+G(d,p) level of theory. <sup>b</sup>Same conformation as the minimized molecule with methyl-capping, but the methyl groups were replaced with hydrogens and the bonds to the hydrogens were adjusted to be of lengths equal to the corresponding bonds in the planar model with hydrogen-capping.

with methyl capping was not in plane with the molecule when minimized, and this agrees with previous studies that show that the C-N-C=O dihedral angle is off-plane by approximately 30 degrees.<sup>12,70,77</sup> Retaining the same geometry that was found by minimizing the structure with methyl-capping, we replaced the methyl groups with H-atoms (coordinates are given in Table A2 in Appendix A) and recomputed the TPA cross sections. The result was a significant increase in the TPA cross section from 1 GM to 5 GM, suggesting the critical role of this out-of-plane carbonyl conformer in the RFP chromophore. Examining the  $S_{\alpha\beta}$  elements contributing to the TPA cross section, see eq 2.3, the major difference between the nonplanar and planar geometries is an increase in the  $S_{xx}$  element. Since the one-photon oscillator strengths for the planar and nonplanar geometries are comparable, i.e., 1.23 and 1.11, respectively, the increase in the TPA cross section can be attributed to an increase in the permanent dipole moment difference between the excited and ground states (in the  $x$ -direction along the long molecular axis) as the FP chromophore goes from the planar to nonplanar geometry. More generally, the results suggest that the chromophore geometry, perhaps influenced by the protein backbone, can strongly impact the TPA as illustrated before.<sup>69</sup>

The case of the BLB chromophore is even more interesting. This chromophore is similar to that of RFP, with the exception of the missing hydroxyl group on the benzene ring. The deletion of the OH-group (and adding a hydrogen) induced a huge enhancement in the TPA cross section of the planar conformer, where the TPA cross section determined using the B3LYP functional increased from 1 GM to 23 GM; the TPA cross section increased similarly for the other functionals. To investigate the role of geometry on the photophysical properties, the TPA cross section for BLB was recomputed after deleting the OH group from the out-of-plane RFP<sub>A</sub> model with H-capping (coordinates are given in Table A2 in Appendix A). The result was a noticeable decrease in TPA cross section to 6 GM which is very close to the computed cross section for the similar RFP chromophore model. Therefore, the out-of-plane rotation of the acylimine moiety does not always result in an increase in the TPA. The planar conformer could be achieved because of the use of hydrogen capping in the model. Of course, this is different from the usual connectivities to the chromophore in the protein that are essentially bulkier and force this out-of-plane rotation. Nevertheless, if it is possible to enforce a planar conformation for BLB through protein engineering, the results suggest the TPA may be strongly enhanced. The large difference between the computed values for the planar, or nonplanar, geometries and the experimentally measured values could be attributed to the strong geometry dependence of the TPA cross section for BLB. The discrepancy between the experimental value for mBlueberry1 and the computed values (whether the planar or the nonplanar models) is further highlighted in section 2.3.2 in the discussion of higher excited states for BLB.

### 2.3.2 High-Energy Transitions to $S_n$

The one-photon energies, oscillator strengths, overlap quantity  $\Lambda$  and TPA cross sections for the next lowest energy bright states for the RFP chromophore and all the other chromophores are given in Table 2.3 and Table 2.4, respectively. The quoted experimental excitation energies and TPA cross sections represent average values over several proteins containing the same chromophore except in the cases of BLB,

KO<sub>A</sub> and OR<sub>A</sub>, where only one instance of each FP chromophore is available (see Table A6). In many cases, the previously reported experimental values for the S<sub>n</sub> excitation did not represent a true spectroscopic peak, but rather the largest TPA cross section determined within the experimentally feasible measurement window in each case.<sup>4</sup> The upper-limit for this window (multiplied by a factor of two) ranges from approximately 4.5 eV for blue fluorescent proteins to approximately 3.5 eV for red fluorescent proteins. The computed data was examined using two criteria: (1) Excitations with zero TPA cross sections are not considered. (2) high-energy transitions that are 0.5 eV larger than the experimental upper-limit are excluded from the discussion. Although predicted to have very large cross sections, they might be of limited experimental utility due to the larger energy difference between the emitted photons and the absorbed ones causing the dissipated energy in the non-fluorescent route to be higher. In addition, fluorescent proteins in general would have high TPA cross sections in the range of 532 nm (equivalent to an energy of 4.66 eV, when multiplied by two) because of their tryptophan residues.<sup>128,129</sup>

The remaining excitations are grouped according to visual inspection of the primary orbitals involved in the excitation (see Table A7 to Table A14 in Appendix A). Orbitals are referenced relative to the HOMO (H) and the LUMO (L) using a subscript. Each group of excitations has four transitions (one for each functional) that involve orbitals of similar shapes, but they do not necessarily correspond to the same energy gaps. This is because computations with different functionals did not necessarily yield orbitals in the same order. The groups are ordered according to the relative B3LYP energies and the number of each state is given for reference. All the excitation groups for RFP are given in Table 2.3 and discussed in detail. For all the other chromophores, only data assumed to be contributing to the corresponding experimental S<sub>n</sub> transition is reported in Table 2.4 and a brief discussion is provided. A full list is available in Table A15 and Table A16 in Appendix A. The experimental spectra that are referred to in this analysis are found in Supplementary Figure 1 of the work of Drobizhev et al.<sup>4</sup>

**Table 2.3:** Data for higher excitations as computed at the TD-DFT/6-31+G(d,p) level of theory in PCM ( $H_2O$ ) for the RFP chromophore.

State <sup>a</sup>	Functional	Energy (eV)	$\Lambda$ Diagnostic	OPA OS	cross section (GM)	Transition
	Experimental <sup>b</sup>	3.47 (3.30-3.68)			197 (35-400)	
3	B3LYP	3.32	0.75	0.02	1	H <sub>-2</sub> to L
3	PBE0	3.47	0.67	0.02	1	H <sub>-2</sub> to L
6	CAM-B3LYP	4.13	0.74	0.09	67	H <sub>-1</sub> to L
6	LC-BLYP	4.45	0.72	0.01	311	H <sub>-1</sub> to L
5	B3LYP noncoplanar	3.78	0.57	0.09	16	H <sub>-2</sub> to L
6	CAM-B3LYP (gas)	4.18	0.73	0.25	150	H <sub>-1</sub> to L
6	B3LYP	3.61	0.41	0.01	3	H <sub>-4</sub> to L
6	PBE0	3.76	0.42	0.02	3	H <sub>-4</sub> to L
7	CAM-B3LYP	4.27	0.40	0.01	16	H <sub>-3</sub> to L
7	LC-BLYP	4.52	0.40	0.01	53	H <sub>-4</sub> to L
6	B3LYP noncoplanar	3.84	0.50	0.08	40	H <sub>-4</sub> to L
7	CAM-B3LYP (gas)	4.36	0.41	0.01	18	H <sub>-4</sub> to L
7	B3LYP	3.83	0.72	0.35	23	H to L <sub>+1</sub>
7	PBE0	3.91	0.71	0.35	19	H to L <sub>+1</sub>
4	CAM-B3LYP	3.96	0.73	0.34	23	H to L <sub>+1</sub>
4	LC-BLYP	4.00	0.71	0.44	283	H to L <sub>+1</sub>
3	B3LYP noncoplanar	3.44	0.71	0.12	18	H to L <sub>+1</sub>
5	CAM-B3LYP (gas)	3.99	0.74	0.11	56	H to L <sub>+1</sub>

<sup>a</sup> States are given in increasing energy order for the B3LYP functional and grouped with the corresponding PBE0, CAM-B3LYP and LC-BLYP results according to the nature of excitation. See the discussion in section 2.3.2. Experimental energies are doubled to match the OPA energies computed by the different methods. <sup>b</sup> The limits of the range of values contributing to the average are given between brackets.

## RFP

In Table 2.3, three of the higher bright excited states are grouped. There is a remarkable variation between functionals concerning the arrangement of orbitals and excited states. The first group of excited states includes the third excitation in the B3LYP and PBE0 computations which corresponds to a transition from H<sub>-2</sub> to L. As deduced from Table A7, this transition is mapped to a H<sub>-1</sub> to L transition in the CAM-B3LYP and LC-BLYP computations which corresponds to the sixth excited state. In this group, B3LYP and PBE0 predict very small cross sections in contrast to CAM-B3LYP and LC-BLYP. The corresponding TPA for the H-capped structure having the off-plane acylimine carbonyl is much larger as computed by B3LYP (16

**Table 2.4:** Data for higher excitations to  $S_n$  for the BFP, CFP, GFP,  $KO_A$ ,  $OR_A$  and BLB chromophores as determined at the TD-DFT/6-31+G(d,p) level of theory in PCM ( $H_2O$ ) solvent and the corresponding experimental measurements<sup>a</sup> by Drobizhev et al.<sup>4</sup>

State	Functional	Energy (eV)	$\Lambda$ Diagnostic	OPA OS	cross section (GM)	Transition
BFP <sub>1</sub>	Experimental	4.50 (4.49-4.51)			12 (10-16)	
3	B3LYP	4.42	0.69	0.05	12	H <sub>-2</sub> to L
3	PBE0	4.54	0.69	0.05	14	H <sub>-2</sub> to L
3	CAM-B3LYP	4.88	0.68	0.04	30	H <sub>-1</sub> to L
3	LC-BLYP	4.98	0.67	0.05	43	H <sub>-2</sub> to L
CFP	Experimental	4.51			21 (15-27)	
4	B3LYP	4.07	0.62	0.04	32	H <sub>-2</sub> to L
4	PBE0	4.22	0.61	0.02	37	H <sub>-2</sub> to L
4	CAM-B3LYP	4.71	0.59	0.00	48	H <sub>-2</sub> to L
4	LC-BLYP	4.92	0.56	0.02	20	H <sub>-2</sub> to L
GFP <sub>N</sub>	Experimental	4.43 (4.34-4.51)			31 (27-36)	
4	B3LYP	4.30	0.62	0.02	30	H <sub>-3</sub> to L
4	PBE0	4.44	0.62	0.02	33	H <sub>-3</sub> to L
4	CAM-B3LYP	4.82	0.66	0.02	60	H <sub>-4</sub> to L
4	LC-BLYP	4.93	0.65	0.03	71	H <sub>-4</sub> to L
GFP <sub>A</sub>	Experimental	4.30 (4.27-4.35)			13 (12-16)	
5	B3LYP	4.23	0.64	0.02	4	H <sub>-2</sub> to L
5	PBE0	4.36	0.70	0.01	6	H <sub>-2</sub> to L
5	CAM-B3LYP	4.76	0.76	0.01	24	H <sub>-2</sub> to L
6	LC-BLYP	4.90	0.73	0.01	47	H <sub>-2</sub> to L
KO <sub>A</sub>	Experimental	3.62			93	
4	B3LYP	3.62	0.67	0.17	19	H to L <sub>+1</sub>
4	PBE0	3.77	0.67	0.19	27	H to L <sub>+1</sub>
3	CAM-B3LYP	4.17	0.62	0.32	205	H to L <sub>+1</sub>
3	LC-BLYP	4.31	0.57	0.31	339	H to L <sub>+1</sub>
OR <sub>A</sub>	Experimental	3.88			200	
3	B3LYP	3.65	0.71	0.11	22	H to L <sub>+1</sub>
3	PBE0	3.80	0.70	0.13	30	H to L <sub>+1</sub>
3	CAM-B3LYP	4.21	0.63	0.30	189	H to L <sub>+1</sub>
3	LC-BLYP	4.35	0.59	0.29	320	H to L <sub>+1</sub>
BLB	Experimental	4.51			22	
5	B3LYP	3.80	0.60	0.02	58	H <sub>-4</sub> to L
5	PBE0	3.96	0.60	0.03	72	H <sub>-4</sub> to L
5	CAM-B3LYP	4.48	0.61	0.07	75	H <sub>-3</sub> to L
6	LC-BLYP	4.66	0.61	0.08	51	H <sub>-3</sub> to L

<sup>a</sup> Experimental values for BFP, CFP and GFP are averages for the experimental values measured by Drobizhev et al.<sup>4</sup> while the values for  $KO_A$ ,  $OR_A$  and BLB refer to one protein for each. Experimental energies are doubled to match the OPA energies computed by the different methods. <sup>b</sup> The limits of the range of values contributing to the average are given between brackets when applicable.

GM compared to 1 GM for the planar model), showing the dependence of this excitation on the geometry of the model in a way similar to, or even greater than, the  $S_1$  excitation. For the full grouped data for the nonplanar model, see Table A17. The second group of transitions have a small value for the  $\Lambda$  diagnostic indicating a long-range excitation. Such a transition is expected to be characterized better by CAM-B3LYP than the B3LYP functional.<sup>124</sup> The third group, however, shows a very strong absorption with all functionals and a relatively large  $\Lambda$  diagnostic indicating a valence-type transition. Examining the data suggests the presence of two peaks of very close energy with the higher energy peak having stronger (or similar) TPA cross section. This behavior is experimentally noticed in some of the proteins represented by the RFP chromophore. The spectrum of mTangerine,<sup>72</sup> for instance, has a small peak at 3.01 eV followed by a stronger rise at 3.57 eV. Another example is seen in the spectrum of tagRFP<sup>130</sup> that shows two close peaks around 3.65 eV and 3.27 eV of comparative strengths. Experimentally, the energy of the second band in the fluorescent proteins having the RFP chromophore averages around 3.47 eV. In all proteins included in this average, this corresponds to the maximum TPA cross section that was recorded in the high-energy region (according to Drobizhev et al.<sup>4</sup>) rather than being a true peak. In addition, this average covers TPA cross section values ranging from 35 GM to 400 GM. Thus, the quoted experimental data are only good for qualitative, or at best semiquantitative, analysis. Comparing this averaged experimental energy to the third group of transitions where all functionals predict a strong absorption, it is noticeable that B3LYP and PBE0 show more accuracy in predicting the energies than CAM-B3LYP and LC-BLYP - similar to what is noticed in the  $S_1$  analysis. Again, LC-BLYP yields the largest TPA cross section.

Factors that could contribute to the large cross sections were investigated. Transition dipole moments were computed with CAM-B3LYP as an exemplary functional in the gas phase and the transitions were mapped to the relevant PCM ones (see Table 2.3) via matching orbital shapes. As reported in Table 2.3, the OPA oscillator strengths for states 5,6 and 7 (for CAM-B3LYP in the gas phase) are very small compared to the large TPA cross sections. Values for contributing dipole moment

**Table 2.5:** *Nonzero dipole-moment elements contributing to the TPA cross section of the 5<sup>th</sup>, 6<sup>th</sup> and 7<sup>th</sup> excited states of the RFP chromophore determined at the CAM-B3LYP/6-31+G(d,p) level of theory in the gas phase*

Dipole Element	value (a.u.)
$ \langle 0 \mu 5\rangle ^2$	1.07
$ \langle 0 \mu 1\rangle ^2$	19.79
$ \langle 1 \mu 5\rangle ^2$	0.13
$ \langle 5 \mu 5\rangle - \langle 0 \mu 0\rangle ^2$	0.72
$ \langle 0 \mu 6\rangle ^2$	2.48
$ \langle 1 \mu 6\rangle ^2$	0.29
$ \langle 6 \mu 6\rangle - \langle 0 \mu 0\rangle ^2$	0.72
$ \langle 0 \mu 7\rangle ^2$	0.10
$ \langle 1 \mu 7\rangle ^2$	0.03
$ \langle 7 \mu 7\rangle - \langle 0 \mu 0\rangle ^2$	4.37

elements are shown in Table 2.5. The contributing transition dipole moments and the difference between permanent dipole moments of excited and ground states are collectively too weak to drive this strong absorption. For the CAM-B3LYP computations in the gas phase, energies of the involved transitions are comparable to twice of the first excited state energy of 2.477 eV. This causes the denominator of eq 2.2 to be small causing the amplification of these cross sections. These near-resonance conditions can be considered a key contributor to the large cross sections obtained confirming what has been shown before.<sup>12,77,103</sup>

### BFP

The values shown in Table 2.4 are for the BFP<sub>1</sub> model. The second excited state in all computations was dim, so there is no intermediate peak between the  $S_1$  and the significant  $S_n$  peak - which is similar to the experimental spectra. The largest TPA cross section belongs to the third excited state for all functionals. The computed energies using B3LYP and PBE0 (4.42 eV and 4.54 eV, respectively) match very well with the average experimental energy of 4.50 eV. On the other hand, CAM-B3LYP and LC-BLYP overestimate the energy. The cross sections determined via B3LYP and PBE0 (12 and 14 GM, respectively) match the experimental average (12 GM),

while those computed by CAM-B3LYP and LC-BLYP are overestimated (30 and 43, respectively). The experimental spectra of all the involved proteins show a very sharp rise at approximately 4.5 eV, but each spectrum is discontinued before showing the descending part of the curve. The sharp slope indicates that the energy of the true spectroscopic peak is probably very close to the reported energy, but the TPA cross section could be larger. The deviation of energies computed using CAM-B3LYP and LC-BLYP functionals from the experimental value is slightly larger in the  $S_n$  states than the behavior observed in the  $S_1$  state.

### CFP

The second lowest-energy transition with nonzero TPA in the case of CFP is given by the second state for B3LYP and PBE0 and the third state for CAM-B3LYP and LC-BLYP as shown in Table A15 in Appendix A. The state is relatively dim in the case of B3LYP and PBE0 (2 GM for both), weakly bright for CAM-B3LYP (6 GM) and bright in case of LC-BLYP (31 GM). These transitions have a small value for the  $\Lambda$  diagnostic (0.42 for both B3LYP and PBE0, 0.46 for CAM-B3LYP and 0.55 for LC-BLYP), characterizing a long-range excitation. The fourth transition computed using all functionals is shown in Table 2.4. Unlike the preceding transition, the fourth state is a valence-type excitation and is strongly bright as computed by all the functionals, so it is considered the main  $S_n$  transition. Experimentally, the spectra for ECFP and mCerulean<sup>131</sup> exhibit a smaller peak at an average energy of 3.78 eV with a TPA cross section of 13 GM. This matches the computed long-range excitation state where B3LYP and PBE0 compute the energies to be 3.71 eV and 3.89 eV (respectively) and the cross sections to be 2 GM. The experimental spectra show a sharp rise at 4.509 eV for both proteins with the average cross section of 21 GM as in Table 2.4. As the spectrum ended abruptly at this energy, the experimentally reported cross section could be underestimated. Our computed values estimate a slightly larger cross section than the largest determined value. When energies are compared, we notice a large variation between functionals (approximately 0.9 eV). The deviation from the experimental energy follows a trend similar to that seen

with the BLB chromophore and to a lesser extent  $OR_A$  where the energy is lower for B3LYP and PBE0 and higher for CAM-B3LYP and LC-BLYP compared to the experimental one.

## **GFP**

For  $GFP_A$ , as in CFP and  $RFP_A$ , there is a lower-energy state for each functional (see Table A15 in Appendix A) with a weak TPA cross section ranging from 2 GM for B3LYP to 4 GM for LC-BLYP. This resembles the experimental spectra of some proteins that have the GFP chromophore such as mWasabi.<sup>132</sup> The experimental average for such a peak is 3.73 eV with a cross section of 11 GM. Using B3LYP and PBE0, this band is computed to have a cross section of 2 GM at 4.19 eV and 4.32 eV, respectively. On the other hand, the energy and (to a lesser extent) the cross section for the higher-energy peak match with the experimental average (4.30 eV, 13 GM) with B3LYP (4.23 eV, 4 GM) and PBE0 (4.36 eV, 6 GM) underestimating the cross section while CAM-B3LYP (4.76 eV, 24 GM) and LC-BLYP (4.90 eV, 47 GM) overestimate it. It is important to mention that the quoted experimental average of the highest energy peak is also for the highest value measured in that region and not for a true peak, but the rising curve is not steep indicating that the peak is most probably not much higher than the reported value. The analysis for  $GFP_N$  is very similar to that of  $GFP_A$ , as can be seen from the Table 2.4 and Table A16 and from the experimental TPA curves of proteins containing the  $GFP_N$  chromophore (mAmetrine,<sup>72</sup> for example).

## **KO<sub>A</sub>**

The experimental spectrum for mKok does not show an intermediate peak between the  $S_1$  absorption and the sharp rise in the higher energy range. It only shows a shoulder peak at 3.4 eV. All functionals predict 4 higher groups of excitations following the transition to  $S_1$  with large TPA cross sections (see A16). 2.4 shows the second group which corresponds to the fourth transition for B3LYP and PBE0 (3.62 eV and 3.77 eV). This group is preceded according to B3LYP and PBE0 by a very close peak,

which is the third transition, at 3.51 eV and 3.69 eV, respectively. This lower energy peak might be interpreted as the shoulder peak in the experimental spectra. The energy of the fourth state for B3LYP and PBE0 (3.62 eV and 3.77 eV, respectively) match the experimental measurements (3.62 eV) and is exaggerated by CAM-B3LYP and LC-BLYP (4.17 eV and 4.31 eV, respectively). Compared to the experimental value (93 GM), the cross section of the fourth state is underestimated by B3LYP and PBE0 (19 GM and 27 GM, respectively) and overestimated by CAM-B3LYP and LC-BLYP (205 GM and 339 GM, respectively). Further computed higher-energy excitations could not be correlated to the experimental spectrum, as it is attenuated at 3.62 eV.

### **OR<sub>A</sub>**

As with KO<sub>A</sub>, four groups of higher excitations with large TPA cross sections were computed in the case of OR<sub>A</sub>. Table 2.4 shows the first group which corresponds to the third excited state. The experimental data refers to the highest measured value in the high energy range of mOrange, but the rising curve starts to bend before it is attenuated suggesting that the measured value is very close to the true peak. As in the previous cases, excitation energies computed by B3LYP and PBE0 (3.65 eV and 3.80 eV, respectively) are closer to the experimental value (3.88 eV) than those computed by CAM-B3LYP and LC-BLYP (4.21 eV and 4.35 eV, respectively). The cross sections are underestimated with B3LYP and PBE0 and overestimated with LC-BLYP (22 GM, 30 GM and 320 GM, respectively) while that computed by CAM-B3LYP (189 GM) is the closest to the experimental value (200 GM). Again, computed higher-energy excitations could not be correlated to the experimental spectrum, as it ended at the aforementioned energy point (3.88 eV).

### **BLB**

The four functionals predict a lower-energy excitation before the one reported in Table 2.4, with very weak nonzero TPA cross section (1 GM according to B3LYP and PBE0). At this lower energy, the experimental curve starts to rise to form the stronger

TPA peak. The experimental curve is attenuated at 4.51 eV at a cross section of 22 GM. For both  $S_1$  and  $S_n$  excitations, there is significant variation between experimental and computed cross sections for mBlueberry1. Excitation energies, OPA oscillator strengths, and TPA cross sections for the model having the acylimine moiety out-of-plane can be found in Table A16. The crystal structure of a similar protein, Rtms5<sup>133</sup> (PDB ID: 3VK1), shows that the rings of the BLB chromophore are noncoplanar and assume a trans conformer (c.f. the cis conformer presented here). Computing TPA for the conformer where the experimental geometry is taken from the experimental X-ray structure (with hydrogen capping at the B3LYP/6-31+G(d,p) level of theory in PCM) resulted in a cross section of approximately zero GM (data not shown). This observation supports the hypothesis that the weak TPA that is experimentally observed for mBlueberry1 might be due to the distorted conformation its chromophore adopts. Unfortunately, there is no crystal structure for mBlueberry1, so we cannot decisively explore the deviation between experimental and computed cross sections.

## 2.4 Conclusions

TD-DFT was used on isolated chromophores to evaluate the efficiency of PBE0, B3LYP, CAM-B3LYP and LC-BLYP functionals in reproducing TPA absorption energies and cross sections of the corresponding FPs both in the gas phase and in water (PCM). Gas-phase computations yield larger TPA cross sections than the corresponding ones in water. Energies and cross sections of the first transition (lowest energy) are less functional dependent than the higher-energy transitions. For this first transition, the TPA energies and cross sections computed by CAM-B3LYP and LC-BLYP are systematically larger than the corresponding B3LYP and PBE0 values. We compared the computed vertical excitation energies to the TPA experimental energies (that are blue-shifted for many systems due to the involved vibronic transition) and to CC2 energies. The experimental lowest-energy was generally lower than the computed vertical excitation energy for the first transition and thus best reproduced with B3LYP (which generally gave the lowest energies). The TD-DFT energies were, in

general, lower than the corresponding CC2 ones. Within the present approximations neglecting vibronic and environmental effects, the experimental TPA cross sections are not quantitatively reproduced by any of the functionals which renders TD-DFT good at only qualitative or semiquantitative analysis. The situation is worse when it comes to higher-energy transitions. This is not very surprising provided the substantial variability in the experimental TPA cross sections of various proteins sharing the same chromophore. This is a preposed limitation as long as the study involves the isolated chromophores and ignores the protein shell. The electric field of the protein and conformational restraints exerted by the surrounding residues can significantly alter the TPA of the chromophore. The conformational dependence was highlighted here in the discussion of the RFP (BLB) chromophore where the carbonyl rotation considerably increased (decreased) the TPA cross section. Therefore, when screening chromophores with conformational flexibility for TPA, one should consider carefully the role of the conformation on the TPA cross section. For the higher-energy transitions, CAM-B3LYP and LC-BLYP predictions for both TPA cross sections and energies deviated significantly from the experimental values discouraging their use in any future TPA analysis. The erroneous overestimation of cross sections is caused by the artificial resonance enhancement of the TPA for the  $S_n$  state. On the other hand, B3LYP (and PBE0) yielded energies that are close to the averaged experimental values. B3LYP and PBE0, however, underestimate TPA cross sections (as compared to the averaged experimental values). Taking into consideration that the experimental cross sections are usually larger than those computed by B3LYP and PBE0, these two functionals seem suitable for computational screening of TPA to  $S_n$ . This conclusion is based on the present formalism adopting undamped response theory. Including a damping coefficient might change the behavior of the functionals, but this is not discussed in the present study. Moreover, care must be taken if one considers a more detailed analysis of the excited states as in some cases (e.g., RFP<sub>A</sub> and KO<sub>A</sub>, see Table 2.3 and Table 2.4), the energetic ordering may be incorrect for B3LYP and PBE0 due to the presence of low-lying states with long-range (charge-transfer) excitation character. Overall for TD-DFT studies on isolated FP chromophores, we recommend,

based on both energetics and absolute TPA cross sections, the use of B3LYP or PBE0 for the qualitative assessment of potential TPA probes accessing the lowest energy  $S_1$  states.

# Chapter 3

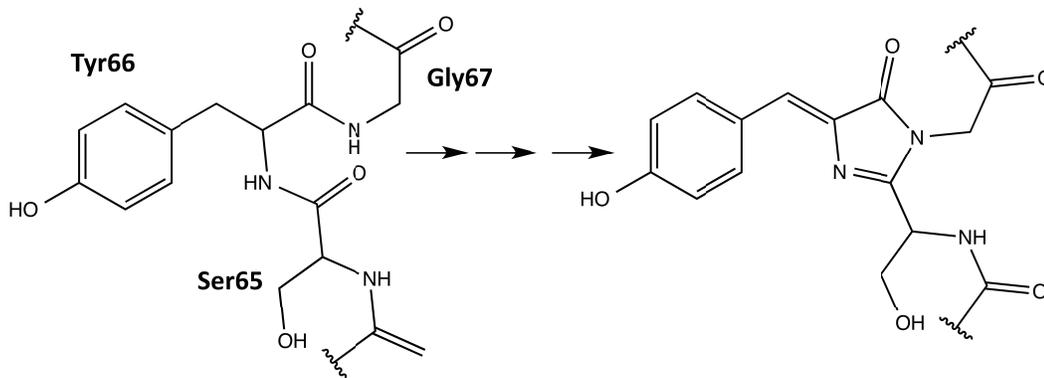
## Two-photon absorption of fluorescent protein chromophores incorporating non-canonical amino acids: TD-DFT screening and classical dynamics<sup>1</sup>

### 3.1 Introduction

Fluorescent proteins (FPs) are the family of homologues of the green fluorescent protein (GFP) of *Aequorea victoria* initially discovered in the 1960s.<sup>56</sup> FPs found great utility as spectroscopic tools after the cloning of the GFP gene<sup>57</sup> and the demonstration that they can be expressed in other organisms while maintaining their fluorescent properties.<sup>58,59</sup> The unique light-absorbing and fluorescence ability for FPs is due to the formation of a chromophore by a post-translational modification of three precursory amino acids within the protein shell (Fig. 3.1).<sup>60,61</sup> The coarse tuning of the colour of fluorescence is generally mediated by the alteration of the precursory amino acids which lead to different chromophores upon maturation. Changes in the micro-environment of the chromophore can also strongly affect the colour of the fluorescence, as seen in some yellow variants of the GFP.<sup>37</sup> In addition to fluorescence

---

<sup>1</sup>A version of this chapter has been published as: M. Alaraby Salem and Alex Brown, *Phys. Chem. Chem. Phys.*, 2015, **17**, 25563-25570.



**Figure 3.1:** *Chromophore maturation in GFP.*

wavelength, the protein environment generally offers fine tuning for all photophysical properties of the FP. A full palette spanning red to blue fluorescent proteins has been synthesized<sup>62,63</sup> with the red-shifted chromophores being of great interest due to their lower cell-toxicity.<sup>64</sup>

Two-photon microscopy of FPs offers many advantages over its conventional one-photon counterpart. It is less phototoxic, as photons of longer wavelength (and less energy) are absorbed and provides better focus and less out-of-focus bleaching enabling it to have deeper penetration into thick tissues.<sup>134,135</sup> These advantages arise because the two-photon absorption (TPA) probability, the so-called TPA cross section, is directly proportional to the square of incident light intensity. This property, however, causes TPA to have less sensitivity compared to its one-photon counterpart. Thus, fluorophores with large TPA cross sections are preferred - a major motivation for this work. TPA is governed by different quantum mechanical selection rules as compared to one-photon absorption (OPA) and so structural modifications in the protein environment can significantly affect the TPA of a FP with minimal effect on its OPA. For example, in the series of red FPs, their measured TPA cross sections range from 15 GM (for mTangerine<sup>72</sup>) to 119 GM (for tdTomato<sup>72</sup>) for the lowest-energy excitation although they all share the same chromophore.<sup>4</sup>

In addition to the known general difficulty of measuring absolute TPA cross sections,<sup>65</sup> measurements in biological systems like FPs are more challenging due to the need for additional calibration. Drobizhev et al. comprehensively explained and cited

the discrepancies in the reported absolute TPA cross sections of FPs.<sup>4</sup> Generally, the TPA spectrum of a FP has two regions of strong absorption: one is at (approximately) double the wavelength of the OPA peak and an additional (strong) band of absorption corresponds to a shorter wavelength. In the FPs with anionic chromophores, the TPA peak is blue-shifted with respect to the corresponding OPA peak (at half the wavelength). This has been rationalized by the enhancement of a vibronic transition in the two-photon process.<sup>29,69,80,102</sup> The additional band that is absent from the corresponding OPA spectrum was first theoretically predicted<sup>12</sup> to be present in the TPA spectra of all FPs and later confirmed through experimental measurements.<sup>4</sup> Theoretical investigations showed that the peak at longer wavelength is caused by the excitation to the first excited state ( $S_0$  to  $S_1$ ), while the other short-wavelength peak is due to a transition to a higher electronic level ( $S_0$  to  $S_n$ ). TPA corresponding to the higher-energy transitions has been shown to be amplified due to a resonance enhancement effect.<sup>12,13,47,103</sup> Being in the near-IR region, the  $S_0$  to  $S_1$  absorptions are of more practical relevance and thus are the focus of the present chapter.

Some theoretical studies of the TPA properties of FPs include the whole protein via combined quantum mechanics/molecular mechanics (QM/MM) approaches. List et al. studied the steric factors and chromophore protein-interactions that result in the enhancement of the TPA peak corresponding to the  $S_0$ - $S_1$  transition in DsRED,<sup>70</sup> a red FP.<sup>69</sup> They attributed the TPA enhancement to the increase in the difference between the permanent dipole moments of the excited and ground states. Another study on GFP succeeded in qualitatively reproducing most of the experimental features of the TPA spectrum.<sup>30</sup> Although the TPA cross section of a chromophore can be largely altered by the protein environment, studying the isolated chromophore can be a good starting point to predict or understand the TPA properties of the protein.<sup>12</sup> A study of the chromophore and close-by residues of a yellow variant of GFP employed RI-CC2 and TD-DFT with CAM-B3LYP to discern the effect of  $\pi - \pi$  stacking on TPA.<sup>26</sup> Although there was good qualitative agreement between the two methods, the values of TPA cross sections had to be scaled for comparison. In a benchmark study, Salem and Brown evaluated the use of several functionals by comparing the TPA of

isolated FP chromophores as computed via TD-DFT to averaged experimental data and higher-level CC2 computations.<sup>13</sup> Results showed that the B3LYP functional can provide a semi-quantitative description of the major TPA peaks. Recently, equation-of-motion coupled-cluster with single and double substitutions (EOM-CCSD)<sup>49,50</sup> was formulated for TPA and applied to chromophores of GFP and photoactive yellow protein.<sup>48</sup> TPA transition moment values computed with this method are comparable to TD-DFT values<sup>13</sup> for similar model chromophores. These studies support the use of computation to design rationally new chromophores.

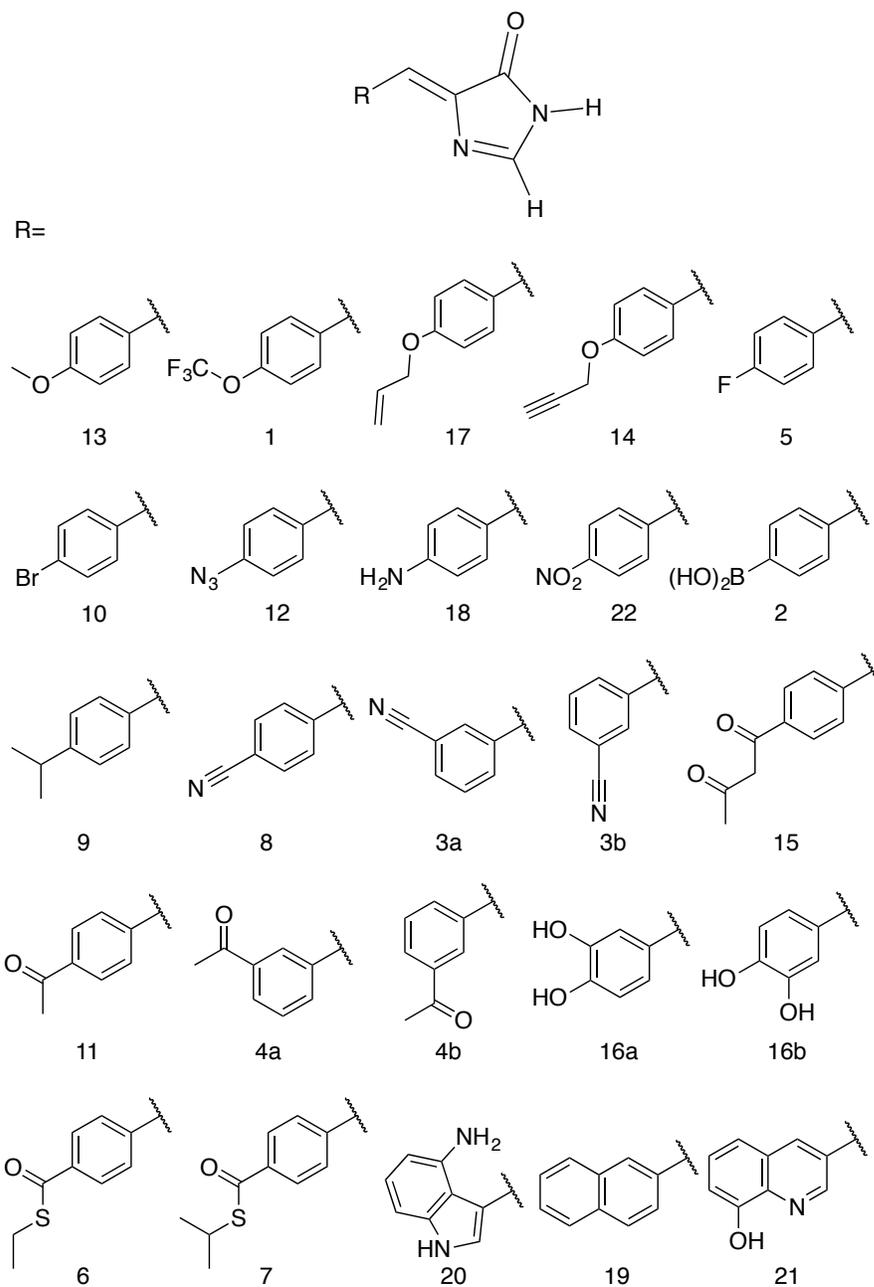
While many FPs have been engineered and a subset scrutinized computationally, they have, in general, been built from the canonical 20 amino acids. However, the protein engineering toolbox has been rapidly expanding as protein chemists have developed methods for incorporating non-canonical amino acids (ncAAs) into proteins.<sup>136–140</sup> Incorporating ncAAs can generate proteins with novel properties. A number of FPs containing ncAAs, which have been incorporated into the chromophore, have been engineered and experimentally characterized for their OPA and fluorescence properties;<sup>141–149</sup> to the best of our knowledge, TPA has not been explored for FPs containing ncAAs. A notable example for OPA is the Gold FP (GdFP),<sup>143</sup> which is represented by model 20 in Figure 3.2 where Trp57 and Trp66 in enhanced cyan FP (ECFP) have been replaced by 4-amino-Trp. These substitutions lead to a strongly red-shifted emission compared to ECFP. Site-specific substitutions of ncAAs for Tyr66 in GFP have also lead to novel chromophore structures with spectral properties notably different from the wild-type GFP.<sup>142,145</sup> As examples for residue-specific mutations, two tyrosine analogues (3-amino-L-tyrosine and 3-fluoro-L-tyrosine) have been incorporated into the DsRed-Monomer FP, leading to shifts in fluorescence wavelengths but, more importantly, increases in quantum yield.<sup>146</sup> While incorporation of ncAAs can directly influence the chromophore structure, ncAAs inserted outside the central chromophore can indirectly change its excitation and/or emission behavior.<sup>150</sup> Although using ncAAs in FP design clearly opens up new possibilities, the use of ncAAs is difficult. Thus, any newly designed FP must function better or differently than one that can be engineered using the 20 canonical amino acids. In this chapter,

we use TD-DFT to screen a variety of possible chromophores that can result from the replacement of the Tyr residue of the tripeptide precursor with one of the ncAAs previously used in protein synthesis. The property at focus is the TPA of the chromophore. The most promising candidate is further simulated in a proposed protein environment using molecular dynamics (MD) to study the protein stability and the steric effects of the protein on the TPA of the chromophore.

## 3.2 Computational Methods

### 3.2.1 Chromophores

A template chromophore model was obtained from GFP by breaking the chromophore connections to the rest of the protein and capping with H-atoms. The  $\pi$ -conjugated system necessary for TPA is preserved. In Chapter 2, we showed that methyl capping is only crucial when the chromophore has an extended conjugation beyond that in the GFP chromophore.<sup>13</sup> Since we only include chromophores derived from a GFP-chromophore template, capping with hydrogen atoms in this present study yields nearly the same TPA values as with methyl groups. We selected candidate amino acids from those compiled by Liu et al.<sup>137</sup> based on the following two criteria: (1) Having an aromatic ring that is necessary for extended conjugation. (2) Excluding bulkier systems that are less likely to fit in the chromophore cavity (without engineering the protein to accommodate the larger sized moiety), or that may not have enough flexibility to mature into the chromophore. For comparison, however, we considered 2 models (19 and 21) with two-cyclic rings that are comparable to that in the GdFP (model 20). The protein with the 2-naphthyl moiety (Model 19) was previously shown to be non-fluorescent probably because the cavity needed to be further manipulated to accommodate the chromophore, as suggested by the authors.<sup>142</sup> Assuming a point mutation at Tyr 66, the phenol ring of the chromophore is replaced with the corresponding moiety in the given ncAA yielding the chromophore models in Figure 3.2. For chromophores 1, 2 and 16, both the E and the Z isomers were considered and labelled a and b, respectively.



**Figure 3.2:** Chromophore models built from the parent GFP chromophore by replacing the phenol of Tyr-66 with the corresponding moiety in a ncAA.

### 3.2.2 Ab initio computations

Following a previous protocol (section 2.2.2),<sup>13</sup> the chromophore models were optimized in the gas phase using the PBE0 functional<sup>41,42</sup> (optimized coordinates are given in Table B1, Appendix B). Excited-states properties were computed with TD-DFT<sup>20</sup> within the response theory framework using the B3LYP functional<sup>40</sup> and the conductor-like polarizable continuum model (commonly referred to as PCM)<sup>120–123</sup> with parameters for water - except where noted. The basis set 6-31+G(d,p)<sup>114–118</sup> in cartesian form, i.e., 6 d-functions, was used in all computations. OPA oscillator strengths (and similarly the transition dipole moments from ground to excited states) were computed via linear response<sup>119</sup> while the two-photon transition matrix elements and the transition dipole moments between excited states were evaluated from the single and double residues of the quadratic response function,<sup>11,21–23</sup> respectively. Dipole moment elements not involving the ground state ( $\langle n|\mu|n\rangle$  and  $\langle n|\mu|m\rangle$ , where  $n, m \neq 0$ ) were computed via the DALTON<sup>51</sup> software package in the gas phase. All other computations were done with GAMESS-US<sup>52</sup> (the May-2013 version). For computations with PCM, default parameters in the GAMESS code were utilized.

For linearly polarized light, the transition moment for TPA is

$$\delta^{TPA} = \frac{1}{15} \sum_{\alpha\beta} [S_{\alpha\alpha}S_{\beta\beta}^* + 2S_{\alpha\beta}S_{\alpha\beta}^*], \quad (3.1)$$

where the elements of the two-photon transition matrix are given by:

$$S_{\alpha\beta} = \sum_n \left[ \frac{\langle 0|\mu_\alpha|n\rangle\langle n|\mu_\beta|f\rangle}{\omega_n - \omega} + \frac{\langle 0|\mu_\beta|n\rangle\langle n|\mu_\alpha|f\rangle}{\omega_n - \omega} \right]. \quad (3.2)$$

In Eq. 3.2  $\mu_\alpha$  and  $\mu_\beta$  refer to the dipole moment operator in a given cartesian direction ( $\alpha, \beta = x, y$  and  $z$ ),  $\omega_n$  is the energy gap from the ground state,  $|0\rangle$ , to a given state  $|n\rangle$ ,  $\omega$  is the photon energy and  $|f\rangle$  is the final excited state.

From the TPA transition moment and excitation energies ( $\omega_f$ ) produced by GAMESS, the TPA cross section is calculated in macroscopic units by:

$$\sigma^{TPA} = \frac{4\pi^2 a_0^5 \alpha \omega^2}{c \Gamma} \delta^{TPA}, \quad (3.3)$$

where  $\alpha$  is the fine structure constant,  $a_0$  is the Bohr radius,  $c$  is the speed of light,  $\omega$  is the photon energy ( $=\frac{\omega_f}{2}$ ) and  $\Gamma$  is the broadening factor derived from a Lorentzian function and chosen to be 0.1 eV, as previously employed<sup>9,12,13</sup> for comparison with experiment. The choice of the conversion equation (Eq. 3.3) and the broadening factor,  $\Gamma$ , affect the resulting values of the TPA cross sections. In a recent study, Beerepot et al.<sup>16</sup> discussed the various forms of Eq. 3.3 and gave recommendations on presenting TPA for computational studies.

### 3.2.3 Molecular Dynamics Simulation

To test the stability of a protein upon the introduction of a selected ncAA, MD trajectories were generated following the same protocol for two FPs: (1) a reference EGFP (PDB ID: 2Y0G) with the corresponding anionic chromophore which we refer to as *control* and (2) the same protein after replacing the chromophore with Model 22 (see Figure 3.2) that is assumed to be formed by replacing Tyr 66 with a Tyrosine-derived ncAA. We refer to this modified protein as *nitro*.

The crystal structure for the *control* and its modified *nitro* version were prepared using the `pdb4amber` and `reduce` programs<sup>151</sup> in `Ambertools 14`. The 2Y0G crystal structure is missing 12 residues from the protein termini and hence these are unlikely to affect the dynamics of the  $\beta$ -barrel or the chromophore environment. The missing residues are not considered and the protein is renumbered, so that the chromophore is formed by residue 63. The all-atom forcefield AMBER ff12SB<sup>152,153</sup> was used to parameterize both protein models except for the chromophore residue. The chromophore in both cases includes all atoms between the LEU 62 and the VAL 64 so that more linker atoms are considered than in the attenuated models used for DFT screening (see Table B2 and Table B3, Appendix B). Although previously validated parameters are available for the *control* chromophore,<sup>154</sup> we adopted a general procedure to parameterize both *nitro* and *control* chromophores and it can be easily extended to test other residues of interest. The parameters generated here serve the purpose of determining the protein stability and conformational freedom of the chromophore. We used ANTECHAMBER<sup>155</sup> to generate parameters for the *nitro*

chromophore that are consistent with the General Amber Force Field (GAFF).<sup>156</sup> We assigned similar parameters to the *control* chromophore. Charges were derived using the online R.E.D. server development tool<sup>157</sup> following the default scheme for amino acid fragments. All parameters are given in Appendix B (see Tables B2 and B3 for atom types and charges). All crystallographic water molecules were removed, including those in the vicinity of the chromophore to enable extra conformational freedom. Each protein model was solvated with approximately 23,670 TIP3P water molecules in a cuboid solvation box with edge length of 20 Å. To neutralize the negatively charged protein, 7 Na<sup>+</sup> ions were added to each model followed by 64 Na<sup>+</sup> and Cl<sup>-</sup> ions to reach a salt concentration of 0.15 M.

The MD simulation was done with the AMBER Molecular Dynamics package<sup>158</sup> following a standard protocol that consists of minimization, heating, density equilibration and production. Minimization was done first with restraints on the protein atoms and then repeated without restraints. Heating was applied gradually for 20 ps with restraints on the protein atoms. Density equilibration was achieved in four 50-ps runs gradually relieving the restraint. This was followed by a production run at constant pressure for 99 ns. Langevin dynamics were employed globally throughout the simulation. Details of the simulation are provided in terms of Amber input files in Appendix B. Trajectories were analyzed via CPPTRAJ.<sup>159</sup>

### 3.3 Results and Discussion

Computing the TPA of FPs involves several levels of complexity. In addition to the intrinsic nature of the chromophore, there are other factors that affect the TPA of a FP. One factor is that the protein shell can change the conformation of the chromophore to enhance or diminish its two-photon absorption cross section.<sup>13,69</sup> This factor can be accounted for via TD-DFT which can capture the change in TPA associated with various conformers in a semi-quantitative fashion.<sup>13</sup> Another level of complexity is added by the protein-chromophore interactions, or the electric field due to the protein around the chromophore, which can greatly influence the TPA cross

section.<sup>4,69</sup> In this chapter, we compute TPA cross sections for isolated chromophores ignoring the protein shell (Section 3.3.1). We then account for part of the influence of the protein shell by running a classical MD simulation for an EGFP-based protein with the chromophore predicted to have the largest cross section (*nitro*) and compare it to an analogous simulation for its native form (*control*). The motivation is to obtain insight into the relative stability of the protein after introducing the new moiety (Section 3.3.2) and to account for part of the influence of the protein shell on the chromophore, through studying its flexibility over the trajectory and computing TPA for different conformations of the isolated chromophore (Section 3.3.3).

### 3.3.1 TPA cross sections

The TPA cross sections for the lowest-energy transition ( $S_0$  to  $S_1$ ) of all the GFP-derived chromophores with natural amino acids have been previously computed (see Chapter 2).<sup>13</sup> Their TPA cross sections at the B3LYP/6-31G+(d,p) level of theory in PCM H<sub>2</sub>O range from 1 GM (for the BFP<sup>111,112</sup> chromophore) to 7 GM (for the CFP<sup>112</sup> chromophore). An equivalent range of TPA cross sections is determined for molecules 5 through 18 of this chapter. TPA cross sections, as computed from Eq. 3.3, are given in Table 3.1 while  $\delta^{\text{TPA}}$  values are given in Table B4 in Appendix B. Proteins with bromo, methoxy and amino substituted chromophores (models 10, 13 and 18, respectively) have been previously synthesized and shown to be fluorescent.<sup>142</sup> Their measured OPA energies are: 3.31 eV, 3.15 eV and 2.85 eV, respectively.<sup>142</sup> Compared to the values 3.37 eV, 3.29 eV and 3.14 eV in Table 3.1, TD-DFT using B3LYP/6-31+G(d,p) captures the proper trend within the expected error. For model 20, the computed energy of 2.69 eV is very close to the measured absorption peak at 2.66 eV for the corresponding GdFP.<sup>143</sup> The trend in Table 3.1 looks very promising because, in general, the chromophores with the largest computed TPA cross sections are the most red-shifted ones.

Although the computation was done in the response theory framework, comparison to a truncated sum-over-states expression gives insight into the factors contributing to the TPA cross section. In a 2-level model (2LM) approximation, the TPA

**Table 3.1:** *One-Photon Excitation Energies, OPA oscillator strengths (OS) and TPA cross sections for the Transition to  $S_1$  as determined at the B3LYP/6-31G+(d,p) Level of Theory in PCM with parameters for  $H_2O$ .*

Model	Energy (eV)	OS	TPA (GM)
1a	3.455	0.640	0
2a	3.468	0.679	0
1b	3.473	0.652	0
2b	3.452	0.557	0
3	3.463	0.705	0
4	3.387	0.787	0
5	3.467	0.704	1
6	3.197	0.912	1
7	3.200	0.927	1
8	3.304	0.796	2
9	3.403	0.841	3
10	3.369	0.848	3
11	3.256	0.812	4
12	3.210	0.984	5
13	3.289	0.853	7
14	3.310	0.894	7
15	3.222	0.837	7
16a	3.218	0.577	8
16b	3.260	0.711	7
17	3.284	0.901	8
18	3.137	0.927	9
19	3.138	0.553	11
20	2.689	0.363	15
21	2.985	0.298	17
22	2.965	0.638	29

Model	$(\Delta\mu_x)^2$	$(\Delta\mu_y)^2$	$\mu_{x01}^2$	$\mu_{y01}^2$	$\sigma_{2LM}$
6	0.0	0.1	9.8	0.1	1
7	0.1	0.1	9.9	0.1	1
12	0.6	0.0	10.5	0.0	5
13	0.8	0.0	8.7	0.0	6
14	0.8	0.0	9.3	0.1	6
16a	5.2	0.0	5.0	0.2	22
18	1.2	0.0	9.6	0.0	10
19	8.8	0.1	4.0	0.1	31
20	10.6	0.7	3.8	0.1	34
22	7.0	0.0	7.6	0.0	46

**Table 3.2:** The (non-zero) dipole elements (in atomic units squared) of the 2-state model computed at the B3LYP/6-31G+(d,p) level in the gas phase.  $\Delta\mu_\alpha$  is the  $\alpha$ -component of the difference between the permanent dipole moment of the first excited and the ground states, i.e.,  $(\langle 1|\mu_\alpha|1\rangle - \langle 0|\mu_\alpha|0\rangle)$ .  $\mu_{\alpha 01}$  is the  $\alpha$ -component of the transition dipole moment for the transition from the ground state to the first excited state.  $(\Delta\mu_z)^2$  and  $\mu_{z01}^2$  are zero (0.0) for all models.  $\sigma_{2LM}$  is the TPA cross section determined via the 2LM directly from the corresponding elements for each chromophore.

cross section is proportional to the square of the difference between the permanent dipole moments of the excited and ground states  $(\langle 1|\mu|1\rangle - \langle 0|\mu|0\rangle)^2$  and that of the transition dipole moment from the ground to the excited state  $(\langle 0|\mu|1\rangle)^2$ . The dipole elements for the chromophore models were determined in the gas phase, as the corresponding PCM computation were difficult to converge in DALTON. This change in medium does not affect the analysis, as the trend of TPA cross sections for the first bright transition is the same whether computed with PCM or in the gas phase (Table B7, Appendix B). The dipole elements and the corresponding cross sections ( $\sigma_{2LM}$ ) calculated directly using Eqs. 3.1 - 3.3 are given in Table 3.2 for the models where the first gas-phase excitation corresponds to the first PCM one.

There is a significant discrepancy between the absolute  $\sigma$  values computed via response theory (see Table B7, Appendix B) and the corresponding 2LM ones (Table 3.2). However, the trend is the same (with the exchange of order for models 16a and 18). Since all studied molecules are nearly planar (symmetry was not enforced

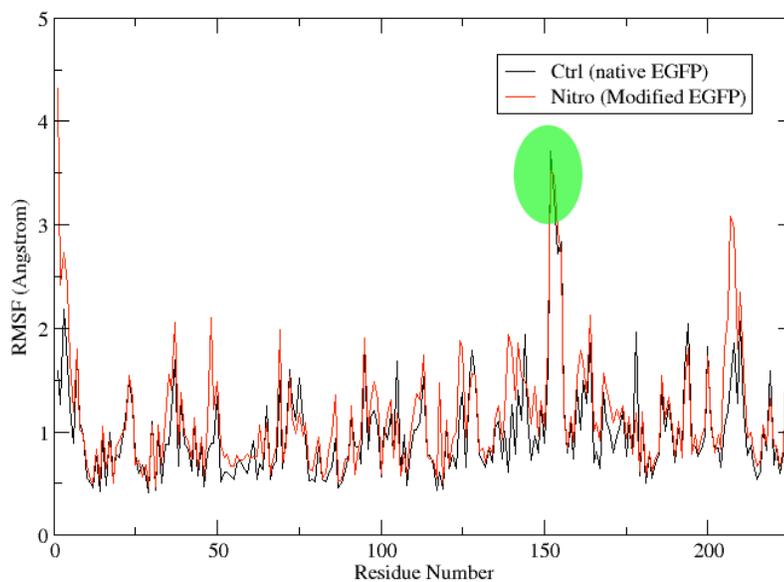
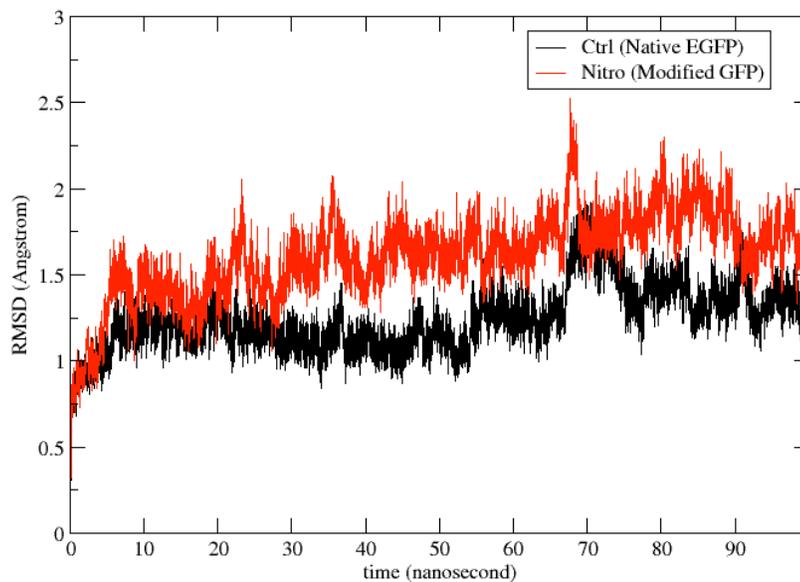
during geometry optimization), there is no contribution from dipole elements along the  $z$ -axis. Most of the contribution comes from the dipole moments along the  $x$ -axis which runs through the  $\pi$ -conjugated system. The nitro-derivative, molecule 22, has both large dipole difference and transition dipole moment which explains the large cross section obtained via response theory computation.

### 3.3.2 Protein Stability

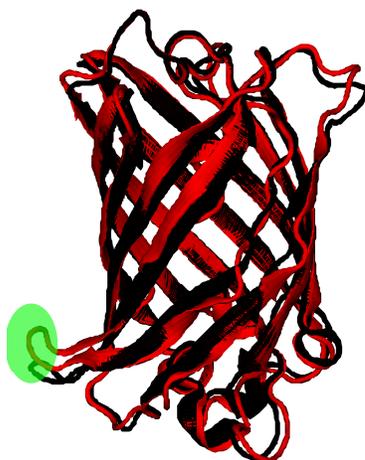
The root-mean-square deviation (RMSD) for the backbone atoms and the root-mean-square fluctuation (RMSF) for the protein residues are shown for the two trajectories in Figure 3.3. As expected due to the modification introduced to the *nitro* model, its RMSD is larger than the *control*. The RMSD deviation, however, is still within range of the crystal-structure resolution of 1.5 Å. A comparison of the average bond lengths in the chromophores from the MD simulations with the DFT-optimized values, and, for the *control*, those from Xray crystallography show no significant deviations (see Tables B5 and B6 in Appendix B). Figure 3.4 shows a superposition of an average structure for each model generated from the trajectory between 60 and 99 ns. Residue 153 with the largest RMSF is a loop residue outside the  $\beta$ -barrel structure and thus is more flexible than the residues composing the  $\beta$ -sheets (see Fig. 3.4). In addition to the duly conserved 3D structure in the modified model of the protein, the unique neutral form of the nitro chromophore should make it, in principal, less sensitive to changes in the surrounding micro-environment.

### 3.3.3 Conformational Analysis of The Chromophore

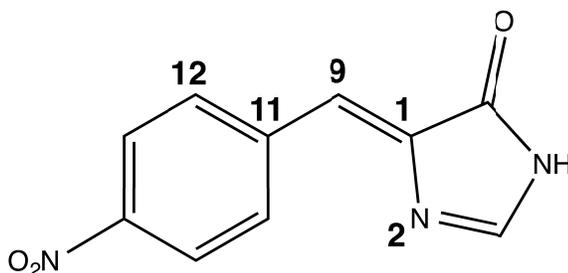
We monitored the conformational flexibility of the chromophore using three characteristic angles: the angle corresponding to the methine bridge, as well as the tilting and the twisting angles of the nitro-benzylidene moiety with respect to the imidazolone ring (see Figure 3.5). The three angles were recorded for 49511 snapshots over the course of 99 ns. The minimum, maximum and average values for these angles are shown in Table 3.3. The methine bridge shows the least flexibility with more than 14% of the snapshots having the average angle of 134° and 95% of the snapshots



**Figure 3.3:** *The RMSD of the protein in reference to the original minimized structure (top) and the RMSF of the protein residues (bottom) over the simulation period of 99 ns. Fluctuation for residue 153 is highlighted (see Fig. 3.4).*



**Figure 3.4:** An overlay of average structures of nitro (red) and control (black) models generated from the interval between 60 and 99 ns. Residue 153 (highlighted green) is the non-terminus residue with the largest RMSF fluctuations.

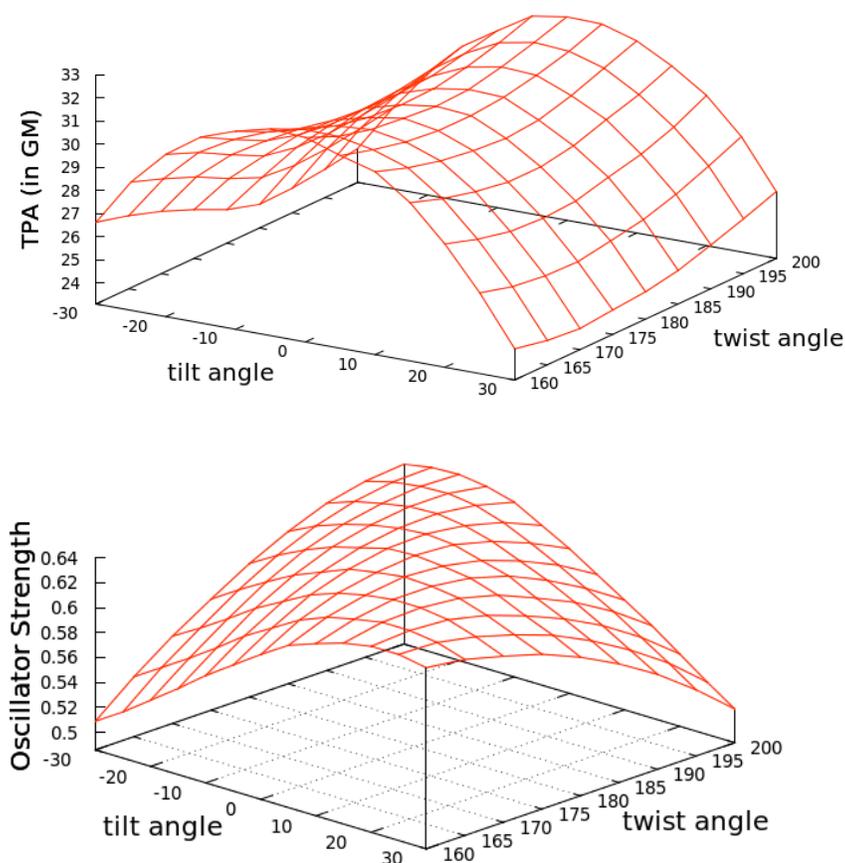


**Figure 3.5:** The nitro chromophore model showing the three angles monitored in the conformational study: the methine bridge is the angle between atoms 11 - 9 - 1, the tilt angle is the dihedral between atoms 11 - 9 - 1 - 2 and the twist is the dihedral between atoms 12 - 11 - 9 - 1. The full  $z$ -matrix is given in Table B8 in Appendix B.

**Table 3.3:** Minimum, maximum and average values for the three angles defined in Figure 3.5 determined from 49511 snapshots of the full simulation trajectory.

	Methine Bridge	tilt	twist
minimum	123°	-49°	152°
maximum	145°	60°	207°
average	134°	-6°	178°

having an angle within  $129^\circ$  and  $139^\circ$ . The twisting and tilting angles show more flexibility where 6.5% and less than 3.5% snapshots have the average angles for each, respectively. In 95% of the snapshots, the twist angle ranges from  $160^\circ$  to  $200^\circ$  and the tilt angle ranges from  $-30^\circ$  to  $30^\circ$ . Hence, we generated 117 conformers by varying the twist and tilt angles by  $5^\circ$  within these ranges and fixing the methine bridge at the average angle of  $134^\circ$ . For each conformer, we computed the first excitation energy, OPA and TPA at the same level of theory used in screening the chromophore models, that is, TD-B3LYP/6-31+(d,p) in PCM with parameters for  $\text{H}_2\text{O}$ . The trends for the TPA cross section and OPA oscillator strength are illustrated in Figure 3.6.

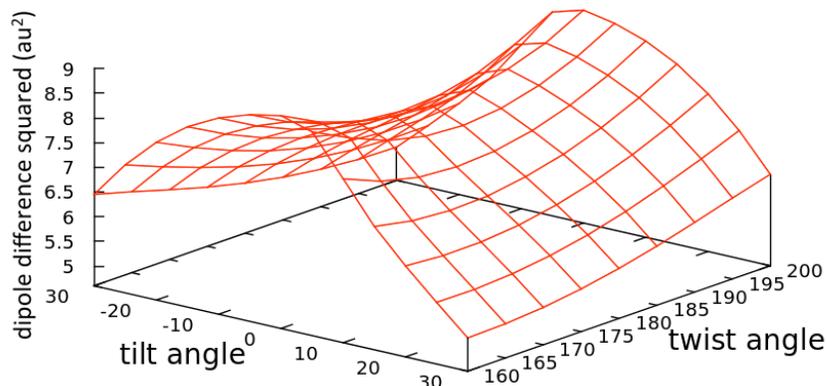


**Figure 3.6:** *The variation of TPA cross section (top) and OPA oscillator strength (bottom) with the tilt and twist angles while fixing the methine bridge at  $134^\circ$  (see Figure 3.5 for the definition of the angles). The tabulated values are given in Table B9 in Appendix B*

The trend in the excitation energy is similar to that for the TPA cross section. Nevertheless, the TPA trend is *not* driven by the change in energy, as the largest energy difference in the set of conformers is less than 0.1 eV (Fig. B1, Appendix B). Further, the same TPA trend is generated even if the same excitation energy is used to calculate the TPA cross section for all conformers. For the OPA oscillator strength, a uniform parabola can be noticed when the tilt angle is fixed to the planar value,  $0^\circ$ , and the twist is varied, or the twist is fixed to  $180^\circ$  and the tilt is varied. In such cases, the oscillator strength decreases upon deviation from planarity. The decrease is the same whether the tilt or the twist is varied. As the fixed angle deviates from the planar value, the curve is skewed. On the other hand, the TPA cross section increases when the twist angle deviates from  $180^\circ$  and decreases when the tilt angle deviates from a planar or near-planar value. The TPA value is significantly more sensitive to the tilt angle than it is to the twist angle. To further investigate the reason for such trends in the TPA cross sections, we computed the difference between the first excited state permanent dipole moment and the ground state dipole moment for each conformer ( $\langle 1|\mu|1\rangle - \langle 0|\mu|0\rangle$ ) at the TD-B3LYP/6-31G+(d,p) level of theory in the gas phase. The square of the x-components of the dipole difference are plotted in Figure 3.7. The resemblance in trend and the magnitude of the difference as the tilt and twist angles change strongly confirm that the TPA cross section variation is driven by the difference between permanent dipoles. These results could further guide the protein engineering of the chromophore cavity to optimize its TPA, where a (near) planar value is needed for the tilt angle and deviation of the twist angle from planarity is desirable.

### 3.4 Conclusion

A group of 22 proposed FP chromophore models derived from ncAAs are screened for their TPA cross sections. TD-DFT employing the B3LYP functional was used in the screening. Most of the studied molecules exhibited poor intrinsic TPA cross sections similar to the naturally occurring GFP-derived chromophore models. Molecules with



**Figure 3.7:** The effect of changing the tilt and twist angles while fixing the methine bridge at  $134^\circ$  (see Figure 3.5) on the square of the difference between the  $x$ -components of the permanent dipoles of the first excited and the ground states ( $\langle 1|\mu_x|1\rangle - \langle 0|\mu_x|0\rangle$ ) of the nitro chromophore (model 20) computed at the TD-B3LYP/6-31G+(d,p) level of theory in the gas phase. The tabulated values are given in Table B9 in Appendix B.

multiple rings, as expected due to the extended conjugation, showed relatively large TPA cross sections. The chromophore with the largest cross section, however, was the one derived from a nitro-based ncAA. Two-state model analysis suggests that the increase in TPA is due to both its large transition dipole moment ( $\langle 0|\mu|1\rangle$ ) and the significant difference between the permanent dipole moments of its first excited and ground state ( $\langle 1|\mu|1\rangle - \langle 0|\mu|0\rangle$ ).

To investigate this model further, MD simulations were run on both a native EGFP (*control*) and a nitro-derived EGFP (*nitro*). Comparison with the *control* showed that the protein was stable after the replacement of the hydroxyl group with the nitro substituent. A conformational analysis was then performed to study the change of TPA with a range of the conformations visited by the chromophore in the protein cavity. Results show that a large TPA cross section (24 – 32 GM) is maintained through the various conformations and that the TPA fluctuation is, again, driven by the change in the difference between the permanent dipole moments of its first excited state and its ground state ( $\langle 1|\mu|1\rangle - \langle 0|\mu|0\rangle$ ).

In this proposed model, we accounted for two degrees of complexity, that is, the nature of the chromophore and the effect of the protein environment on the chromophore conformation. There still remains the consideration of the chromophore-protein interactions and the electric field due to the protein shell; both can affect the TPA cross section. The sensitivity to the surrounding electrostatic environment of the chromophore is due to the dependence of the TPA cross section on  $\langle 1|\mu|1\rangle - \langle 0|\mu|0\rangle$ . The red FPs share the same chromophore that has an intrinsic TPA cross section of about 5 GM (see Chapter 2), as computed previously at the same level of theory used in this chapter (TD-B3LYP/6-31+G(d,p)).<sup>13</sup> However, due to the protein shell, some red FP proteins reach an experimental TPA cross section of 139 GM;<sup>4</sup> a 27-fold amplification. This amplification has been attributed to the sensitivity of the difference between permanent dipoles ( $\langle 1|\mu|1\rangle - \langle 0|\mu|0\rangle$ ) to the electric field of the protein.<sup>4</sup> The *nitro* model, having most of its TPA driven by a large difference between permanent dipoles, seems to be a promising FP target especially if properly engineered to amplify its large intrinsic cross section.

# Chapter 4

## Prediction of two-photon absorption enhancement in red fluorescent protein chromophores made from non-canonical amino acids<sup>1</sup>

### 4.1 Introduction

Fluorescent proteins (FPs) make up a family of homologues of the *Aequorea victoria* green fluorescent protein (avGFP) initially discovered in the 1960s.<sup>56</sup> The FP chromophore is made by a post-translational modification of three precursory amino acids within the protein shell.<sup>60,61</sup> In some FPs, that are of interest in this chapter, there is an additional maturation step, resulting in a chromophore with an extra acylimine moiety (see Figure 4.1 for a depiction of the red FP (RFP) chromophore).<sup>70</sup> The chromophore structure together with the surrounding protein environment influence the photophysical properties of the protein.<sup>37</sup> Modifications to the precursory amino acids have enabled protein engineers to develop a full spectrum of FPs, ranging from blue-to-red fluorescence.<sup>63,64</sup> Currently many photophysical studies on FPs have focused on their one-photon absorption (OPA), with a recent increasing interest in

---

<sup>1</sup>A version of this chapter has been accepted for publication as: M. Alaraby Salem, Isaac Twelves and Alex Brown, *Phys. Chem. Chem. Phys.*, 2016, Accepted.

two-photon absorption (TPA) properties.<sup>62,160</sup>

In TPA microscopy, light of longer wavelength (smaller energy) is absorbed, decreasing the chance of irreparable cell damage associated with higher-energy photons and enabling deeper penetration into thick samples. As TPA varies with the square of the incident light intensity, there is less out-of-focus bleaching and more focused imaging.<sup>134,135</sup> This advantage comes at the expense of sensitivity and thus fluorophores with large TPA probabilities (cross sections) are needed.

There is a known challenge in measuring absolute TPA cross sections, especially in biological molecules.<sup>65,160</sup> As discussed by Drobizhev et al.,<sup>4</sup> there exists large deviations - up to two orders of magnitude - between TPA cross section measurements for the same FP; for example, measurements made for the lower energy peak of an enhanced green-fluorescent protein (EGFP)<sup>66</sup> range from 1.5 GM<sup>67</sup> to 180 GM.<sup>68</sup> In any case, while large deviations may occur measurement-to-measurement, the experimental TPA spectra of the FPs consistently exhibit two regions of absorption: one at roughly double the wavelength of the OPA peak and another, stronger band shifted to a shorter wavelength.<sup>4,27</sup> The longer wavelength band is blue-shifted relative to double the wavelength of the corresponding OPA peak in the anionic chromophore spectra due to the enhancement of a vibronic transition.<sup>29,69,80,102</sup> The peak at a longer wavelength was determined theoretically as the excitation to the first (electronic) excited-state,  $S_0$  to  $S_1$ , while the second, higher energy peak corresponds to a transition to a higher excited-state,  $S_0$  to  $S_n$ . The main focus of this chapter, however, is on the  $S_0$  to  $S_1$  absorptions, as they occur in the near IR-region, and thus are more pertinent to imaging. Proteins of orange to far-red fluorescence have efficient TPA in the range between 1,000 nm and 1,200 nm (beyond the tuning range for Ti:sapphire lasers) where there is higher tissue transparency, weak scattering and very little autofluorescence.<sup>4</sup> Measurements of TPA cross sections of a series of RFPs showed how the protein environment strongly affects the TPA of the chromophore.<sup>4</sup> Although they have the same chromophore structure (see the RFP chromophore model in Figure 4.1), the first bright state of mTangerine<sup>72</sup> has a measured TPA cross section of 15 GM while that of a monomer of tdTomato<sup>72</sup> is 139 GM.<sup>4</sup>

The computation of TPA cross sections is likewise challenging. Time-dependent density functional theory (TD-DFT) is used extensively to compute TPA cross sections for large molecules due to its reasonable computational expense and relatively good accuracy.<sup>12,24-31</sup> Until recently, higher level ab initio methods were restricted only to the study of the TPA of small molecules. An investigation into the yellow fluorescent protein (YFP) showed that the CAM-B3LYP functional yields similar qualitative TPA results to the resolution-of-identity (RI) CC2 method.<sup>26</sup> Another study that compared TD-DFT and full CC2 results showed that the B3LYP functional with the modest 6-31+G(d,p) basis can be used for a semi-quantitative comparison of TPA for the lowest energy excitation of FP chromophores.<sup>13,47</sup> Beerepoot et al. duly benchmarked CAM-B3LYP and RI-CC2 results against equation-of-motion CCSD (EOM-CCSD)<sup>48-50</sup> for a set of neutral FP chromophores.<sup>16</sup> Their results show that CC2 results are slightly overestimated as compared to EOM-CCSD ones within a factor of 1.4 while CAM-B3LYP results are significantly underestimated by a factor of 1.5 to 3.<sup>16</sup>

In recent years, great strides have been made in protein engineering following the development of methods to incorporate non-canonical amino acids (ncAAs) into proteins.<sup>136-140</sup> Barring ongoing research, only OPA has been experimentally explored for FPs containing ncAAs:<sup>141-149</sup> for example, the Gold FP (GdFP) - based off of the enhanced cyan FP (ECFP) with the replacement of Trp66 and Trp57 with a ncAA - saw a significant red-shifted emission compared to the ECFP as a result of the ncAA substitutions.<sup>143</sup> While ncAA incorporation can have a direct effect on chromophore structure, it can also have an indirect effect on the nature of excitation and/or emission when placed outside the central chromophore.<sup>150</sup> Important to realize, however, is that ncAA incorporation remains a difficult task. Consequently, each ncAA incorporated must ultimately generate a protein with sufficiently differing or enhanced functionality.

In Chapter 3,<sup>75</sup> 22 chromophore models (constructed from the replacement of the Tyr66 residue of the tri-peptide precursor of a GFP template chromophore with a ncAA) were screened for excited-states properties; mainly TPA. Molecular dynamics

simulations were further run to test the stability of a proposed FP containing the chromophore with maximal TPA cross section (a nitro-substituted chromophore; similar to Model 22 in Figure 4.1). Recently, interest into mutating red-fluorescent proteins (RFPs) with ncAAs has been piqued.<sup>161</sup> In the present chapter, we computed OPA and TPA properties for the same set of 22 chromophores considered previously,<sup>75</sup> each of them now having an acylimine moiety resembling the extra maturation step in the RFPs. As the RFP chromophores occur in various conformations in reported crystal structures, we explore the effect of rotating the carbonyl of the acylimine moiety on the OPA and TPA of the chromophores with relatively large TPA cross sections. We further investigate the effect of altering the tilt and twist angles between the rings of the chromophore with the largest intrinsic TPA cross section. Inter-functional comparison between the TPA cross sections computed by B3LYP and CAM-B3LYP is presented. We also compare our findings to the previous data<sup>75</sup> from GFP-derived chromophores.

## 4.2 Computational Methods

The chromophore structures were adopted from Chapter 3<sup>75</sup> with two modifications: (1) an extra extension is added to account for the acylimine moiety characteristic of RFP; (2) the broken connections between the chromophore and the rest of the protein are capped by methyl groups rather than hydrogen atoms. A subset of the chromophores considered is shown in Figure 4.1 while the full list is given in Figure C1 in Appendix C. Within the protein shell, the carbonyl of the acylimine is not coplanar with the rings of the chromophore, as observed in the crystal structures we analyzed (see the discussion in Section 4.3.2 and Table C1 in Appendix C). Hence, methyl capping is necessary when RFP-based chromophores are optimized to avoid obtaining a planar structure which would be far from the conformation (potentially) dictated by the protein shell.<sup>13</sup> For the models with large TPA cross sections, the effect of acylimine rotation on the TPA is further investigated.

We followed our previous protocol for optimization and computation of excited-

states properties (see Chapters 2 and 3):<sup>13,75</sup> Dipole moments used in the two-level model (2LM) analysis were computed with DALTON<sup>51</sup> (2016 version) while all other computations were done with GAMESS-US (May-2013 version).<sup>52</sup> Optimization was done in the gas phase using the PBE0 functional.<sup>41,42</sup> The optimized coordinates for all structures are given in Table C2 of Appendix C.

Excited-state properties were computed with TD-DFT<sup>20</sup> within the response theory framework; OPA oscillator strengths (OS) and energies were computed via linear response,<sup>119</sup> while the two-photon transition matrix elements and the difference between permanent dipole moments of the excited and ground states were evaluated from the single and double residues of the quadratic response function,<sup>11,21–23</sup> respectively.

Assuming linearly polarized light, the transition moment for TPA is

$$\delta^{\text{TPA}} = \frac{1}{15} \sum_{\alpha\beta} [S_{\alpha\alpha}S_{\beta\beta}^* + 2S_{\alpha\beta}S_{\alpha\beta}^*], \quad (4.1)$$

where the elements of the two-photon transition matrix are given by:

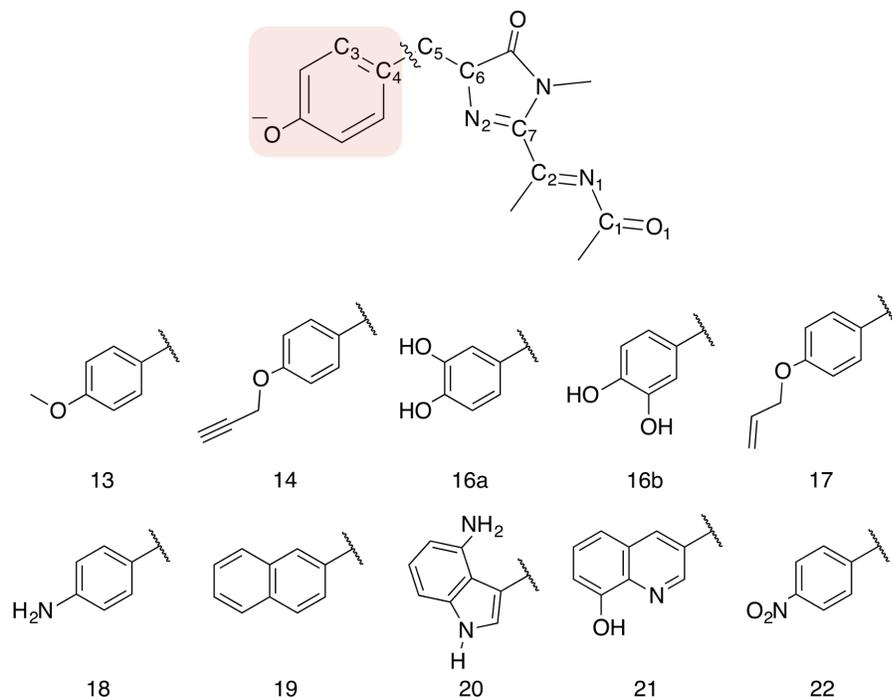
$$S_{\alpha\beta} = \sum_n \left[ \frac{\langle 0|\mu_\alpha|n\rangle\langle n|\mu_\beta|f\rangle}{\omega_n - \omega} + \frac{\langle 0|\mu_\beta|n\rangle\langle n|\mu_\alpha|f\rangle}{\omega_n - \omega} \right]. \quad (4.2)$$

In Eq. 4.2,  $\mu_\alpha$  and  $\mu_\beta$  refer to the dipole moment operator in a given Cartesian direction ( $\alpha, \beta = x, y$  and  $z$ ),  $\omega_n$  is the energy gap from the ground state,  $|0\rangle$ , to a given state  $|n\rangle$ ,  $\omega$  is the photon energy ( $=\frac{\omega_f}{2}$ ) and  $|f\rangle$  is the final excited state.

The TPA cross section,  $\sigma^{\text{TPA}}$ , is commonly reported in units of  $10^{-50} \text{ cm}^4 \text{ s molecule}^{-1} \text{ photon}^{-1}$  that is referred to as one GM (in honour of the work<sup>1</sup> of ‘‘Maria G6ppert-Mayer’’). From the TPA transition moment and excitation energies ( $\omega_f$ ) produced by GAMESS, the TPA cross section is calculated in macroscopic units (GM) and properly scaled to match single-beam experiments by:

$$\sigma^{\text{TPA}} = \frac{4\pi^2 a_0^5 \alpha \omega^2}{c \Gamma} \delta^{\text{TPA}}, \quad (4.3)$$

where  $\alpha$  is the fine structure constant,  $a_0$  is the Bohr radius,  $c$  is the speed of light,  $\omega$  is the photon energy ( $=\frac{\omega_f}{2}$ ) and  $\Gamma$  is the broadening factor derived from a Lorentzian



**Figure 4.1:** A RFP chromophore model showing various angles that were varied in the TPA conformational study. The dihedral angle of the acylimine moiety ( $\theta_{\text{acylimine}}$ ) is made by atoms O<sub>1</sub>, C<sub>1</sub>, N<sub>1</sub> and C<sub>2</sub>. The twist angle is the dihedral between atoms C<sub>3</sub>, C<sub>4</sub>, C<sub>5</sub> and C<sub>6</sub> while the tilt angle is the dihedral between atoms C<sub>4</sub>, C<sub>5</sub>, C<sub>6</sub> and N<sub>2</sub>. Model chromophores are generated by replacing the highlighted part with moieties derived from *n*CAAs. A subset of these moieties is given below the parent structure while the full list is given in Figure C1 in Appendix C.

function and chosen to be 0.1 eV for comparison with experiment, as previously employed.<sup>9,12,13</sup>

TPA computations were done using the conductor-like polarizable continuum model (referred to herein as PCM)<sup>120–123</sup> with parameters for water. Gas-phase TPA computations were used for comparison. The B3LYP functional<sup>40</sup> was mostly used for excited-states properties, while CAM-B3LYP<sup>36</sup> was used for comparison, as noted. The basis set 6-31+G(d,p)<sup>114–118</sup> in Cartesian form, i.e., 6 d-functions, was used in all computations.

### 4.3 Results and Discussion

There are some important limitations associated with studying the TPA of isolated FP chromophores. Firstly, the protein environment has a large influence on its TPA through the following factors: (1) a steric factor affecting the conformation of the chromophore, (2) protein-chromophore interactions via residues that are in the vicinity of the chromophore, and (3) the electric field surrounding the chromophore.<sup>4,70</sup> Two previous computational studies of TPA accounted for the environmental effects via polarizable embedding (PE) mixed quantum mechanics and molecular mechanics (QM/MM) approaches.<sup>30,69</sup> Both studies confirmed the enhancement of TPA of the chromophore when embedded in the protein shell. Although all environmental factors affecting the TPA of the chromophore were theoretically considered, the first study on GFP did not quantitatively reproduce the experimental TPA spectrum.<sup>30</sup> Further, List et al. computed (at the CAM-B3LYP/6-31+G(d) level of theory) the TPA cross section for the isolated RFP chromophore of DsRed<sup>70</sup> at its vacuum-optimized geometry (17 GM), its protein-influenced geometry (47 GM), and for the whole protein via PE (106 GM).<sup>69</sup> The PE result was in good agreement with the experimental value (96 GM).<sup>84</sup> Including effective external field effects (due to the external electromagnetic field), however, nullified the enhancement due to the protein environment and caused the computed cross section (30 GM) to fall even below that of the isolated chromophore (at its native protein conformation).<sup>85</sup> From these attempts, it seems that the computational toolbox needs more validation before it can reliably predict the quantitative effect of the protein environment on the TPA cross section of the chromophore. In this chapter, we span some of the conformational space for the chromophore, based on our previous classical dynamics study<sup>75</sup> and data from crystal structures (Table C1 in Appendix C). Electrostatic effects from the protein shell or close-by residues are not considered. Secondly, we limit our study to vertical excitations and thus temperature and non-Condon effects are not considered. The present TD-DFT scan highlights promising chromophores with large intrinsic TPA cross sections that might stimulate experimental interest.

The oxygen of the acylimine moiety is below the plane of the molecule in the optimized structures (coordinates given in Table C2 in Appendix C). In this conformation, the acylimine dihedral angle ( $\theta_{\text{acylimine}}$  in Figure 4.1) is less than  $180^\circ$  which is similar to the angles found in the majority of the studied RFP-like crystal structures (Table C1 in Appendix C). We first discuss the OPA and TPA properties for the chromophores at their optimized conformations (Section 4.3.1). We then explore the change of TPA cross section with the acylimine angle for some of the chromophores with relatively large cross sections (Section 4.3.2). Finally, we compute the TPA cross sections for a portion of the accessible conformational space of the chromophore with the largest TPA (Model 21) at fixed acylimine dihedral angles (Section 4.3.3).

### 4.3.1 TPA cross sections

One-photon energies, OPA oscillator strengths, and the TPA cross sections (corresponding to the transition to  $S_1$ , computed with B3LYP) for all chromophores are given in Table 4.1. Other data, including previous results of the GFP-derived chromophores<sup>75</sup> and comparison between the GFP and the RFP-derived ones, are given in Tables C3 and C4 in Appendix C. The comparison between the gas phase and PCM results in Table C5 and Figure C2 in Appendix C demonstrate that the computation is not sensitive to the dielectric constant of the medium. As expected, the extra acylimine moiety results in a red-shift for the absorption of all chromophores (as compared to their GFP-derived counterparts).<sup>75</sup> The average red-shift is 0.446 eV, and Model 5 (a fluoro-derivative) has the maximum shift (0.522 eV). In terms of wavelengths, Model 20 (the Gold FP derivative) has the largest red-shift of 93 nm from 461 nm (2.689 eV)<sup>75</sup> to 554 nm (2.239 eV) corresponding to the GFP-derived and RFP-derived models, respectively.

In general, introducing the acylimine moiety is accompanied by an increase in TPA cross section, as illustrated in Figure 4.2. Within the 2LM approximation, the

sum-over-states (SOS) expression in Equation 4.2 becomes:<sup>160,162</sup>

$$\begin{aligned} S_{\alpha\beta} &= \frac{2}{\omega_1} [\mu_{\beta 01} (\mu_{\alpha 11} - \mu_{\alpha 00}) + \mu_{\alpha 01} (\mu_{\beta 11} - \mu_{\beta 00})] \\ &= \frac{2}{\omega_1} [\mu_{\beta 01} (\Delta\mu_\alpha) + \mu_{\alpha 01} (\Delta\mu_\beta)], \end{aligned} \quad (4.4)$$

where  $\mu_{\alpha mn}$  is the  $\alpha^{\text{th}}$  component of the dipole moment vector from state  $|m\rangle$  to state  $|n\rangle$ ; i.e.,  $\langle m|\mu_\alpha|n\rangle$ , and  $\omega_1$  is the energy gap to the first excited state,  $|1\rangle$ . The difference between the permanent dipole moments of the first excited state and the ground state for the  $\alpha^{\text{th}}$  component is denoted  $\Delta\mu_\alpha$ . As previously derived,<sup>162,163</sup> Equation 4.4 can be inserted in Equation 4.1 and manipulated using the vector nature of the dipole moment elements to give:

$$\delta^{\text{TPA}} = \frac{16}{15} \left( \frac{|\mu_{01}| |\Delta\mu|}{\omega_1} \right)^2 (2\cos^2\theta + 1), \quad (4.5)$$

where  $\theta$  is the angle between the  $\mu_{01}$  and the  $\Delta\mu$  vectors. Within the 2LM, equations 4.3 and 4.5 show that the increase in TPA cross section can be due to larger  $\mu_{01}$ ,  $\Delta\mu_{01}$  or having both vectors more aligned. The square of the transition dipole moment to the excited state is directly related to the corresponding OPA oscillator strength. There was slight variation in OPA oscillator strengths (within  $\pm 14\%$ ) upon the inclusion of the acylimine moiety except for Models 19, 20 and 21 where it increased by 0.24 (44%), 0.12 (33%) and 0.17 (55%), respectively. The average percent change in TPA cross sections is 1227% while that of OPA oscillator strengths is 4%. In 13 models (approximately half of the studied set), the increase in TPA cross section is actually accompanied by a decrease in OPA oscillator strength. In Models 11 and 15, the percent decrease in TPA cross section was approximately 16 times the decrease in OPA oscillator strength. In Model 22, the TPA cross section decreased by 35%, while the OPA oscillator strength increased by 8%. The percent increase of OPA oscillator strengths in Models 19 (44%), 20 (33%) and 21 (55%) is still significantly less than that of the TPA cross sections (228%, 228% and 154%, respectively). It can thus be inferred that the change in TPA cross sections is mainly driven by the change in permanent dipole differences ( $\Delta\mu$ ) and not the change in transition dipole moments. The factors in Equation 4.5 are explicitly computed in the gas phase at

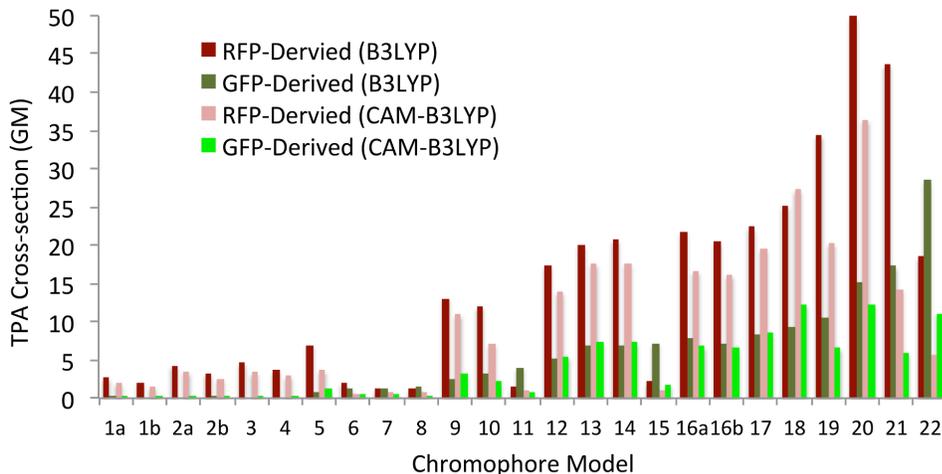
the B3LYP and CAM-B3LYP/6-31+G(d,p) levels of theory and the complete set of results is reported in Table C6 and Figure C3 in Appendix C. A comparison between the TPA cross sections obtained via quadratic response and the truncated SOS approach (within the 2LM) using the B3LYP functional is shown in Figure 4.3. The 2LM successfully captures the trend of relative TPA cross sections while there are deviations in terms of absolute values. Interestingly, the CAM-B3LYP results (see Figure C3 in Appendix C) show a better agreement between the quadratic response and 2LM absolute values (with the exception of Model 21). The interplay between the three components affecting the 2LM cross sections is shown in Figure C3. While the transition dipole moment is nearly unchanged for most chromophores, the models with larger  $\Delta\mu$  and with more aligned  $\mu_{01}$  and  $\Delta\mu$  vectors yield the largest cross sections. Models 16a through 22 have nearly aligned  $\mu_{01}$  and  $\Delta\mu$  vectors ( $\cos^2\theta > 0.8$ ). Model 20 and, according to the B3LYP results, Model 21 get their special TPA enhancement due to the relatively large magnitude of  $\Delta\mu$ . Deviations from the 2LM, especially for Model 21, strongly suggest the involvement of channel interference<sup>162</sup> via higher excited states to the overall TPA cross section of the relevant models; the detailed analysis of the photophysics within N-level ( $N > 2$ ) models is beyond the focus of this chapter.

The CAM-B3LYP functional has been shown to significantly underestimate the TPA strength of neutral chromophores as compared to the more accurate CC2 method.<sup>16,26</sup> A previous benchmark showed that the B3LYP functional seemed to (slightly) underestimate the TPA even more than CAM-B3LYP (a difference of less than 3 GM, according to the properly scaled values).<sup>13,47</sup> As there is a growing interest to benchmark the performance of CAM-B3LYP,<sup>26</sup> we computed the same properties for the same set of GFP- and RFP-derived chromophores using the CAM-B3LYP functional. Full data can be found in Table C7 in Appendix C. A comparison between the TPA results computed with CAM-B3LYP/6-31G+(d,p) in PCM ( $\text{H}_2\text{O}$ ) for the GFP-derived and RFP-derived chromophores is shown in Figure 4.2 together with the analogous B3LYP computations; all TD-DFT results are determined at the PBE0/6-31+G(d,p) optimized geometries (see Table C8 in Appendix C). Similar to the B3LYP results, the

**Table 4.1:** *One-photon excitation energies [in eV], OPA oscillator strengths and TPA cross sections [in GM] for the transition to  $S_1$  as determined at the B3LYP/6-31G+(d,p) level of theory and PCM with parameters for  $H_2O$ . Between brackets is the difference between the property computed for the RFP-derived chromophore and that previously computed for the corresponding GFP-derived chromophore.*<sup>75</sup>

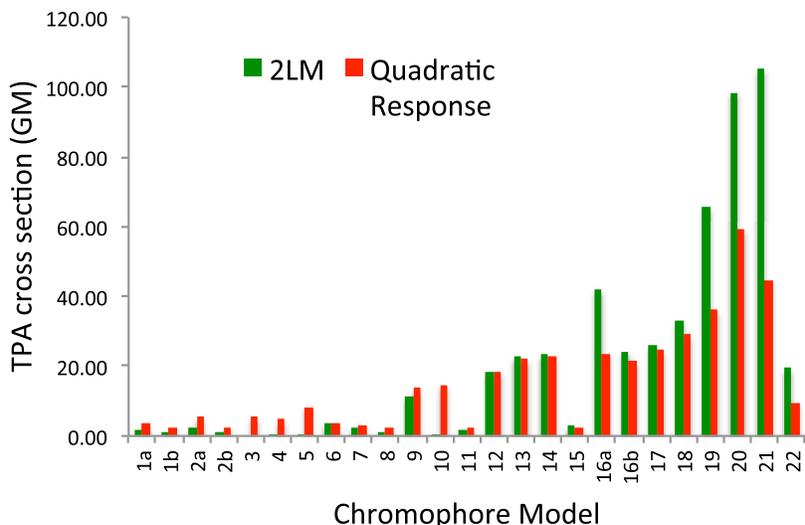
Model	Energy [eV]	OS	TPA [GM]
1a	2.977 (-0.478)	0.64 (0.00)	3 (2)
1b	2.987 (-0.481)	0.61 (-0.07)	2 (2)
2a	2.968 (-0.505)	0.67 (0.02)	4 (4)
2b	2.995 (-0.457)	0.63 (0.08)	3 (3)
3	2.968 (-0.495)	0.68 (-0.03)	5 (4)
4	2.923 (-0.464)	0.75 (-0.03)	4 (4)
5	2.945 (-0.522)	0.70 (0.00)	7 (6)
6	2.881 (-0.316)	0.83 (-0.08)	2 (0)
7	2.819 (-0.381)	0.88 (-0.05)	1 (0)
8	2.883 (-0.421)	0.75 (-0.05)	1 (0)
9	2.897 (-0.506)	0.82 (-0.03)	13 (11)
10	2.898 (-0.471)	0.81 (-0.04)	12 (8)
11	2.859 (-0.397)	0.78 (-0.03)	1 (-3)
12	2.772 (-0.438)	0.92 (-0.07)	17 (12)
13	2.785 (-0.504)	0.83 (-0.02)	20 (13)
14	2.801 (-0.509)	0.85 (-0.04)	21 (14)
15	2.840 (-0.382)	0.80 (-0.04)	2 (-5)
16a	2.733 (-0.485)	0.64 (0.06)	22 (14)
16b	2.783 (-0.477)	0.77 (0.06)	20 (13)
17	2.775 (-0.509)	0.86 (-0.04)	22 (14)
18	2.630 (-0.507)	0.91 (-0.02)	25 (16)
19	2.732 (-0.406)	0.80 (0.24)	34 (24)
20	2.239 (-0.450)	0.48 (0.12)	50 (35)
21	2.654 (-0.331)	0.46 (0.17)	44 (26)
22	2.710 (-0.255)	0.69 (0.05)	19 (-10)

RFP-derived chromophores have larger TPA cross sections than their GFP-derived counterparts (with some exceptions). The distribution of TPA cross sections with CAM-B3LYP is of similar qualitative nature to that of B3LYP and both functionals agree that the RFP-derived form of Model 20 has the largest intrinsic TPA strength (at the optimized geometry). Contrary to B3LYP,<sup>75</sup> CAM-B3LYP does not predict that the nitro derivative (Model 22) has the largest TPA cross section amongst



**Figure 4.2:** TPA cross sections of the GFP- and RFP-derived chromophores for the transition to  $S_1$  as determined at the B3LYP/6-31G+(d,p) and the CAM-B3LYP/6-31G+(d,p) levels of theory with PCM ( $H_2O$ ). For B3LYP, cross sections of the RFP- and GFP-derived chromophores are given in Table 4.1 and taken from our previous work,<sup>75</sup> respectively. The data for cross sections computed with CAM-B3LYP are given in Table C7 in Appendix C.

the GFP-derived models. Oposing the trend previously observed on natural chromophores,<sup>13</sup> the TPA cross sections computed by CAM-B3LYP are 1 to 3 times smaller than those computed by B3LYP (corresponding to a difference of 1 to 29 GM). The largest variation in TPA cross sections is in Models 19, 20 and 21. To investigate the extent of charge-transfer in the studied excitations, the overlap quantity  $\Lambda$  is reported<sup>124</sup> for the  $S_0$  to  $S_1$  transitions of the RFP-derived models (see Table C9 in Appendix C). The  $\Lambda$  parameter is a non-unique diagnostic value (ranging from 0 to 1) that measures the degree of overlap between virtual and occupied orbitals for a given excitation where small values ( $<0.4$ ) indicate evidence for long-range excitations. All chromophore models have a  $\Lambda$ -diagnostic value of greater than 0.6 indicating the absence of significant charge-transfer in the transitions to the first excited state.



**Figure 4.3:** TPA cross sections calculated via the components of the 2-level model (2LM) and via quadratic response at the B3LYP/6-31+G(d,p) level of theory in the gas phase. The contributing components to the 2LM expression are given in Table C6 in Appendix C.

### 4.3.2 TPA change with acylimine rotation

As various crystal structures for RFPs are examined, we note that the acylimine carbonyl always assumes an out-of-plane conformation with respect to the rest of the chromophore (See Table C1 in Appendix C). In this set of chromophores (from crystal structures),  $\theta_{\text{acylimine}}$  ranges from  $129^\circ$  to  $359^\circ$ . In addition, the carbon and nitrogen of the acylimine extension are usually off-plane, as can be shown by studying the dihedral angles  $\theta_1$  (made by  $C_1$ ,  $N_1$ ,  $C_2$  and  $C_7$ ) and  $\theta_2$  (made by  $N_1$ ,  $C_2$ ,  $C_7$  and  $N_2$ ), respectively (see Figure 4.1). As given in Table C1 in Appendix C,  $\theta_1$  ranges from  $91^\circ$  to  $284^\circ$  while  $\theta_2$  ranges from  $165^\circ$  to  $256^\circ$ . For simplicity, we limit our conformational study of the extended acylimine to the rotation of  $\theta_{\text{acylimine}}$  while keeping  $\theta_1$  and  $\theta_2$  at the computationally optimized (near-planar) values. The RFP-derived chromophores with relatively large TPA cross sections ( $\geq 19$  GM) as computed with B3LYP were considered in the acylimine rotation scan. The role of out-of-plane conformation of the acylimine moiety in increasing the TPA cross section of DsRed<sup>70</sup> was previously

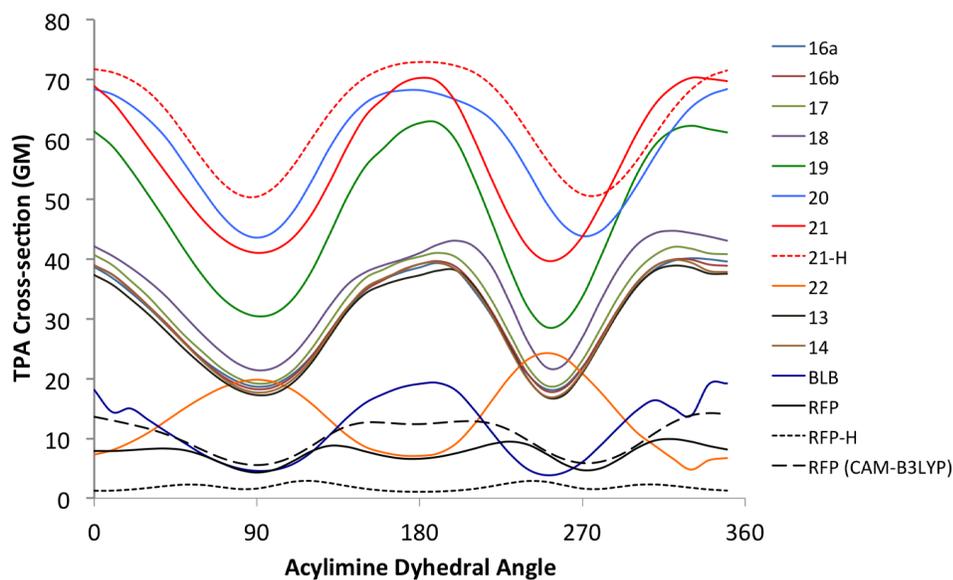
discussed.<sup>13,69</sup> To the best of our knowledge, no study considered a full rotation of the acylimine moiety of the RFP chromophore, so we also included chromophore models of DsRed (See Figure 4.1) and mBlueberry1.<sup>74</sup> The mBlueberry1 chromophore (BLB) is similar to that of DsRed, but with no substitution on the benzene ring, i.e., no O<sup>-</sup>. The scan considered a full rotation of the carbonyl moiety where the excited-state properties were computed at 10° intervals. Full TD-DFT results for the  $S_0$  to  $S_1$  excitations are given in Table C10 in Appendix C. As the acylimine is rotated, the rest of the structure is not re-optimized. As shown in Figure 4.4, the TPA cross sections follow an oscillatory behaviour with the rotation of the carbonyl of the acylimine moiety. The results given in Table C11 and plotted in Figure 4.4 are for the TPA cross sections associated with  $S_0$  to  $S_1$  transition with some exceptions: in Model 22 and BLB, the first two excited states of the near-planar structures ( $\theta_{\text{acylimine}} = 0^\circ, 10^\circ, 340^\circ$  and  $350^\circ$ ) are nearly-degenerate with the first state being dark while the second is bright (See Table C12 in Appendix C). For the conformer with  $\theta_{\text{acylimine}} = 330^\circ$ , the magnitude of the OPA oscillator strength is divided among the two states, with the second state having a slightly larger value. For all these near-planar cases for Model 22 and BLB, we included the TPA cross section of the  $S_0$  to  $S_2$  transition in Figure 4.4. With the exception of Model 22 (NO<sub>2</sub>-substituted), peak TPA cross section is achieved when the acylimine is coplanar with the molecule. The difference between the maximum and minimum TPA cross sections is more than 50% of the maximum value in nearly all chromophores. Interestingly, the RFP model follows neither trends. Having relatively weak TPA, it is difficult to discern the relationship between the acylimine angle and TPA at this level of theory. The type of capping (CH<sub>3</sub> vs H) and the variation in bond lengths and angles cause slight variation in cross sections that is more pronounced if the molecule has already low intrinsic TPA. To demonstrate this inadequacy, we repeated the scan using a H-capped model for RFP after optimizing it at the same level of theory adopted for the methyl-capped model. The rotational scan produced a very similar trend to the methyl-capped counterpart, but with a lower cross section (see Figure 4.4). As the chromophore is optimized at a certain acylimine angle and not re-optimized upon its rotation, this introduces

some “noise” in the TPA computation. This “noise” could be ignored only if the chromophore model has a large intrinsic TPA, as can be seen in the other models. As a side note, methyl capping does not always increase the TPA cross section, as demonstrated by another scan done on model 21 with H-capping (see comparison in Figure 4.4). In addition, using CAM-B3LYP for the RFP chromophore reproduces the same B3LYP trend.

The driving force for the trend in TPA cross sections observed with the rotation of the acylimine dihedral of Model 21 was investigated. The components contributing to the TPA cross section within a 2LM were computed at the B3LYP and CAM-B3LYP/6-31+G(d,p) levels of theory in the gas phase for conformers of Model 21 varying  $\theta_{\text{acylimine}}$  from  $0^\circ$  to  $90^\circ$  (with  $10^\circ$  intervals). Cross sections calculated via the truncated SOS approach within a 2LM follow the same oscillatory trend observed with the quadratic response results. Similar to what is discussed in Section 4.3.1, this trend is driven by the variation in  $\Delta\mu$  and not in  $\mu_{01}$ . Interestingly, B3LYP and CAM-B3LYP yield opposite trends for the alignment between the  $\Delta\mu$  and  $\mu_{01}$  vectors. Having the vectors nearly aligned under the studied spectrum of  $\theta_{\text{acylimine}}$ , the overall trend of TPA cross sections is the same whether computed by B3LYP or CAM-B3LYP. Relevant results are given in Table C13 and the significant trends are illustrated in Figure C4 in Appendix C.

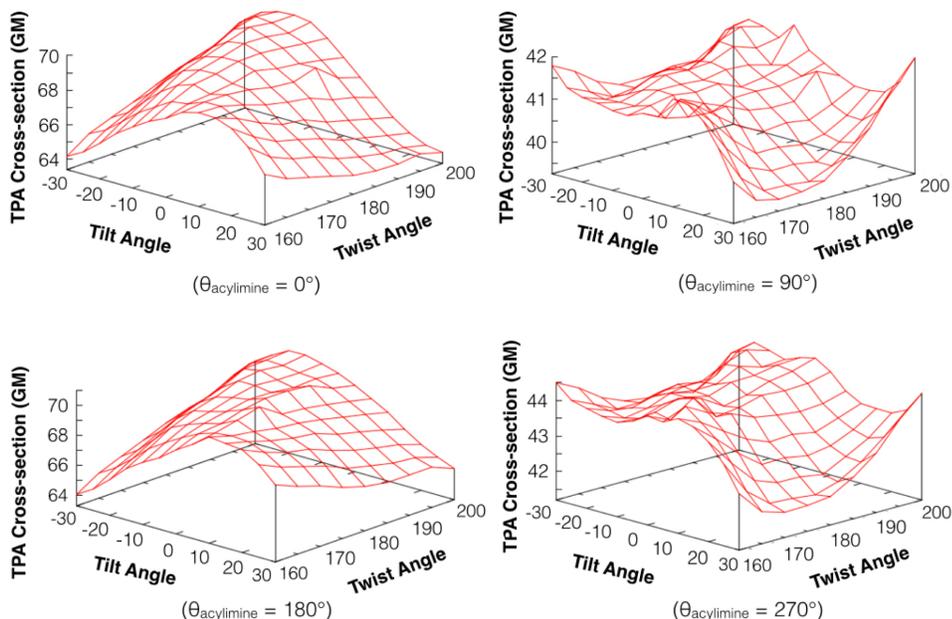
### 4.3.3 TPA change with twist and tilt

Similar to what is observed in the acylimine moiety in the studied set of crystal structures, the rest of the chromophore body is usually distorted from planarity (See Table C1 in Appendix C). This distortion can be represented via the twist and tilt angles between the rings of the chromophore (see Figure 4.1). We computed the TPA of various tilt and twist angles for Model 21 which exhibits the largest intrinsic TPA (see Figure 4.4). Following our previous dynamics results,<sup>75</sup> the conformational flexibility of the chromophore is expected to range from  $-30^\circ$  to  $+30^\circ$  and from  $160^\circ$  to  $200^\circ$  for the twist and tilt angles, respectively. These angle ranges mirror those observed in the examined crystal structures for the proteins containing the RFP chromophore



**Figure 4.4:** Variation of TPA with rotation of the dihedral angle of the acylimine moiety ( $\theta_{\text{acylimine}}$  in Figure 4.1). The TPA values are for the transition to  $S_1$  (with some exceptions noted in the main text) as determined at the B3LYP/6-31G+(d,p) level of theory and PCM with parameters for  $H_2O$ . The dashed curve corresponds to a scan of the RFP model using CAM-B3LYP at the same level of theory. Only the two dotted curves were generated using H-capping while all other computations were done on  $CH_3$ -capped chromophores.

(See Table C1 in Appendix C). We varied the tilt and twist angles by  $5^\circ$  increments within these ranges at four fixed acylimine angles:  $0^\circ$ ,  $90^\circ$ ,  $180^\circ$  and  $270^\circ$ . Full data is available in Table C14 in Appendix C. The change in TPA cross section due to the variation of tilt and twist angles is less than 10% of the maximum TPA at a given acylimine dihedral (Figure 4.5). An interesting observation is that the OPA oscillator strength is maximum when the two rings are coplanar (regardless the value of  $\theta_{\text{acylimine}}$ ) and this maximum value is maintained as the tilt angle is varied from  $0^\circ$  to  $30^\circ$  or to  $-30^\circ$  as long as it is accompanied with a variation in the twist angle from  $180^\circ$  to  $160^\circ$  or to  $200^\circ$ , respectively (see Figure C5 in Appendix C). Nevertheless, energies and OPA oscillator strengths still experience relatively little total variation



**Figure 4.5:** Variation of TPA cross sections with tilt and twist angles (see Figure 4.1) for Model 21 at fixed  $\theta_{\text{acylimine}}$  of  $0^\circ$ ,  $90^\circ$ ,  $180^\circ$  and  $270^\circ$ . The TPA values are for the transition to  $S_1$  as determined at the B3LYP/6-31G+(d,p) level of theory and PCM with parameters for  $H_2O$ .

(within 0.1 eV and 0.17, respectively). These results, together with the results from Section 4.3.2, strongly suggest that the acylimine orientation is the strongest driver of change in the intrinsic TPA cross sections for RFP-like chromophores.

## 4.4 Conclusion

In this chapter, we investigated the OPA and TPA properties for 22 RFP-like chromophores made from non-canonical amino acids. Interestingly, the extra acylimine moiety significantly alters the TPA cross section of the chromophores along with the expected redshift in OPA energies. In terms of magnitude, the RFP-derived chromophores are determined to have larger TPA cross sections than their GFP-derived counterparts that were previously computed.<sup>75</sup> Computing the TPA cross sections with the B3LYP and CAM-B3LYP functionals yields similar trends but with some

subtle quantitative differences. Results for both functionals at the optimized geometries agree that the RFP-derived Model 20 has the largest intrinsic TPA cross section. The trend of TPA cross sections as computed with PCM is mirrored by the gas-phase computation at the same level of theory. In addition, the simplified 2LM also reproduces the same trend computed via quadratic response. The 2LM analysis shows that the amplification of TPA cross section in Model 20 is due to its relatively large  $\Delta\mu$ . Further, we studied the variation of TPA with the dihedral angle of the acylimine moiety ( $\theta_{\text{acylimine}}$ ) for the models with relatively large TPA cross sections and the natural RFP-like chromophores. We noticed a large variation of TPA with the rotation of the acylimine dihedral assuming an oscillatory trend that peaks at the planar conformation for all models except the nitro-substituted one (Model 22). The trend with  $\theta_{\text{acylimine}}$  in the RFP chromophore is difficult to follow due to its low intrinsic TPA cross section. Though larger than all other models, the TPA cross sections determined for Models 21 and 20 are significantly lower at their optimized geometries (44 GM and 50 GM, respectively) than their planar conformations (70 GM and 68 GM, respectively). Studying the TPA trend for Model 21 with varying  $\theta_{\text{acylimine}}$  using the 2LM reveals that the trend of cross sections is, again, driven by the variation in  $\Delta\mu$ . We further computed the excited state properties for Model 21 with various tilt and twist angles spanning the most accessible conformational space. What is clear is that the acylimine angle has a much stronger effect on the TPA of the chromophore than its tilt and twist angles. Model 20 refers to a GdFP chromophore (but with the additional acylimine moiety) while Model 21 has a quinoline-like structure that has not yet been experimentally incorporated in a FP. Either model appears to be very promising in terms of intrinsic TPA cross section that is more than 7 times that of the native RFP chromophore. In this chapter, we provide a rational basis to the experimental synthesis of FPs that are expected to have improved TPA cross sections.

# Chapter 5

## Conclusions

### 5.1 Summary of Thesis Research

In this thesis, I highlight some fluorescent protein chromophores (FPs) that can be made from non-canonical amino acids (ncAAs) and are expected to have largely enhanced two-photon absorption (TPA).

I started by validating the computational technique that was used in screening possible chromophore structures. I compared the TD-DFT results using four functionals on a given set of natural chromophores to a higher wave-function-based method (CC2) and to averaged experimental data, see Chapter 2. Comparison to experimental data was for a qualitative purpose and not strict benchmarking, as these are for the whole proteins while my computations were for isolated amino acids. The comparison shows that TD-DFT with B3LYP and PBE0 underestimate the absolute TPA cross sections as compared to CC2 and experimental values (see Table 2.1). However, they can be used in a semi-quantitative fashion to study the trends of transition to the first excited-state across various structures. Studying the RFP chromophore at different conformations of the acylimine moiety and different levels of theory showed that B3LYP also captures the variation of TPA cross sections with conformation (see Table 2.2). Further comparison between B3LYP and CAM-B3LYP was done in a later stage of my thesis, see Chapter 4. The presence of low-lying states with long-range (charge-transfer) excitation character in transitions to higher-excited states makes their computation with B3LYP and PBE0 less reliable.

I further used TD-DFT at the B3LYP/6-31+G(d,p) level of theory to screen twenty-two possible chromophores that can be formed upon replacing a precursory amino acid of the GFP chromophore with a ncAA, see Chapter 3. A proposed chromophore with a nitro substituent was found to have a large TPA cross section (29 GM). This is more than 7 times that of the TPA cross section of the native GFP chromophore computed at the same level of theory (see Table 2.1 and Table 3.1). To motivate the feasibility of having a mature protein with such mutation, a classical MD study was performed on a nitro-modified FP and its parent EGFP structure. The stability of the modified protein and conformational flexibility of the chromophore were determined over a 99-ns trajectory. Studying the RMSD and RMSF of the nitro-modified and native EGFP models demonstrate that the modified protein is comparably stable (see Figure 3.3). Theoretical results show that the large cross section of the nitro-modified chromophore is primarily due to the difference between its permanent dipole moments of the excited and ground states (see Table 3.2). This large difference is maintained through the various conformations assumed by the chromophore in the protein cavity (see Figure 3.6). This finding is very interesting as previous experimental work on RFPs demonstrated that the sensitivity of the chromophore to the surrounding electrostatic environment is due to the dependence of the TPA cross section on this difference between permanent dipoles. Due to the protein environment, some RFPs reach an experimental TPA cross section of 139 GM which is 27-times the computed intrinsic TPA cross section (5 GM) at the same level theory. Following this line of reasoning, the nitro-modified chromophore is an excellent candidate for future experimental investigations for its large intrinsic TPA cross section which is amenable to further enhancement via the protein environment.

I further screened a similar set of 22 modified chromophores bearing an acylimine extension that characterizes the majority of RFPs, see Chapter 4. Both B3LYP and CAM-B3LYP functionals were used in the screening of these RFP-based chromophores. The previously studied GFP-derived ones were recomputed with CAM-B3LYP to complete the inter-functional comparison. Computing TPA cross sections with B3LYP and CAM-B3LYP yield similar overall trends (see Figure 2). Results

using both functionals agree that the RFP-derived model of the gold FP (GdFP) chromophore has the largest intrinsic TPA cross section. According to the B3LYP computation, the GdFP chromophore has an intrinsic cross section of 50 GM. One chromophore bearing a hydroxyquinoline ring is also shown to be comparable to that of the GdFP in terms of TPA cross section (44 GM with B3LYP). With an intrinsic TPA ability that is more than 7 times that of the native RFP chromophore, the GdFP- and the hydroxyquinoline-derived chromophores are very promising candidates for future experimental investigation. TPA was further computed for selected chromophores following conformational changes: variation of the dihedral angle of the acylimine moiety and the tilt and twist angles between the rings of the chromophore. The TPA cross section assumed an oscillatory trend with the rotation of the acylimine dihedral (see Figure 3). Excluding the nitro-modified RFP-based chromophore, the TPA is maximized in the planar conformation for all models. The conformational study on the hydroxyquinoline-derived chromophore shows that the acylimine angle has a much stronger effect on TPA than its tilt and twist angles (see Figure 4).

While the TPA of many of the scanned ncAA-modified chromophores is equivalent to their natural counterparts, a number of candidates that have significantly larger TPA cross sections is presented. Their large intrinsic TPA cross section is confirmed in various conformations that FP chromophores are likely to have with the protein shell. this work will hopefully guide and motivate future experimental synthesis of new FPs with enhanced TPA.

## 5.2 Future Perspective

All of the TPA computations in this work considered isolated chromophore structures and the steric effect of the protein environment. Nevertheless, the conformational freedom of the RFP-like chromophores within the protein was not fully explored. The flexibility between the rings of a RFP-based chromophore was studied based on data from crystal structures (which provide a static picture) or from a MD trajectory that was actually done on a GFP-like protein. The rotation of only one dihedral angle

within the acylimine extension (between the carbonyl and imine) was studied. The available RFP-like crystal structures reveal variation of other angles associated with this extension. A straight-forward MD study of the available RFP crystal structures can enhance our understanding of the conformational space available to the chromophore within the FP barrel. Sampling of the MD trajectory followed by TPA computation can be done for the chromophore in different orientations.

To have a complete picture on the performance of candidate chromophores, other protein effects need to be considered. To achieve this, two components are required: a reliable computational method and a reliable protein structure. Clearly, TPA computation with TD-DFT needs a more rigorous benchmarking against experimental values for fluorescent chromophores of comparable size to FP ones. Experimental data should be carefully used in comparisons due to the variations in experimental measurements of TPA for a given fluorophore. On the other hand, further computational work can include the use of QM approaches including other nearby residues or QM/MM methods to compute the TPA cross sections of available RFP homologues. It is essential to verify that either of these approaches can capture the proper trend of the TPA variation for RFPs that have the same chromophore. Once this is achieved, the expanded QM or QM/MM methods can be used to predict the TPA cross section of computationally altered FPs incorporating ncAAs. A good starting point can be the GdFP, as a crystal structure of its (GFP-based) structure is available.

So far, TD-DFT has been used in the computations for its efficiency and comparable accuracy to other wave-function-based methods. Nevertheless, much faster methods are to be implemented for computing the TPA of a whole FP. That is why QM/MM is usually the method of choice in this case. Another alternative to QM/MM would be using a (reliable) semi-empirical method to compute the property for a given FP. In ongoing research (unpublished), we computed the one-photon absorption of a series of FP chromophores using the TD-density functional tight binding (TD-DFTB) method; a much faster alternative to TD-DFT. The enhanced speed and comparable results to TD-DFT make TD-DFTB very promising to be used in the computation of larger systems; i.e., a much larger QM region in QM/MM simulations or, perhaps,

even a full FP. Being an unexplored avenue, the use of TD-DFTB to compute OPA and TPA of FP chromophores and/or full FPs needs to be validated via comparison to other available computational methods and to experimental data.

# Bibliography

- [1] Göppert-Mayer, M. *Ann. d. Phys.* **1931**, *9*, 273-295.
- [2] Kaiser, W.; Garrett, C. G. B. *Phys. Rev. Lett.* **1961**, *7*, 229-231.
- [3] Denk, W.; Strickler, J.; Webb, W. *Science* **1990**, *248*, 73-76.
- [4] Drobizhev, M.; Makarov, N. S.; Tillo, S. E.; Hughes, T. E.; Rebane, A. *Nat. Methods* **2011**, *8*, 393-399.
- [5] Pawlicki, M.; Collins, H.; Denning, R.; Anderson, H. *Angew. Chem., Int. Ed.* **2009**, *48*, 3244-3266.
- [6] Ciuciu, A. I.; Firmansyah, D.; Hugues, V.; Blanchard-Desce, M.; Gryko, D. T.; Flamigni, L. *J. Mater. Chem. C* **2014**, *2*, 4552-4565.
- [7] Morales, A. R.; Frazer, A.; Woodward, A. W.; Ahn-White, H.-Y.; Fonari, A.; Tongwa, P.; Timofeeva, T.; Belfield, K. D. *J. Org. Chem.* **2013**, *78*, 1014-1025..
- [8] Campbell, R. E.; Davidson, M. W. *Molecular Imaging with Reporter Genes*; Cambridge University Press, 2010; pp 3-40, Cambridge Books Online.
- [9] Cronstrand, P.; Luo, Y.; Ågren, H. *Adv. Quantum Chem.* **2005**, *50*, 1.
- [10] Yang, G.; Qin, C.; Su, Z.; Dong, S. *Journal of Molecular Structure: THEOCHEM* **2005**, *726*, 61 - 65.
- [11] Olsen, J.; Jørgensen, P. *J. Chem. Phys.* **1985**, *82*, 3235-3264.
- [12] Nifosì, R.; Luo, Y. *J. Phys. Chem. B.* **2007**, *111*, 14043-14050.

- [13] Salem, M. A.; Brown, A. *J. Chem. Theory Comput.* **2014**, *10*, 3260–3269.
- [14] Friese, D. H.; Hattig, C.; Ruud, K. *Phys. Chem. Chem. Phys.* **2012**, *14*, 1175–1184.
- [15] Silva, D. L.; Krawczyk, P.; Bartkowiak, W.; Mendonça, C. R. *J. Chem. Phys.* **2009**, *131*, 244516.
- [16] Beerepoot, M. T. P.; Friese, D. H.; List, N. H.; Kongsted, J.; Ruud, K. *Phys. Chem. Chem. Phys.* **2015**, *17*, 19306–19314.
- [17] Meath, W. J.; Jagatap, B. N. *J. Chem. Phys.* **2013**, *139*, 144104.
- [18] Meath, W. J.; Jagatap, B. N. *J. Phys. B* **2011**, *44*, 205401.
- [19] Meath, W. J.; Jagatap, B. N.; Kondo, A. E. *J. Phys. B* **2006**, *39*, S605–S620.
- [20] Runge, E.; Gross, E. K. U. *Phys. Rev. Lett.* **1984**, *52*, 997–1000.
- [21] Salek, P.; Vahtras, O.; Guo, J.; Luo, Y.; Helgaker, T.; Ågren, H. *Chem. Phys. Lett.* **2003**, *374*, 446–452.
- [22] Frediani, L.; Rinkevicius, Z.; Ågren, H. *J. Chem. Phys.* **2005**, *122*, 244104.
- [23] Tretiak, S.; Chernyak, V. *J. Chem. Phys.* **2003**, *119*, 8809–8823.
- [24] Yuan, L.; Lin, W.; Chen, H.; Zhu, S.; He, L. *Angew. Chem., Int. Ed.* **2013**, *52*, 10018–10022.
- [25] Terenziani, F.; Katan, C.; Badaeva, E.; Tretiak, S.; Blanchard-Desce, M. *Adv. Mater.* **2008**, *20*, 4641–4678.
- [26] Beerepoot, M. T. P.; Friese, D. H.; Ruud, K. *Phys. Chem. Chem. Phys.* **2014**, *16*, 5958–5964.
- [27] Nifosì, R.; Luo, Y. *J. Phys. Chem. B* **2007**, *111*, 505–507.

- [28] Vivas, M. G.; Silva, D. L.; Misoguti, L.; Zalesny, R.; Bartkowiak, W.; Mendonca, C. R. *J. Phys. Chem. A* **2010**, *114*, 3466–3470.
- [29] Kamarchik, E.; Krylov, A. I. *J. Phys. Chem. Lett.* **2011**, *2*, 488–492.
- [30] Steindal, A. H.; Olsen, J. M. H.; Ruud, K.; Frediani, L.; Kongsted, J. *Phys. Chem. Chem. Phys.* **2012**, *14*, 5440–5451.
- [31] Nayyar, I. H.; Tretiak, S. *J. Phys. Chem. C* **2013**, *117*, 18170–18189.
- [32] Kristensen, K.; Kauczor, J.; Thorvaldsen, A. J.; Jørgensen, P.; Kjærgaard, T.; Rizzo, A. *J. Chem. Phys.* **2011**, *134*, 214104.
- [33] Koch, H.; Christiansen, O.; Jørgensen, P.; Sánchez de Merás, A. M.; Helgaker, T. *J. Chem. Phys.* **1997**, *106*, 1808–1818.
- [34] Christiansen, O.; Koch, H.; Jørgensen, P. *J. Chem. Phys.* **1995**, *103*, 7429–7441.
- [35] Paterson, M. J.; Christiansen, O.; Pawłowski, F.; Jørgensen, P.; Hättig, C.; Helgaker, T.; Salek, P. *J. Chem. Phys.* **2006**, *124*, 054322.
- [36] Yanai, T.; Tew, D. P.; Handy, N. C. *Chem. Phys. Lett.* **2004**, *393*, 51 – 57.
- [37] Wachter, R. M.; Elsliger, M. a.; Kallio, K.; Hanson, G. T.; Remington, S. J. *Structure* **1998**, *6*, 1267–1277.
- [38] Christiansen, O.; Koch, H.; Jørgensen, P. *Chem. Phys. Lett.* **1995**, *243*, 409 – 418.
- [39] Hättig, C.; Christiansen, O.; Jørgensen, P. *J. Chem. Phys.* **1998**, *108*, 8331–8354.
- [40] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- [41] (a) Perdew, J. P.; Burke, K.; Ernzerhoff, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868; (b) Perdew, J. P.; Burke, K.; Ernzerhoff, M. *Phys. Rev. Lett.* **1996**, *78*, 1396.

- [42] Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158–6169.
- [43] Iikura, H.; Tsuneda, T.; Yanai, T.; Hirao, K. *J. Chem. Phys.* **2001**, *115*, 3540–3544.
- [44] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.
- [45] Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.
- [46] Tawada, Y.; Tsuneda, T.; Yanagisawa, S.; Yanai, T.; Hirao, K. *J. Chem. Phys.* **2004**, *120*, 8425–8433.
- [47] Note: In this benchmark, the reported TPA cross-sections were too large by a factor of 4 when given in macroscopic units (GM) due to the use of excitation energy ( $\omega_f$ ) in Equation 3 rather than photon energy ( $\frac{\omega_f}{2}$ ).
- [48] Nanda, K. D.; Krylov, A. I. *J. Chem. Phys.* **2015**, *142*, 064118.
- [49] Stanton, J. F.; Bartlett, R. J. *J. Chem. Phys.* **1993**, *98*, 7029–7039.
- [50] Bartlett, R. J. *Mol Phys* **2010**, *108*, 2905–2920.
- [51] *DALTON2013*, a molecular electronic structure program, 2013. Available online: <http://www.daltonprogram.org>.
- [52] Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347–1363.
- [53] Shao, Y. et al. *Mol. Phys.* **2015**, *113*, 184–215.
- [54] Krylov, A. I.; Gill, P. M. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2013**, *3*, 317–326.
- [55] TURBOMOLE V7.0 2015, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989 - 2007, TURBOMOLE GmbH, since 2007; available from <http://www.turbomole.com>

- [56] Shimomura, O.; Johnson, F. H.; Saiga, Y. *J. Cell. Comp. Physiol.* **1962**, *59*, 223–229.
- [57] Prasher, D. C.; Eckenrode, V. K.; Ward, W. W.; Prendergast, F. G.; Cormier, M. J. *Gene* **1992**, *111*, 229–233.
- [58] Chalfie, M.; Tu, Y.; Euskirchen, G.; Ward, W. W.; Prasher, D. *Science* **1994**, *263*, 802–805.
- [59] Inouye, S.; Tsuji, F. I. *FEBS Lett.* **1994**, *341*, 277–280.
- [60] Ormo, M.; Cubitt, A. B.; Kallio, K.; Gross, L. A.; Tsien, R. Y.; Remington, S. *Science* **1996**, *273*, 1392–1395.
- [61] Yang, F.; Moss, L. G.; Phillips, G. *Nat. Biotechnol.* **1996**, *14*, 1246–1251.
- [62] Nifosì, R.; Tozzini, V. In *Fluorescent Proteins I*; Jung, G., Ed.; Springer Series on Fluorescence; Springer Berlin Heidelberg, 2012; Vol. 11; pp 3–40.
- [63] Shaner, N. C.; Steinbach, P. A.; Tsien, R. Y. *Nat. Methods* **2005**, *2*, 905–909.
- [64] Weissleder, R. *Nat. Biotechnol.* **2001**, *19*, 316–317.
- [65] Oulianov, D.; Tomov, I.; Dvornikov, A.; Rentzepis, P. *Opt. Comm.* **2001**, *191*, 235 – 243.
- [66] Zhang, G.; Gurtu, V.; Kain, S. R. *Biochem. Biophys. Res. Commun.* **1996**, *227*, 707–711.
- [67] Hosoi, H.; Yamaguchi, S.; Mizuno, H.; Miyawaki, A.; Tahara, T. *J. Phys. Chem. B* **2008**, *112*, 2761–2763.
- [68] Kawano, H.; Kogure, T.; Abe, Y.; Mizuno, H.; Miyawaki, A. *Nat. Methods* **2008**, *5*, 373–374.
- [69] List, N. H.; Olsen, J. M. H.; Jensen, H. J. A.; Steindal, A. H.; Kongsted, J. *J. Phys. Chem. Lett.* **2012**, *3*, 3513–3521.

- [70] Gross, L. A.; Baird, G. S.; Hoffman, R. C.; Baldrige, K. K.; Tsien, R. Y. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 11990–11995.
- [71] Kikuchi, A.; Fukumura, E.; Karasawa, S.; Mizuno, H.; Miyawaki, A.; Shiro, Y. *Biochemistry* **2008**, *47*, 11573–11580.
- [72] Shaner, N. C.; Campbell, R. E.; Steinbach, P. A.; Giepmans, B. N. G.; Palmer, A. E.; Tsien, R. Y. *Nat. Biotechnol.* **2004**, *22*, 1567–1572.
- [73] Shu, X.; Shaner, N. C.; Yarbrough, C. A.; Tsien, R. Y.; Remington, S. J. *Biochemistry* **2006**, *45*, 9639–9647.
- [74] Ai, H.-W.; Shaner, N. C.; Cheng, Z.; Tsien, R. Y.; Campbell, R. E. *Biochemistry* **2007**, *46*, 5904–5910.
- [75] Alaraby Salem, M.; Brown, A. *Phys. Chem. Chem. Phys.* **2015**, *17*, 25563–25571.
- [76] Olsen, J. M.; Aidas, K.; Kongsted, J. *J. Chem. Theory Comp.* **2010**, *6*, 3721–3734.
- [77] Note: In this study, the reported TPA cross-sections were too large by a factor of 4 when given in macroscopic units (GM) due to the use of excitation energy ( $\omega_f$ ) in Equation 4 rather than photon energy ( $\frac{\omega_f}{2}$ ).
- [78] Epifanovsky, E.; Polyakov, I.; Grigorenko, B.; Nemukhin, A.; Krylov, A. I. *J. Chem. Theory Comput.* **2009**, *5*, 1895–1906.
- [79] Polyakov, I. V.; Grigorenko, B. L.; Epifanovsky, E. M.; Krylov, A. I.; Nemukhin, A. V. *J. Chem. Theory Comput.* **2010**, *6*, 2377–2387.
- [80] Ai, Y.; Tian, G.; Luo, Y. *Mol. Phys.* **2013**, *111*, 1316–1321.
- [81] Chakrabarti, S.; Ruud, K. *Phys. Chem. Chem. Phys.* **2009**, *11*, 2592–2596.
- [82] Chakrabarti, S.; Ruud, K. *J. Phys. Chem. A* **2009**, *113*, 5485–5488.

- [83] Drobizhev, M.; Callis, P. R.; Nifosì, R.; Wicks, G.; Stoltzfus, C.; Barnett, L.; Hughes, T. E.; Sullivan, P.; Rebane, A. *Sci. Rep.* **2015**, *5*, 13223.
- [84] Drobizhev, M.; Tillo, S.; Makarov, N. S.; Hughes, T. E.; Rebane, A. *J. Phys. Chem. B* **2009**, *113*, 12860–12864.
- [85] List, N. H.; Jensen, H. J. A.; Kongsted, J. *Phys. Chem. Chem. Phys.* **2016**, *18*, 10070–10080.
- [86] Alemán, E. A.; de Silva, C.; Patrick, E. M.; Musier-Forsyth, K.; Rueda, D. *J. Phys. Chem. Lett.* **2014**, *5*, 777–781.
- [87] Song, G.; Mandal, M. *RNA Nanotechnology*; Pan Stanford, 2014; pp 153–184.
- [88] Ha, T. *Biochemistry* **2004**, *43*, 4055–4063.
- [89] Xavier Michalet, S. W.; Jäger, M. *Chem. Rev.* **2006**, *106*, 1785–1813.
- [90] Katilius, E.; Woodbury, N. W. *J. Biomed. Opt.* **2006**, *11*, 044004.
- [91] Pitchiaya, S.; Heinicke, L. A.; Custer, T. C.; Walter, N. G. *Chem. Rev.* **2014**, *114*, 3224–3265.
- [92] Sinkeldam, R. W.; Greco, N. J.; Tor, Y. *Chem. Rev. (Washington, DC, U. S.)* **2010**, *110*, 2579–2619.
- [93] Matsika, S. *Top. Curr. Chem.* **2015**, *355*, 209–243.
- [94] Samanta, P. K.; Pati, S. K. *Phys. Chem. Chem. Phys.* **2015**, *17*, 10053–10058.
- [95] Robert J. Stanley, A. Y., Zhanjia Hou; Hawkins, M. E. *J. Phys. Chem. B* **2005**, *109*, 3690–3695.
- [96] Lane, R. S. K.; Magennis, S. W. *RSC Adv.* **2012**, *2*, 11397–11403.
- [97] Lane, R. S. K.; Jones, R.; Sinkeldam, R. W.; Tor, Y.; Magennis, S. W. *ChemPhysChem* **2014**, *15*, 867–871.

- [98] Gedik, M.; Brown, A. *J. Photochem. Photobiol., A* **2013**, *259*, 25 – 32.
- [99] Samanta, P. K.; Manna, A. K.; Pati, S. K. *J. Phys. Chem. B* **2012**, *116*, 7618–7626.
- [100] Drobizhev, M.; Tillo, S.; Makarov, N. S.; Hughes, T. E.; Rebane, A. *J. Phys. Chem. B* **2009**, *113*, 855–859.
- [101] Blab, G. A.; Lommerse, P. H.; Cognet, L.; Harms, G. S.; Schmidt, T. *Chem. Phys. Lett.* **2001**, *350*, 71–77.
- [102] Drobizhev, M.; Makarov, N. S.; Tillo, S. E.; Hughes, T. E.; Rebane, A. *J. Phys. Chem. B* **2012**, *116*, 1736–1744.
- [103] Drobizhev, M.; Makarov, N. S.; Hughes, T.; Rebane, A. *J. Phys. Chem. B* **2007**, *111*, 14051–14054.
- [104] Marchant, J. S.; Stutzmann, G. E.; Leissring, M. a.; LaFerla, F. M.; Parker, I. *Nat. Biotechnol.* **2001**, *19*, 645–649.
- [105] Tillo, S. E.; Hughes, T. E.; Makarov, N. S.; Rebane, A.; Drobizhev, M. *BMC biotechnol.* **2010**, *10*, 6.
- [106] Paterson, M. J.; Christiansen, O.; Pawłowski, F.; Jørgensen, P.; Hättig, C.; Helgaker, T.; Salek, P. *J. Chem. Phys.* **2006**, *124*, 054322.
- [107] Nayyar, I. H.; Tretiak, S. *J. Phys. Chem. C* **2013**, *117*, 18170–18189.
- [108] Day, P. N.; Nguyen, K. A.; Pachter, R. *J. Phys. Chem. B* **2005**, *109*, 1803–1814.
- [109] Matz, M. V.; Fradkov, A. F.; Labas, Y. A.; Savitsky, A. P.; Zaraisky, A. G.; Markelov, M. L.; Lukyanov, S. A. *Nat. Biotechnol.* **1999**, *17*, 969–973.
- [110] Wachter, R. M.; Watkins, J. L.; Kim, H. *Biochemistry* **2010**, *49*, 7417–7427.
- [111] Wachter, R. M.; King, B. a.; Heim, R.; Kallio, K.; Tsien, R. Y.; Boxer, S. G.; Remington, S. J. *Biochemistry* **1997**, *36*, 9759–9765.

- [112] Heim, R.; Prasher, D. C.; Tsien, R. Y. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 12501–12504.
- [113] Chatteraj, M.; King, B. a.; Bublitz, G. U.; Boxer, S. G. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8362–8367.
- [114] Ditchfield, R.; Hehre, W. J.; Pople, J. A. *J. Chem. Phys.* **1971**, *54*, 724–728.
- [115] Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257–2261.
- [116] Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. *J. Chem. Phys.* **1982**, *77*, 3654–3665.
- [117] Hariharan, P.; Pople, J. *Theor. Chim. Acta* **1973**, *28*, 213–222.
- [118] Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. R. *J. Comput. Chem.* **1983**, *4*, 294–301.
- [119] Casida, M. E. In *Recent Advances in Density Functional Methods, Part I*; Chong, D. P., Ed.; World Scientific: Singapore, 1995; Chapter 5, pp 155–192.
- [120] Barone, V.; Cossi, M. *J. Phys. Chem. A* **1998**, *102*, 1995–2001.
- [121] Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. *J. Comput. Chem.* **2003**, *24*, 669–681.
- [122] Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999–3094.
- [123] Cossi, M.; Barone, V. *J. Chem. Phys.* **2001**, *115*, 4708–4717.
- [124] Peach, M. J. G.; Benfield, P.; Helgaker, T.; Tozer, D. J. *J. Chem. Phys.* **2008**, *128*, 044118.
- [125] Bode, B. M.; Gordon, M. S. *J. Mol. Graphics Modell.* **1998**, *16*, 133–138.
- [126] Subach, O. M.; Gundorov, I. S.; Yoshimura, M.; Subach, F. V.; Zhang, J.; Grünwald, D.; Souslova, E. a.; Chudakov, D. M.; Verkhusha, V. V. *Chem. Biol.* **2008**, *15*, 1116–1124.

- [127] Timerghazin, Q. K.; Carlson, H. J.; Liang, C.; Campbell, R. E.; Brown, A. J. *Phys. Chem. B* **2008**, *112*, 2533–2541.
- [128] Xu, Y. W.; Zhang, J. R.; Deng, Y. M.; Hui, L. K.; Jiang, S. P.; Lian, S. H. *J. Photochem. Photobiol. B* **1987**, *1*, 223–227.
- [129] Sengupta, P.; Balaji, J.; Mukherjee, S.; Philip, R.; Ravindra Kumar, G.; Maiti, S. *Proc. SPIE-Int. Soc. Opt. Eng.* **2001**, *4262*, 336–339.
- [130] Merzlyak, E. M.; Goedhart, J.; Shcherbo, D.; Bulina, M. E.; Shcheglov, A. S.; Fradkov, A. F.; Gaintzeva, A.; Lukyanov, K. A.; Lukyanov, S.; Gadella, T. W. J.; Chudakov, D. M. *Nat. Biotechnol.* **2007**, *4*, 555–557.
- [131] Rizzo, M. A.; Springer, G. H.; Granada, B.; Piston, D. W. *Nat. Biotechnol.* **2004**, *22*, 445–449.
- [132] Ai, H.-W.; Olenych, S. G.; Wong, P.; Davidson, M. W.; Campbell, R. E. *BMC Biol.* **2008**, *6*, 13.
- [133] Battad, J. M.; Traore, D. A. K.; Byres, E.; Rossjohn, J.; Devenish, R. J.; Olsen, S.; Wilce, M. C. J.; Prescott, M. *PLoS ONE* **2012**, *7*, e47331.
- [134] Xu, C.; Zipfel, W.; Shear, J. B.; Williams, R. M.; Webb, W. W. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 10763–10768.
- [135] Zipfel, W. R.; Williams, R. M.; Webb, W. W. *Nat. Biotechnol.* **2003**, *21*, 1369–1377.
- [136] Budisa, N. *Angew. Chem., Int. Ed.* **2004**, *43*, 6426–6463.
- [137] Liu, C. C.; Schultz, P. G. *Ann. Rev. Biochem.* **2010**, *79*, 413–444.
- [138] Johnson, J. A.; Lu, Y. Y.; Deventer, J. A. V.; Tirrell, D. A. *Curr. Op. Chem. Biol.* **2010**, *14*, 774 – 780.
- [139] Hoesl, M. G.; Budisa, N. *Curr. Op. Biotechnol.* **2012**, *23*, 751 – 757, Tissue, cell and pathway engineering.

- [140] Quast, R. B.; Mrusek, D.; Hoffmeister, C.; Sonnabend, A.; Kubick, S. *FEBS Lett.* **2015**, *589*, 1703 – 1712.
- [141] Taki, M.; Hohsaka, T.; Murakami, H.; Taira, K.; Sisido, M. *FEBS Lett.* **2001**, *507*, 35–38.
- [142] Wang, L.; Xie, J.; Deniz, A. A.; Schultz, P. G. *J. Org. Chem.* **2003**, *68*, 174–176.
- [143] Bae, J. H.; Rubini, M.; Jung, G.; Wiegand, G.; Seifert, M. H.; Azim, M.; Kim, J.-S.; Zumbusch, A.; Holak, T. A.; Moroder, L.; Huber, R.; Budisa, N. *J. Mol. Biol.* **2003**, *328*, 1071–1081.
- [144] Pal, P. P.; Bae, J. H.; Azim, M. K.; Hess, P.; Friedrich, R.; Huber, R.; Moroder, L.; Budisa, N. *Biochem.* **2005**, *44*, 3663–3672.
- [145] Kajihara, D.; Hohsaka, T.; Sisido, M. *Protein Eng. Des. Sel.* **2005**, *18*, 273–278.
- [146] Goulding, A.; Shrestha, S.; Dria, K.; Hunt, E.; Deo, S. K. *Protein Eng. Des. Sel.* **2008**, *21*, 101–106.
- [147] Wang, F.; Niu, W.; Guo, J.; Schultz, P. G. *Angew. Chem., Int. Ed.* **2012**, *51*, 10132–10135.
- [148] Padilla, M. S.; Young, D. D. *Bioorg. Medicinal Chem. Lett.* **2015**, *25*, 470–473.
- [149] Chen, S.; Chen, Z.-J.; Ren, W.; Ai, H.-W. *J. Am. Chem. Soc.* **2012**, *134*, 9589–9592.
- [150] Kuhn, S. M.; Rubini, M.; Müller, M. A.; Skerra, A. *J. Am. Chem. Soc.* **2011**, *133*, 3708–3711.
- [151] Word, J.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C. *J. Mol. Biol.* **1999**, *285*, 1735 – 1747.
- [152] Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.

- [153] Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins: Structure, Function, and Bioinformatics* **2006**, *65*, 712–725.
- [154] Reuter, N.; Lin, H.; Thiel, W. *J. Phys. Chem. B* **2002**, *106*, 6310–6321.
- [155] Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. *J. Mol. Graphics Modell.* **2006**, *25*, 247 – 260.
- [156] Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157–1174.
- [157] Vanquelef, E.; Simon, S.; Marquant, G.; Garcia, E.; Klimerak, G.; Delepine, J. C.; Cieplak, P.; Dupradeau, F.-Y. *Nucleic Acids Research* **2011**, *39*, W511–W517.
- [158] Case, D. et al. AMBER 12. 2012; University of California, San Francisco.
- [159] Roe, D. R.; Cheatham, T. E. *J. Chem. Theory Comput.* **2013**, *9*, 3084–3095.
- [160] Salem, M. A.; Gedik, M.; Brown, A. In *Handbook of Computational Chemistry*; Leszczynski, J., Ed.; Springer Netherlands: Dordrecht, 2016; Chapter Two Photon Absorption in Biological Molecules, pp 1–19.
- [161] H.-W. Ai, private communication, Dec 2015.
- [162] Alam, M. M.; Chattopadhyaya, M.; Chakrabarti, S.; Ruud, K. *Acc. Chem. Res.* **2014**, *47*, 1604–1612..
- [163] Alam, M. M.; Chattopadhyaya, M.; Chakrabarti, S. *Phys. Chem. Chem. Phys.* **2012**, *14*, 1156–1165.

# Appendix A

## Appendix to Chapter 2

**Table A1:** *Coordinates of the optimized models at the PBE0/6-31+G(d,p) level of theory in the gas phase.*

<hr/>			
	RFP		
C	-0.62880	-4.71297	0.00000
C	0.33307	-3.75191	0.00000
C	-2.04742	-4.37247	0.00000
C	-2.34924	-2.94281	0.00000
C	0.01819	-2.34807	0.00000
C	-1.37643	-1.99216	0.00000
C	1.06059	-1.42934	0.00000
C	1.05877	-0.02308	0.00000
N	-0.00848	0.81007	0.00000
C	0.48341	2.04666	0.00000
C	2.32141	0.74666	0.00000
N	1.86385	2.05273	0.00000
O	-2.93951	-5.23909	0.00000
H	2.43197	2.88530	0.00000
O	3.49412	0.37753	0.00000
H	-0.37808	-5.77047	0.00000
H	1.38550	-4.03288	0.00000
H	-3.40192	-2.67149	0.00000
H	-1.62665	-0.93567	0.00000
H	2.07120	-1.83984	0.00000
C	-0.30977	3.20936	0.00000
H	-1.39440	3.03297	0.00000
N	0.21064	4.41185	0.00000
C	-0.70660	5.43697	0.00000
H	-1.77824	5.11628	0.00000
O	-0.42618	6.62188	0.00000
<hr/>			
	CFP		
C	-3.24822	-2.39416	0.00000

H	-4.33260	-2.40829	0.00000
C	-1.16183	-3.23519	0.00000
O	-0.23367	-4.02497	0.00000
C	-1.20450	-1.75633	0.00000
N	-2.52240	-3.56144	0.00000
N	-2.53295	-1.31557	0.00000
H	-2.87955	-4.50327	0.00000
C	-0.08636	-0.98426	0.00000
H	0.84962	-1.54069	0.00000
C	0.00061	0.44059	0.00000
C	1.21929	1.22099	0.00000
C	-1.04314	1.35552	0.00000
C	0.84112	2.58223	0.00000
C	2.58190	0.89680	0.00000
N	-0.54014	2.61883	0.00000
C	1.77074	3.62193	0.00000
C	3.51480	1.92388	0.00000
H	2.90624	-0.13999	0.00000
H	-1.09733	3.45718	0.00000
C	3.11437	3.27197	0.00000
H	1.46161	4.66345	0.00000
H	4.57435	1.68557	0.00000
H	3.86853	4.05333	0.00000
H	-2.10536	1.15648	0.00000
O	-6.11448	-1.25468	0.00000
<hr/>			
BLB			
C	5.13290	-0.29788	0.00000
C	4.05807	0.58099	0.00000
C	4.90230	-1.67316	0.00000
C	3.59360	-2.16160	0.00000
C	2.73251	0.10453	0.00000
C	2.51504	-1.28859	0.00000
C	1.66745	1.08002	0.00000
C	0.31782	0.90892	0.00000
N	-0.40064	-0.27708	0.00000
C	-1.65624	0.08250	0.00000
C	-0.62419	2.07062	0.00000
N	-1.86374	1.43885	0.00000
H	5.74018	-2.36457	0.00000
H	-2.76861	1.88715	0.00000
O	-0.41019	3.26559	0.00000
H	6.14831	0.08708	0.00000
H	4.23544	1.65371	0.00000
H	3.41589	-3.23320	0.00000
H	1.49829	-1.66520	0.00000
H	1.97426	2.12603	0.00000
C	-2.74688	-0.86792	0.00000
H	-2.45292	-1.92609	0.00000
N	-3.96224	-0.45587	0.00000
C	-4.93281	-1.47796	0.00000

H	-4.53125	-2.51613	0.00000
<hr/>			
BFP1			
C	-1.98449	-1.88950	0.00000
H	-2.19141	-2.95369	0.00000
C	-2.43283	0.31792	0.00000
O	-3.04161	1.37120	0.00000
C	-0.98712	0.01112	0.00000
N	-2.99691	-0.96475	0.00000
N	-0.79436	-1.37780	0.00000
C	-0.00025	0.94552	0.00000
H	-0.30511	1.98880	0.00000
C	1.39307	0.66757	0.00000
H	-3.98777	-1.14776	0.00000
C	2.50629	1.50352	0.00000
N	3.65395	0.78226	0.00000
C	3.26423	-0.48394	0.00000
N	1.92283	-0.60568	0.00000
H	1.33934	-1.43770	0.00000
H	3.92808	-1.33777	0.00000
H	2.51306	2.58568	0.00000
<hr/>			
BFP2			
C	-2.18320	-1.94702	0.00000
H	-2.49221	-2.98684	0.00000
C	-2.43903	0.28318	0.00000
O	-2.93860	1.39426	0.00000
C	-1.01512	-0.16017	0.00000
N	-3.11869	-0.93138	0.00000
N	-0.95308	-1.55686	0.00000
C	-0.00650	0.74217	0.00000
H	-0.33699	1.78030	0.00000
C	1.41499	0.54078	0.00000
H	4.46262	1.40907	0.00000
H	-4.12176	-1.02255	0.00000
C	2.34856	1.56628	0.00000
N	2.06233	-0.67795	0.00000
N	3.56444	0.95221	0.00000
C	3.33578	-0.39728	0.00000
H	2.23675	2.64047	0.00000
H	4.14271	-1.11766	0.00000
<hr/>			
GFP <sub>N</sub>			
C	0.12607	-3.49151	0.00000
H	-0.54354	-4.34467	0.00000
C	2.08893	-2.39676	0.00000
O	3.28411	-2.16679	0.00000
C	0.90682	-1.49613	0.00000
N	1.49056	-3.65936	0.00000
N	-0.26717	-2.25975	0.00000
H	1.99892	-4.52904	0.00000
C	1.03134	-0.14507	0.00000
H	2.06366	0.20478	0.00000

C	0.01133	0.87739	0.00000
C	0.40983	2.22877	0.00000
C	-1.36941	0.59672	0.00000
C	-0.51399	3.25973	0.00000
H	1.47086	2.46591	0.00000
C	-2.29988	1.62298	0.00000
H	-1.69470	-0.43791	0.00000
C	-1.87822	2.95726	0.00000
H	-0.20429	4.29960	0.00000
H	-3.36320	1.39075	0.00000
O	-2.74694	3.99655	0.00000
H	-3.65241	3.66826	0.00000
<hr/>			
GFP <sub>A</sub>			
C	-0.39733	3.32069	0.00000
C	0.45830	2.25437	0.00000
C	-1.83999	3.14551	0.00000
C	-2.28490	1.75998	0.00000
C	0.00143	0.89959	0.00000
C	-1.41672	0.70408	0.00000
C	0.94854	-0.13905	0.00000
C	0.81235	-1.51951	0.00000
N	-0.37511	-2.25819	0.00000
C	-0.01684	-3.50332	0.00000
C	1.96088	-2.41553	0.00000
N	1.34945	-3.67769	0.00000
O	-2.63786	4.11023	0.00000
H	1.86003	-4.54452	0.00000
H	-0.70020	-4.34528	0.00000
O	3.18290	-2.23045	0.00000
H	-0.02521	4.34266	0.00000
H	1.53496	2.42908	0.00000
H	-3.36055	1.59660	0.00000
H	-1.78851	-0.31669	0.00000
H	1.99439	0.17397	0.00000
<hr/>			
OR <sub>A</sub>			
C	-5.25494	0.02039	-0.11494
C	-4.24175	0.93105	-0.06185
C	-4.99483	-1.41157	-0.16013
C	-3.58507	-1.78412	-0.14737
C	-2.86020	0.54647	-0.04553
C	-2.58340	-0.86094	-0.09325
C	-1.88139	1.54169	0.01441
C	-0.48672	1.47339	0.05487
N	0.30168	0.34993	0.05417
C	1.54638	0.77535	0.10212
C	0.34751	2.68210	0.10933
N	1.63198	2.15046	0.13783
O	-5.90950	-2.25846	-0.20764
H	2.49060	2.67708	0.15500
O	0.06141	3.88019	0.12840

H	-6.29670	0.33096	-0.12542
H	-4.47154	1.99598	-0.02860
H	-3.36554	-2.84854	-0.18495
H	-1.54305	-1.17170	-0.08856
H	-2.24329	2.57089	0.03598
C	2.70548	-0.06610	0.09959
C	4.79098	-0.72330	0.01587
N	3.91074	0.39882	0.12534
C	2.64935	-1.57585	0.03421
O	4.03312	-1.91202	0.13655
H	2.25681	-1.88896	-0.94398
C	1.86826	-2.25175	1.14482
H	2.23791	-1.92546	2.12173
H	0.81060	-1.99177	1.06146
H	1.98496	-3.33735	1.07208
O	5.42928	-0.65371	-1.23658
H	5.93708	-1.46695	-1.33355
H	5.54582	-0.73254	0.81618
<hr/>			
KO <sub>A</sub>			
C	5.39277	-0.11863	0.08704
C	4.40389	0.81892	0.05639
C	5.09509	-1.54420	0.08547
C	3.67629	-1.88003	0.04915
C	3.01281	0.47031	0.02013
C	2.69884	-0.93053	0.01821
C	2.06084	1.49159	-0.01031
C	0.66419	1.45954	-0.05085
N	-0.15182	0.35777	-0.07691
C	-1.38511	0.81594	-0.11489
C	-0.13925	2.69054	-0.07632
N	-1.43685	2.19250	-0.11856
O	5.98727	-2.41521	0.11334
H	-2.28185	2.74052	-0.13520
O	0.17755	3.88084	-0.06676
H	6.44214	0.16397	0.11409
H	4.66070	1.87814	0.05886
H	3.42992	-2.93920	0.04792
H	1.65085	-1.21346	-0.00853
H	2.44887	2.51139	-0.00166
C	-2.55573	-0.01164	-0.14757
C	-4.78449	-0.46354	-0.18761
N	-3.73834	0.49613	-0.23563
C	-2.39281	-1.51196	-0.05160
S	-4.03972	-2.16191	-0.42185
H	-1.65969	-1.87389	-0.77524
H	-2.04830	-1.78968	0.94906
O	-5.44613	-0.32002	1.04788
H	-6.07489	-1.04481	1.13186
H	-5.49100	-0.31113	-1.01443

**Table A2:** *Coordinates of the model for RFP<sub>A</sub> with methyl capping optimized at the PBE0/6-31+G(d,p) level of theory in the gas phase and the corresponding model that was generated by replacing the methyl groups with hydrogen atoms (a non-planar model with H-capping). A similar non-planar BLB model is built by replacing the hydroxyl group on the benzene ring with a hydrogen atom.*

RFP <sub>A</sub> (CH <sub>3</sub> -capping) with co-planar acylimine			
C	-5.40977	0.15799	0.05968
C	-4.31665	0.97182	0.04862
C	-5.28268	-1.29241	0.04480
C	-3.91326	-1.79368	0.01531
C	-2.97609	0.46139	0.01704
C	-2.82982	-0.96666	0.00656
C	-1.91112	1.36417	0.00275
C	-0.52636	1.17007	-0.02172
N	0.15647	-0.01265	-0.02929
C	1.44199	0.28769	-0.03840
C	0.41422	2.28938	-0.02118
N	1.65951	1.65890	-0.03568
O	-6.27139	-2.05291	0.06118
O	0.24183	3.51189	-0.01863
H	-6.41872	0.56307	0.06585
H	-4.44864	2.05380	0.04444
H	-3.79286	-2.87505	0.00942
H	-1.82169	-1.37362	-0.01024
H	-2.17809	2.42210	0.01818
C	2.46489	-0.73065	-0.05701
N	3.70984	-0.40288	0.08555
C	4.75710	-1.28481	-0.05084
O	5.03740	-1.85747	-1.09128
C	1.98998	-2.15043	-0.21078
H	1.05738	-2.29372	0.33679
H	1.78473	-2.36344	-1.26435
RFP <sub>A</sub> (H-capping) with out-of-plane acylimine			
C	4.72554	0.26170	0.06860
C	3.60206	1.03307	0.08040
C	4.65429	-1.19095	-0.00045
C	3.30520	-1.74290	-0.05274
C	2.28205	0.47291	0.03084
C	2.19095	-0.95790	-0.04215
C	1.18311	1.33387	0.05182
C	-0.19303	1.08811	0.01061
N	-0.82930	-0.11739	-0.07329
C	-2.12492	0.13468	-0.09900
C	-1.17559	2.17051	0.03351
N	-2.39517	1.49501	-0.03583
O	5.67166	-1.91263	-0.01882
O	-1.05076	3.39675	0.10428
H	5.71755	0.70318	0.12460
H	3.69175	2.11723	0.14864

H	3.22668	-2.82659	-0.11044
H	1.19945	-1.40181	-0.08760
H	1.40926	2.40008	0.10455
C	-3.10755	-0.91989	-0.17647
N	-4.35801	-0.63049	-0.35040
C	-5.37470	-1.55642	-0.30392
O	-5.67160	-2.19616	0.69188
H	-2.72537	-1.94774	-0.10702
H	-5.96769	-1.62015	-1.25026
H	-3.27727	1.98448	-0.02137
H	2.75272	-2.85485	0.12453
C	5.61207	-1.40930	1.19122
H	5.01525	-1.82646	2.01032
H	6.45912	-2.06780	0.98887
H	5.96338	-0.42665	1.51645
C	2.89133	2.40652	-0.06476
H	3.49391	2.23148	0.82853
H	3.48910	2.13282	-0.93934
H	2.60194	3.45933	-0.12559
<hr/>			
BLB (H-capping) with out-of-plane acylimine			
C	4.72554	0.26170	0.06860
C	3.60206	1.03307	0.08040
C	4.65429	-1.19095	-0.00045
C	3.30520	-1.74290	-0.05274
C	2.28205	0.47291	0.03084
C	2.19095	-0.95790	-0.04215
C	1.18311	1.33387	0.05182
C	-0.19303	1.08811	0.01061
N	-0.82930	-0.11739	-0.07329
C	-2.12492	0.13468	-0.09900
C	-1.17559	2.17051	0.03351
N	-2.39517	1.49501	-0.03583
O	-1.05076	3.39675	0.10428
H	5.71755	0.70318	0.12460
H	3.69175	2.11723	0.14864
H	3.22668	-2.82659	-0.11044
H	1.19945	-1.40181	-0.08760
H	1.40926	2.40008	0.10455
C	-3.10755	-0.91989	-0.17647
N	-4.35801	-0.63049	-0.35040
C	-5.37470	-1.55642	-0.30392
O	-5.67160	-2.19616	0.69188
H	-2.72537	-1.94774	-0.10702
H	-5.96769	-1.62015	-1.25026
H	-3.27727	1.98448	-0.02137
H	5.49694	-1.78663	-0.01158

**Table A3:** *Excitation energies and TPA cross-sections for 8 transitions per model as determined at the TD-DFT/6-31+G(d,p) level of theory in PCM (water). Models are given in Fig. 1 of Chapter 2. (CAM = CAM-B3LYP and LC = LC-BLYP)*

transition	Energy (eV)				TPA Cross-Section (GM)			
	B3LYP	PBE0	CAM	LC	B3LYP	PBE0	CAM	LC
RFP <sub>A</sub>								
1	2.28	2.31	2.32	2.29	1	1	2	3
2	2.89	3.01	3.68	3.83	0	0	0	0
3	3.32	3.47	3.91	3.88	1	1	0	0
4	3.34	3.47	3.96	4.00	0	0	23	283
5	3.42	3.49	4.05	4.14	0	0	0	0
6	3.61	3.76	4.13	4.45	3	3	67	311
7	3.83	3.91	4.27	4.52	23	19	16	53
8	4.47	4.60	4.80	4.77	74	127	1	1
BFP1								
1	3.40	3.46	3.55	3.59	1	1	2	3
2	3.81	3.91	4.25	4.22	0	0	0	0
3	4.42	4.54	4.88	4.98	12	14	30	43
4	4.73	4.93	5.60	5.68	0	0	0	0
5	5.01	5.19	5.69	5.84	2	2	3	1
6	5.42	5.63	5.76	5.92	2	0	1	4
7	5.47	5.69	5.94	5.92	0	2	0	0
BFP2								
1	3.59	3.66	3.79	3.83	1	1	2	3
2	3.79	3.89	4.23	4.18	0	0	0	0
3	4.53	4.67	5.05	5.17	16	18	34	44
4	4.85	5.03	5.56	5.56	0	0	0	0
5	5.18	5.35	5.77	5.86	2	2	1	1
6	5.41	5.65	5.85	6.01	1	0	2	0
7	5.50	5.68	6.01	6.10	0	1	0	5
8	5.58	5.79	6.13	6.28	6	6	6	3
CFP								
1	3.14	3.23	3.37	3.41	8	7	7	8
2	3.71	3.89	4.29	4.24	2	2	0	0
3	3.85	3.94	4.49	4.74	0	0	6	31
4	4.07	4.22	4.71	4.92	32	37	48	20
5	4.37	4.52	4.93	5.12	1	1	21	33
6	4.87	5.00	5.17	5.33	0	0	0	17
7	4.98	5.15	5.45	5.53	24	26	1	1
8	5.07	5.36	5.72	5.74	1	1	56	0
GFP <sub>A</sub>								
1	2.96	3.00	3.02	2.97	1	1	2	4
2	3.56	3.68	4.30	4.41	0	0	0	0
3	4.00	4.11	4.51	4.54	0	0	3	4
4	4.19	4.32	4.55	4.57	2	2	0	0
5	4.23	4.36	4.76	4.86	4	6	24	1
6	4.42	4.58	4.78	4.90	3	4	2	47
7	4.42	4.69	5.03	5.26	2	2	6	11

8	4.74	5.03	5.15	5.27	1	1	1	2
GFP <sub>N</sub>								
1	3.34	3.41	3.56	3.62	4	5	5	6
2	3.74	3.83	4.19	4.14	0	0	0	0
3	4.18	4.32	4.70	4.81	8	11	6	3
4	4.30	4.44	4.82	4.93	30	33	60	71
5	4.84	5.00	5.47	5.64	1	1	2	0
6	5.27	5.42	5.73	5.72	0	0	0	3
7	5.38	5.65	5.80	5.90	1	18	1	1
8	5.47	5.68	6.02	6.13	12	1	70	49
KO <sub>A</sub>								
1	2.45	2.50	2.59	2.58	7	7	8	8
2	3.20	3.32	4.00	4.15	0	0	0	0
3	3.51	3.69	4.17	4.31	17	19	205	339
4	3.62	3.77	4.35	4.41	19	27	0	1
5	3.68	3.82	4.43	4.62	0	0	24	3
6	3.91	4.04	4.51	4.72	19	33	3	289
7	3.99	4.12	4.56	4.78	55	41	32	34
8	4.34	4.47	4.78	4.85	19	32	2	3
OA <sub>A</sub>								
1	2.48	2.53	2.60	2.59	6	6	6	6
2	3.21	3.33	4.00	4.14	0	0	0	0
3	3.65	3.80	4.21	4.35	22	30	189	320
4	3.68	3.81	4.34	4.39	0	0	0	1
5	3.92	4.05	4.47	4.61	9	19	3	3
6	4.00	4.13	4.52	4.76	38	29	0	217
7	4.35	4.47	4.80	4.88	21	29	84	5
8	4.50	4.74	4.82	5.02	67	113	8	167
BLB								
1	2.68	2.77	3.04	3.19	23	24	25	23
2	3.14	3.26	3.67	3.63	0	0	0	0
3	3.24	3.31	3.79	3.81	1	0	0	0
4	3.25	3.43	4.12	4.46	0	1	1	4
5	3.80	3.96	4.48	4.61	58	72	75	0
6	4.16	4.29	4.67	4.66	65	87	0	51
7	4.25	4.35	4.71	5.00	0	0	317	525
8	4.54	4.68	5.09	5.11	0	0	0	0

**Table A4:** *Lambda diagnostic and OPA oscillator strength for 8 transitions per model as determined at the TD-DFT/6-31+G(d,p) level of theory in PCM (water). Models are given in Fig. 1 of Chapter 2. (CAM = CAM-B3LYP and LC = LC-BLYP)*

transition	Lambda Diagnostic (eV)				OPA oscillator strength (GM)			
	B3LYP	PBE0	CAM	LC	B3LYP	PBE0	CAM	LC
RFP <sub>A</sub>								
1	0.72	0.71	0.69	0.67	1.23	1.27	1.34	1.36
2	0.31	0.31	0.30	0.30	0.00	0.00	0.00	0.00
3	0.75	0.67	0.30	0.30	0.02	0.02	0.00	0.00
4	0.28	0.36	0.73	0.71	0.00	0.00	0.34	0.44
5	0.31	0.30	0.29	0.29	0.00	0.00	0.00	0.00
6	0.41	0.42	0.74	0.72	0.01	0.02	0.09	0.01
7	0.72	0.71	0.40	0.40	0.35	0.35	0.01	0.01
8	0.45	0.45	0.34	0.33	0.08	0.08	0.00	0.00
BFP1								
1	0.81	0.80	0.79	0.78	0.71	0.73	0.77	0.76
2	0.40	0.40	0.40	0.40	0.00	0.00	0.00	0.00
3	0.69	0.69	0.68	0.67	0.05	0.05	0.04	0.05
4	0.27	0.27	0.27	0.30	0.00	0.00	0.00	0.00
5	0.51	0.51	0.50	0.27	0.01	0.01	0.00	0.00
6	0.32	0.36	0.29	0.49	0.00	0.00	0.00	0.00
7	0.37	0.31	0.35	0.32	0.00	0.00	0.00	0.00
BFP2								
1	0.81	0.80	0.79	0.78	0.67	0.69	0.74	0.73
2	0.43	0.43	0.43	0.43	0.00	0.00	0.00	0.00
3	0.68	0.67	0.67	0.65	0.06	0.06	0.08	0.11
4	0.34	0.34	0.35	0.35	0.00	0.00	0.00	0.00
5	0.56	0.56	0.30	0.28	0.07	0.07	0.00	0.00
6	0.33	0.38	0.55	0.36	0.00	0.00	0.03	0.01
7	0.38	0.32	0.37	0.53	0.00	0.00	0.00	0.02
8	0.61	0.61	0.55	0.50	0.05	0.05	0.08	0.08
CFP								
1	0.75	0.75	0.73	0.72	0.64	0.67	0.74	0.75
2	0.42	0.42	0.39	0.39	0.04	0.04	0.00	0.00
3	0.40	0.40	0.46	0.55	0.00	0.00	0.05	0.03
4	0.62	0.61	0.59	0.56	0.04	0.02	0.00	0.02
5	0.65	0.65	0.62	0.61	0.09	0.10	0.19	0.21
6	0.59	0.61	0.62	0.53	0.11	0.08	0.02	0.07
7	0.57	0.55	0.29	0.27	0.07	0.10	0.00	0.00
8	0.32	0.31	0.55	0.34	0.00	0.00	0.11	0.00
GFP <sub>A</sub>								
1	0.74	0.74	0.71	0.70	1.05	1.08	1.13	1.14
2	0.33	0.33	0.32	0.32	0.00	0.00	0.00	0.00
3	0.38	0.38	0.53	0.52	0.00	0.00	0.05	0.06
4	0.64	0.58	0.37	0.35	0.02	0.03	0.00	0.00
5	0.64	0.70	0.76	0.24	0.02	0.01	0.01	0.00
6	0.47	0.49	0.26	0.73	0.07	0.08	0.00	0.01

7	0.31	0.28	0.49	0.45	0.01	0.00	0.09	0.09
8	0.28	0.27	0.26	0.26	0.00	0.00	0.00	0.01
<hr/>								
GFP <sub>N</sub>								
1	0.79	0.79	0.78	0.76	0.82	0.84	0.86	0.86
2	0.41	0.41	0.41	0.40	0.00	0.00	0.00	0.00
3	0.65	0.65	0.62	0.61	0.00	0.01	0.01	0.01
4	0.62	0.62	0.66	0.65	0.02	0.02	0.02	0.03
5	0.57	0.57	0.60	0.39	0.10	0.12	0.20	0.00
6	0.39	0.40	0.39	0.61	0.00	0.00	0.00	0.26
7	0.30	0.75	0.29	0.27	0.00	0.04	0.00	0.00
8	0.74	0.31	0.68	0.53	0.05	0.00	0.02	0.04
<hr/>								
KO <sub>A</sub>								
1	0.71	0.71	0.69	0.68	1.05	1.09	1.19	1.21
2	0.31	0.31	0.29	0.48	0.00	0.00	0.00	0.00
3	0.59	0.59	0.62	0.57	0.00	0.00	0.32	0.31
4	0.67	0.67	0.35	0.37	0.17	0.19	0.00	0.00
5	0.33	0.34	0.69	0.44	0.00	0.00	0.01	0.03
6	0.53	0.58	0.43	0.54	0.08	0.10	0.01	0.02
7	0.53	0.48	0.46	0.49	0.12	0.07	0.01	0.00
8	0.49	0.49	0.25	0.25	0.09	0.10	0.00	0.00
<hr/>								
OA <sub>A</sub>								
1	0.71	0.71	0.69	0.68	1.04	1.08	1.18	1.20
2	0.31	0.31	0.30	0.29	0.00	0.00	0.00	0.00
3	0.71	0.70	0.63	0.59	0.11	0.13	0.30	0.29
4	0.34	0.35	0.35	0.35	0.00	0.00	0.00	0.00
5	0.55	0.63	0.70	0.40	0.08	0.12	0.01	0.03
6	0.59	0.52	0.49	0.72	0.14	0.08	0.02	0.01
7	0.48	0.48	0.46	0.31	0.09	0.10	0.12	0.00
8	0.30	0.30	0.30	0.41	0.00	0.00	0.00	0.13
<hr/>								
BLB								
1	0.74	0.74	0.72	0.71	0.80	0.83	0.95	0.97
2	0.33	0.34	0.35	0.35	0.00	0.00	0.00	0.00
3	0.31	0.36	0.35	0.36	0.02	0.00	0.00	0.00
4	0.36	0.31	0.32	0.35	0.00	0.02	0.02	0.02
5	0.60	0.60	0.61	0.37	0.02	0.03	0.07	0.00
6	0.77	0.77	0.38	0.61	0.28	0.26	0.00	0.08
7	0.40	0.40	0.74	0.71	0.00	0.00	0.19	0.23
8	0.40	0.39	0.37	0.36	0.00	0.00	0.00	0.00

**Table A5:** *Excitation energies, TPA cross-sections and OPA oscillator strengths for 8 transitions per model as determined at the TD-DFT/6-31+G(d,p) level of theory in the gas phase. Models are given in Fig. 1 of Chapter 2. (CAM = CAM-B3LYP)*

transition	Energy (eV)			TPA Cross-Section (GM)			OPA Oscillator Strength		
	B3LYP	PBE0	CAM	B3LYP	PBE0	CAM	B3LYP	PBE0	CAM
<b>RFP<sub>A</sub></b>									
1	2.43	2.46	2.48	2	2	3	1.08	1.12	1.20
2	2.65	2.78	3.41	0	0	0	0.00	0.00	0.00
3	3.18	3.32	3.88	0	0	0	0.00	0.00	0.00
4	3.29	3.44	3.91	3	6	0	0.02	0.02	0.00
5	3.35	3.45	3.99	0	0	36	0.00	0.00	0.11
6	3.71	3.86	4.18	-	5	96	0.01	0.01	0.25
7	3.93	4.02	4.36	-	120	11	0.28	0.29	0.01
8	4.32	4.46	4.60	-	56	34	0.06	0.06	0.00
<b>BFP1</b>									
1	3.55	3.61	3.70	1	1	3	0.62	0.64	0.68
2	3.67	3.78	4.13	0	0	0	0.00	0.00	0.00
3	4.40	4.60	5.02	0	0	35	0.00	0.00	0.02
4	4.53	4.66	5.26	13	15	0	0.03	0.03	0.00
5	4.93	5.10	5.59	2	2	3	0.01	0.01	0.00
6	5.11	5.42	5.60	1	2	2	0.00	0.00	0.00
7	5.46	5.62	5.88	0	0	0	0.00	0.00	0.00
8	5.49	5.78	5.99	1	20	1	0.00	0.04	0.00
<b>BFP2</b>									
1	3.52	3.63	3.96	0	0	0	0.00	0.00	0.00
2	3.74	3.82	3.96	1	1	1	0.53	0.55	0.61
3	4.60	4.74	5.15	21	25	43	0.09	0.09	0.11
4	4.68	4.88	5.35	0	0	0	0.00	0.00	0.00
5	4.77	5.11	5.39	0	0	0	0.00	0.00	0.00
6	5.20	5.37	5.75	0	0	0	0.00	0.00	0.00
7	5.24	5.43	5.80	4	3	1	0.09	0.09	0.00
8	5.27	5.59	5.85	1	1	5	0.00	0.00	0.05
<b>CFP</b>									
1	3.30	3.39	3.54	9	8	8	0.53	0.57	0.64
2	3.64	3.75	4.10	0	0	0	0.00	0.00	0.00
3	3.83	4.01	4.58	2	2	10	0.03	0.03	0.04
4	4.13	4.30	4.79	35	42	33	0.02	0.02	0.01
5	4.32	4.47	4.96	1	1	45	0.08	0.09	0.14
6	4.59	4.90	5.12	0	0	0	0.00	0.00	0.00
7	4.88	5.00	5.19	5	4	0	0.01	0.01	0.00
8	4.93	5.23	5.52	2	17	2	0.00	0.16	0.00
<b>GFP<sub>A</sub></b>									
1	3.06	3.12	3.13	2	2	2	0.90	0.93	1.00
2	3.15	3.28	3.83	0	0	3	0.00	0.00	0.00
3	3.23	3.52	3.87	2	2	0	0.00	0.00	0.00
4	3.94	4.13	4.34	0	0	0	0.00	0.00	0.00
5	4.00	4.19	4.47	0	0	4	0.00	0.00	0.05
6	4.01	4.27	4.52	1	1	1	0.00	0.00	0.00

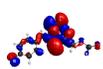
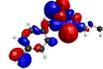
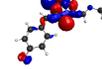
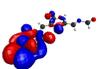
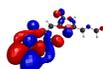
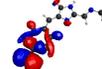
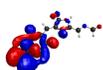
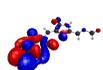
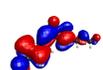
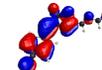
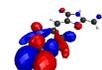
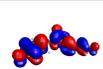
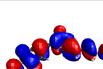
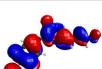
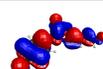
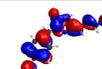
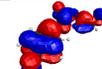
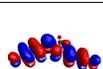
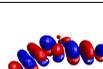
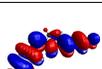
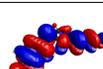
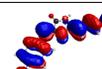
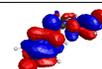
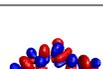
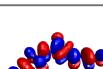
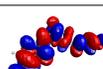
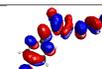
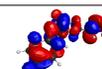
7	4.18	4.30	4.57	2	2	0	0.05	0.05	0.00
8	4.27	4.44	4.76	11	6	76	0.01	0.00	0.06
<hr/>									
GFP <sub>N</sub>									
1	3.49	3.57	3.72	5	5	6	0.72	0.75	0.87
2	3.57	3.68	4.03	0	0	0	0.00	0.00	0.00
3	4.23	4.38	4.72	7	8	3	0.00	0.01	0.11
4	4.37	4.51	4.92	39	45	76	0.01	0.02	0.00
5	4.83	5.01	5.53	1	2	0	0.10	0.00	0.00
6	4.92	5.27	5.55	0	0	2	0.00	0.16	0.00
7	5.15	5.31	5.60	1	0	0	0.00	0.00	0.02
8	5.16	5.48	5.72	0	2	1	0.00	0.00	0.16
<hr/>									
KO <sub>A</sub>									
1	2.45	2.52	2.61	19	15	15	0.87	0.92	1.03
2	2.80	2.94	3.60	0	0	0	0.00	0.00	0.00
3	3.26	3.48	3.97	162	138	331	0.11	0.11	0.22
4	3.48	3.64	4.22	0	0	0	0.00	0.00	0.00
5	3.58	3.78	4.23	127	201	4	0.00	0.00	0.00
6	3.74	3.99	4.40	19	27	184	0.00	0.00	0.07
7	3.88	4.03	4.42	13	22	399	0.02	0.04	0.04
8	3.96	4.10	4.51	0	28	40	0.16	0.20	0.02
<hr/>									
OR <sub>A</sub>									
1	2.51	2.57	2.64	12	11	11	0.86	0.91	1.01
2	2.83	2.96	3.61	0	0	0	0.00	0.00	0.00
3	3.43	3.62	4.04	81	52	222	0.11	0.11	0.21
4	3.50	3.66	4.22	0	0	0	0.00	0.00	0.00
5	3.72	3.95	4.35	234	272	3	0.01	0.05	0.03
6	3.92	4.08	4.44	18	40	154	0.01	0.00	0.06
7	4.00	4.17	4.51	89	50	1	0.03	0.16	0.02
8	4.03	4.26	4.66	-	180	590	0.17	0.00	0.01
<hr/>									
BLB									
1	2.77	2.86	3.15	26	28	30	0.64	0.68	0.80
2	2.93	3.07	3.55	0	0	0	0.00	0.00	0.00
3	3.07	3.15	3.60	0	0	0	0.00	0.00	0.00
4	3.37	3.55	4.22	1	1	3	0.01	0.01	0.01
5	3.84	3.99	4.51	68	84	0	0.03	0.04	0.00
6	4.10	4.21	4.53	0	0	109	0.00	0.00	0.08
7	4.24	4.37	4.77	62	82	329	0.28	0.26	0.16
8	4.36	4.51	4.93	0	0	0	0.00	0.00	0.00

**Table A6:** *Experimental data generated from the spectra in Supplementary Figure 1 by Drobizhev et al.<sup>4</sup> Some proteins were excluded as discussed in the text. “True Peak of  $S_n$ ” refers to the first fully resolved spectroscopic peak higher in energy than the lowest-energy transition (denoted here as  $S_1$ ). “largest  $S_n$ ” refers to the largest TPA cross-section (not a true spectroscopic peak) measured in the highest accessible energy region of the spectrum.*

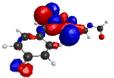
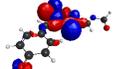
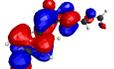
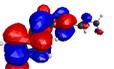
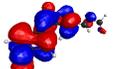
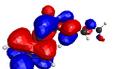
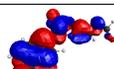
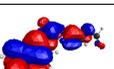
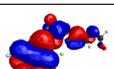
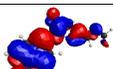
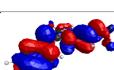
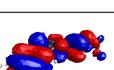
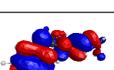
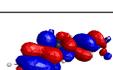
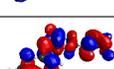
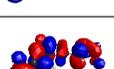
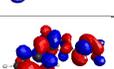
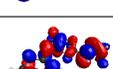
Protein	largest $S_n$		True Peak of $S_n$		$S_1$	
	Energy (eV)	Cross-section (GM)	Energy (eV)	Cross-section (GM)	Energy (eV)	Cross-section (GM)
Ebfp1.2	4.493	10	-	-	3.307	11
EBFP1.5	4.509	11	-	-	3.307	11
EBFP2.0	4.493	16	-	-	3.307	13
BFP Average	4.498	12	-	-	3.307	12
mBlueberry1	4.509	22	-	-	3.047	4
ECFP	4.509	15	3.815	12	2.894	23
Cerulean	4.509	27	3.735	13	2.890	23
CFP Average	4.509	21	3.775	13	2.892	23
Wt-GFP A	4.509	36	-	-	3.062	15
mAmetrine	4.336	27	3.786	19	3.066	56
mKalama1	4.429	31	3.845	23	3.212	36
GFP <sub>N</sub> Average	4.425	31	3.8155	21	3.113	36
MTFP0.7	-	-	3.707	10	2.867	85
MTFP0.8	-	-	3.718	8	2.802	65
MTFP1.0	-	-	3.718	6	2.834	70
Wt-GFP B	-	-	-	-	2.696	12
EGFP (pH8)	-	-	3.758	22	2.675	39
G1	4.351	16	3.701	11	2.696	63
G3	4.269	12	3.663	7	2.647	40
TagGFP2	-	-	3.906	12	2.768	34
mWasabi	4.269	12	3.701	8	2.675	42
GFP <sub>A</sub> Average	4.296	13	3.734	11	2.833	46
mKok	3.615	93	-	-	2.375	41
mOrange	3.875	200	-	-	2.296	67
mBanana	3.680	102	-	-	2.318	64
TagRFP	3.647	330	3.267	300	2.362	95
tdTomato	3.626	158	-	-	2.362	159
mTangerine	3.573	64	3.010	12	2.351	15
DsRED2	3.543	112	-	-	2.362	103
mStrawberry	3.543	35	-	-	2.318	20
mRFP	3.421	125	-	-	2.296	44
mCherry	3.351	101	-	-	2.296	27
mRaspberry	3.523	346	-	-	2.218	31
E2-Crimson	3.298	245	-	-	2.179	23
mPlum	3.425	114	-	-	2.244	22
Katushka	3.435	125	-	-	2.296	66
Katushka2	3.533	283	-	-	2.175	62
tdKatushka2	3.493	323	-	-	2.255	72
mKate (pH8)	3.518	231	-	-	2.218	52

mKate2	3.483	216	-	-	2.175	72
Mkate-S158A	3.523	363	-	-	2.234	79
Mkate-S158C	3.523	400	-	-	2.234	73
mGrape3	3.374	105	-	-	2.175	32
Neptune	3.425	368	-	-	2.214	72
eqFP670	3.360	100	-	-	2.214	23
eqFP650	3.316	114	-	-	2.230	45
<hr/> RFP <sub>A</sub> Average	3.465	197	3.139	-	2.260	57

**Table A7:** Orbitals involved in the studied excitations for the  $RFP_A$  model. The orbitals for the model where the acylimine carbonyl is non-coplanar with the rest of the molecule are added here for convenience. All orbitals shown in this supplementary information come from TDDFT computations.

Orbital	B3LYP	PBE0	CAM-B3LYP	LC-BLYP	CAM-B3LYP (gas-phase)	B3LYP (non-coplanar)
$H_{-4}$						
$H_{-3}$						
$H_{-2}$						
$H_{-1}$						
H						
L						
$L_{+1}$						

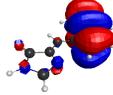
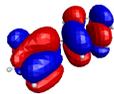
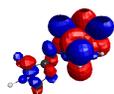
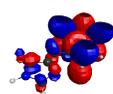
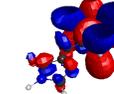
**Table A8:** Orbitals involved in the studied excitations for the  $RFP_A$  model where the acylimine carbonyl is non-coplanar with the rest of the molecule.

Orbital	B3LYP	PBE0	CAM-B3LYP	LC-BLYP
$H_{-4}$				
$H_{-3}$				
$H_{-2}$				
$H_{-1}$				
H				
L				
$L_{+1}$				

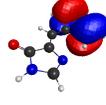
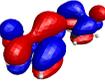
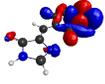
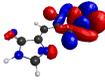
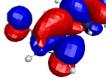
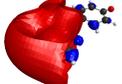
**Table A9:** Orbitals involved in the studied excitations for the BFP1 model.

Orbital	B3LYP	PBE0	CAM-B3LYP	LC-BLYP
H <sub>-4</sub>				
H <sub>-3</sub>				
H <sub>-2</sub>				
H <sub>-1</sub>				
H				
L				
L <sub>+1</sub>				

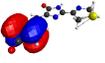
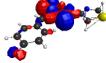
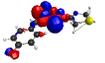
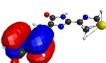
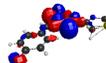
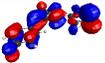
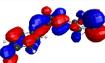
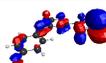
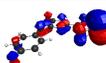
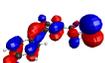
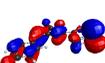
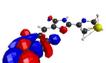
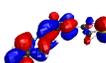
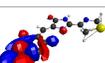
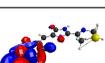
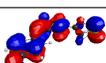
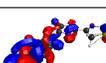
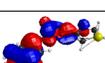
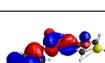
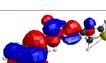
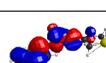
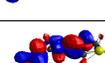
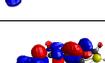
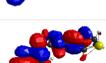
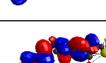
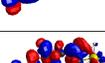
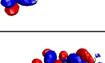
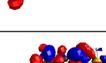
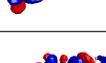
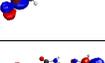
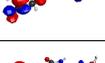
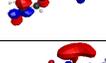
**Table A10:** Orbitals involved in the studied excitations for the CFP model.

Orbital	B3LYP	PBE0	CAM-B3LYP	LC-BLYP
H <sub>-4</sub>				
H <sub>-3</sub>				
H <sub>-2</sub>				
H <sub>-1</sub>				
H				
L				
L <sub>+1</sub>				

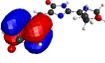
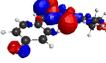
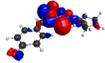
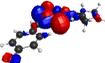
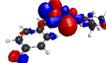
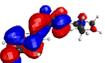
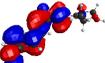
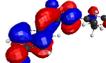
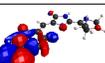
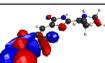
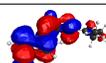
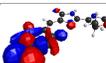
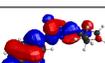
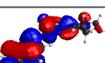
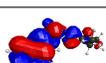
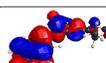
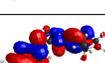
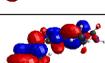
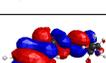
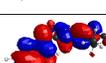
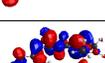
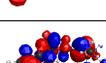
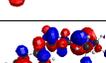
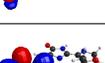
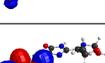
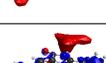
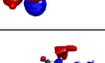
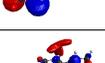
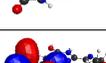
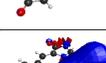
**Table A11:** Orbitals involved in the studied excitations for the  $GFP_A$  model.

Orbital	B3LYP	PBE0	CAM-B3LYP	LC-BLYP
$H_{-3}$				
$H_{-2}$				
$H_{-1}$				
H				
L				
$L_{+1}$				
$L_{+2}$				

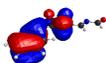
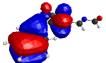
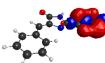
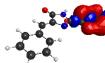
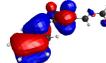
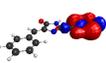
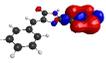
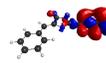
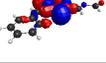
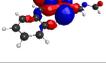
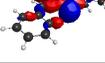
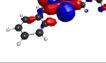
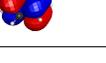
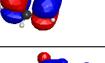
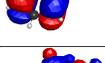
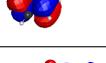
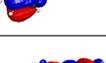
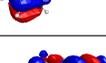
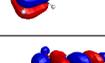
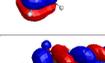
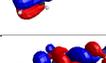
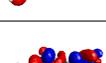
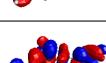
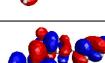
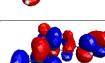
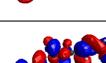
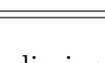
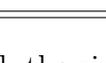
**Table A12:** Orbitals involved in the studied excitations for the  $KO_A$  model.

Orbital	B3LYP	PBE0	CAM-B3LYP	LC-BLYP
$H_{-5}$				
$H_{-4}$				
$H_{-3}$				
$H_{-2}$				
$H_{-1}$				
H				
L				
$L_{+1}$				
$L_{+2}$				

**Table A13:** Orbitals involved in the studied excitations for the  $OR_A$  model.

Orbital	B3LYP	PBE0	CAM-B3LYP	LC-BLYP
$H_{-4}$				
$H_{-3}$				
$H_{-2}$				
$H_{-1}$				
H				
L				
$L_{+1}$				
$L_{+2}$				
$L_{+3}$				

**Table A14:** Orbitals involved in the studied excitations for the BLB model.

Orbital	B3LYP	PBE0	CAM-B3LYP	LC-BLYP	B3LYP (non-coplanar) <sup>a</sup>
H <sub>-4</sub>					
H <sub>-3</sub>					
H <sub>-2</sub>					
H <sub>-1</sub>					
H					
L					
L <sub>+1</sub>					

<sup>a</sup> Refers to the model having the acylimine moiety non-coplanar with the rings; with H-capping as described in text.

**Table A15:** Grouped excitations for RFP, BFP1, CFP and GFP<sub>A</sub> from the data in Table A3 and Table A4 matched by visual inspection of orbital plots from Table A7 to Table A11. ( $\Lambda$  is the overlap diagnostic)

State	Functional	Energy (eV)	$\Lambda$	OPA OS	Cross section (GM)	Transition
RFP <sub>A</sub>						
1	B3LYP	2.28	0.72	1.23	1	H to L
1	PBE0	2.31	0.71	1.27	1	H to L
1	CAM-B3LYP	2.32	0.69	1.34	2	H to L
1	LC-BLYP	2.29	0.67	1.36	3	H to L
3	B3LYP	3.32	0.75	0.02	1	H <sub>-2</sub> to L
3	PBE0	3.47	0.67	0.02	1	H <sub>-2</sub> to L
6	CAM-B3LYP	4.13	0.74	0.09	67	H <sub>-1</sub> to L
6	LC-BLYP	4.45	0.72	0.01	311	H <sub>-1</sub> to L
6	B3LYP	3.61	0.41	0.01	3	H <sub>-4</sub> to L
6	PBE0	3.76	0.42	0.02	3	H <sub>-4</sub> to L
7	CAM-B3LYP	4.27	0.40	0.01	16	H <sub>-3</sub> to L
7	LC-BLYP	4.52	0.40	0.01	53	H <sub>-4</sub> to L
7	B3LYP	3.83	0.72	0.35	23	H to L+1
7	PBE0	3.91	0.71	0.35	19	H to L+1
4	CAM-B3LYP	3.96	0.73	0.34	23	H to L+1
4	LC-BLYP	4.00	0.71	0.44	283	H to L+1
BFP1						
1	B3LYP	3.40	0.81	0.71	1	H to L
1	PBE0	3.46	0.80	0.73	1	H to L
1	CAM-B3LYP	3.55	0.79	0.77	2	H to L
1	LC-BLYP	3.59	0.78	0.76	3	H to L
3	B3LYP	4.42	0.69	0.05	12	H <sub>-2</sub> to L
3	PBE0	4.54	0.69	0.05	14	H <sub>-2</sub> to L
3	CAM-B3LYP	4.88	0.68	0.04	30	H <sub>-1</sub> to L
3	LC-BLYP	4.98	0.67	0.05	43	H <sub>-2</sub> to L
5	B3LYP	5.01	0.51	0.01	2	H <sub>-4</sub> to L
5	PBE0	5.19	0.51	0.01	2	H <sub>-4</sub> to L
5	CAM-B3LYP	5.69	0.50	0.00	3	H <sub>-3</sub> to L
6	LC-BLYP	5.92	0.49	0.00	4	H <sub>-4</sub> to L
6	B3LYP	5.42	0.32	0.00	2	H to L+1
7	PBE0	5.69	0.31	0.00	0	H to L+1
6	CAM-B3LYP	5.76	0.29	0.00	3	H to L+1
5	LC-BLYP	5.84	0.27	0.00	0	H to L+1
CFP						
1	B3LYP	3.14	0.75	0.64	8	H to L

1	PBE0	3.23	0.75	0.67	7	H to L
1	CAM-B3LYP	3.37	0.73	0.74	7	H to L
1	LC-BLYP	3.41	0.72	0.75	8	H to L
2	B3LYP	3.71	0.42	0.04	2	H <sub>-1</sub> to L
2	PBE0	3.89	0.42	0.04	2	H <sub>-1</sub> to L
3	CAM-B3LYP	4.49	0.46	0.05	6	H <sub>-1</sub> to L
3	LC-BLYP	4.74	0.55	0.03	31	H <sub>-1</sub> to L
4	B3LYP	4.07	0.62	0.04	32	H <sub>-2</sub> to L
4	PBE0	4.22	0.61	0.02	37	H <sub>-2</sub> to L
4	CAM-B3LYP	4.71	0.59	0.00	48	H <sub>-2</sub> to L
4	LC-BLYP	4.92	0.56	0.02	20	H <sub>-2</sub> to L
5	B3LYP	4.37	0.65	0.09	1	H to L+1
5	PBE0	4.52	0.65	0.10	1	H to L+1
5	CAM-B3LYP	4.93	0.62	0.19	21	H to L+1
5	LC-BLYP	5.12	0.61	0.21	33	H to L+1
GFP <sub>A</sub>						
1	B3LYP	2.96	0.74	1.05	1	H to L
1	PBE0	3.00	0.74	1.08	1	H to L
1	CAM-B3LYP	3.02	0.71	1.13	2	H to L
1	LC-BLYP	2.97	0.70	1.14	4	H to L
4	B3LYP	4.19	0.64	0.02	2	H to L+1
4	PBE0	4.32	0.58	0.03	2	H to L+1
3	CAM-B3LYP	4.51	0.53	0.05	3	H to L+2
3	LC-BLYP	4.54	0.52	0.06	4	H to L+4
5	B3LYP	4.23	0.64	0.02	4	H <sub>-2</sub> to L
5	PBE0	4.36	0.70	0.01	6	H <sub>-2</sub> to L
5	CAM-B3LYP	4.76	0.76	0.01	24	H <sub>-2</sub> to L
6	LC-BLYP	4.90	0.73	0.01	47	H <sub>-2</sub> to L
7	B3LYP	4.42	0.31	0.01	2	H to L+2
7	PBE0	4.69	0.28	0.00	2	H to L+2
6	CAM-B3LYP	4.78	0.26	0.00	2	H to L+1
5	LC-BLYP	4.86	0.24	0.00	1	H to L+1

**Table A16:** *Grouped excitations for  $KO_A$ ,  $OR_A$  and  $BLB$  from Table A3 and Table A4 matched by visual inspection of orbital plots from Table A12 to Table A14. As can be noticed from the relevant orbital shapes, the higher-energy orbitals don't perfectly match amongst functionals for  $GFP_N$ , so we grouped them with energies. "Non-coplanar" Refers to the model having the acylimine moiety non-coplanar with the rings. Done with H-capping at the same geometry of the similar RFP model and computed with B3LYP at the same level of theory as the planar version. (NET = no equivalent transition)*

State	Functional	Energy (eV)	$\Lambda$	OPA OS	Cross section (GM)	Transition
$GFP_N$						
1	B3LYP	3.34	0.79	0.82	4	H to L
1	PBE0	3.41	0.79	0.84	5	H to L
1	CAM-B3LYP	3.56	0.78	0.86	5	H to L
1	LC-BLYP	3.62	0.76	0.86	6	H to L
3	B3LYP	4.18	0.65	0.00	8	H <sub>1</sub> to L
3	PBE0	4.32	0.65	0.01	11	H <sub>4</sub> to L
3	CAM-B3LYP	4.70	0.62	0.01	6	H to L+2
3	LC-BLYP	4.81	0.61	0.01	3	H to L+4
4	B3LYP	4.30	0.62	0.02	30	H <sub>3</sub> to L
4	PBE0	4.44	0.62	0.02	33	H <sub>3</sub> to L
4	CAM-B3LYP	4.82	0.66	0.02	60	H <sub>4</sub> to L
4	LC-BLYP	4.93	0.65	0.03	71	H <sub>4</sub> to L
$KO_A$						
1	B3LYP	2.45	0.71	1.05	7	H to L
1	PBE0	2.50	0.71	1.09	7	H to L
1	CAM-B3LYP	2.59	0.69	1.19	8	H to L
1	LC-BLYP	2.59	0.68	1.21	8	H to L
3	B3LYP	3.51	0.59	0.00	17	H <sub>-2</sub> to L
3	PBE0	3.69	0.59	0.00	19	H <sub>-2</sub> to L
5	CAM-B3LYP	4.43	0.69	0.01	24	H <sub>-1</sub> to L
6	LC-BLYP	4.72	0.54	0.02	289	H <sub>-1</sub> to L
4	B3LYP	3.62	0.67	0.17	19	H to L+1
4	PBE0	3.77	0.67	0.19	27	H to L+1
3	CAM-B3LYP	4.17	0.62	0.32	205	H to L+1
3	LC-BLYP	4.31	0.57	0.31	339	H to L+1
6	B3LYP	3.91	0.53	0.08	19	H <sub>-5</sub> to L
7	PBE0	4.12	0.48	0.07	41	H <sub>-5</sub> to L
6	CAM-B3LYP	4.51	0.43	0.01	3	H <sub>-4</sub> to L
4	LC-BLYP	4.62	0.44	0.03	3	H <sub>-5</sub> to L
7	B3LYP	3.99	0.53	0.12	55	H <sub>-3</sub> to L

6	PBE0	4.04	0.58	0.10	33	H <sub>-3</sub> to L
7	CAM-B3LYP	4.56	0.46	0.01	32	H <sub>-3</sub> to L
7	LC-BLYP	4.78	0.49	0.00	34	H <sub>-3</sub> to L
<hr/>						
OR <sub>A</sub>						
1	B3LYP	2.48	0.71	1.04	6	H to L
1	PBE0	2.53	0.71	1.08	6	H to L
1	CAM-B3LYP	2.60	0.69	1.18	6	H to L
1	LC-BLYP	2.59	0.68	1.20	6	H to L
<hr/>						
3	B3LYP	3.65	0.71	0.11	22	H to L+1
3	PBE0	3.80	0.70	0.13	30	H to L+1
3	CAM-B3LYP	4.21	0.63	0.30	189	H to L+1
3	LC-BLYP	4.35	0.59	0.29	320	H to L+1
<hr/>						
5	B3LYP	3.92	0.55	0.08	9	H <sub>-2</sub> to L
5	PBE0	4.05	0.63	0.12	19	H <sub>-2</sub> to L
5	CAM-B3LYP	4.47	0.70	0.01	3	H <sub>-1</sub> to L
6	LC-BLYP	4.76	0.72	0.01	217	H <sub>-2</sub> to L
<hr/>						
6	B3LYP	4.00	0.59	0.14	38	H <sub>-4</sub> to L
6	PBE0	4.13	0.52	0.08	29	H <sub>-4</sub> to L
	CAM-B3LYP	NET				
8	LC-BLYP	5.02	0.41	0.13	167	H <sub>-4</sub> to L
<hr/>						
7	B3LYP	4.35	0.48	0.09	21	H to L+2
7	PBE0	4.47	0.48	0.10	29	H to L+2
7	CAM-B3LYP	4.80	0.46	0.12	84	H to L+3
	LC-BLYP	NET				
<hr/>						
BLB						
1	B3LYP	2.68	0.74	0.80	23	H to L
1	PBE0	2.77	0.74	0.83	24	H to L
1	CAM-B3LYP	3.04	0.72	0.95	25	H to L
1	LC-BLYP	3.19	0.71	0.97	23	H to L
1	Non-coplanar	2.63	0.81	0.88	6	H to L
<hr/>						
3	B3LYP	3.24	0.31	0.02	1	H <sub>-1</sub> to L
4	PBE0	3.43	0.31	0.02	1	H <sub>-1</sub> to L
4	CAM-B3LYP	4.12	0.32	0.02	1	H <sub>-1</sub> to L
4	LC-BLYP	4.46	0.35	0.02	4	H <sub>-1</sub> to L
4	Non-coplanar	3.45	0.41	0.02	1	H <sub>-1</sub> to L
<hr/>						
5	B3LYP	3.80	0.60	0.02	58	H <sub>-4</sub> to L
5	PBE0	3.96	0.60	0.03	72	H <sub>-4</sub> to L
5	CAM-B3LYP	4.48	0.61	0.07	75	H <sub>-3</sub> to L
6	LC-BLYP	4.66	0.61	0.08	51	H <sub>-3</sub> to L
5	Non-coplanar	3.75	0.68	0.06	48	H <sub>-4</sub> to L
<hr/>						
6	B3LYP	4.16	0.77	0.28	65	H to L+1
6	PBE0	4.29	0.77	0.26	87	H to L+1

7	CAM-B3LYP	4.71	0.74	0.19	317	H to L+1
7	LC-BLYP	5.00	0.71	0.23	525	H to L+1
6	Non-coplanar	4.03	0.72	0.22	47	H to L <sub>+1</sub>

---

**Table A17:** *Grouped excitations for the non-planar model of RFP<sub>A</sub>. Involved orbitals that were used to match the transitions are shown in Table A8.*

State	Functional	Energy (eV)	$\Lambda$	OPA OS	Cross section (GM)	Transition
1	B3LYP	2.39	0.70	1.10	5	H to L
1	PBE0	2.44	0.70	1.14	6	H to L
1	CAM-B3LYP	2.51	0.68	1.25	7	H to L
1	LC-BLYP	2.51	0.66	1.27	8	H to L
3	B3LYP	3.44	0.71	0.12	18	H to L <sub>+1</sub>
3	PBE0	3.59	0.71	0.14	23	H to L <sub>+1</sub>
3	CAM-B3LYP	4.00	0.66	0.33	159	H to L <sub>+1</sub>
3	LC-BLYP	4.13	0.64	0.36	339	H to L <sub>+1</sub>
5	B3LYP	3.78	0.57	0.09	16	H <sub>-2</sub> to L
5	PBE0	3.92	0.63	0.13	22	H <sub>-1</sub> to L
5	CAM-B3LYP	4.36	0.60	0.02	3	H <sub>-2</sub> to L
7	LC-BLYP	4.67	0.71	0.00	67	H <sub>-2</sub> to L
6	B3LYP	3.84	0.50	0.08	40	H <sub>-4</sub> to L
7	PBE0	4.00	0.47	0.06	32	H <sub>-4</sub> to L
7	CAM-B3LYP	4.46	0.42	0.01	6	H <sub>-3</sub> to L
6	LC-BLYP	4.62	0.45	0.01	4	H <sub>-4</sub> to L
7	B3LYP	3.89	0.35	0.00	3	H <sub>-5</sub> to L
6	PBE0	3.98	0.36	0.00	2	H <sub>-5</sub> to L
6	CAM-B3LYP	4.39	0.49	0.01	9	H <sub>-5</sub> to L
5	LC-BLYP	4.40	0.35	0.00	6	H <sub>-5</sub> to L

# Appendix B

## Appendix to Chapter 3

**Table B1:** *Coordinates of the optimized models at the PBE0/6-31+G(d,p) level of theory in the gas phase.*

Model Number 1a			
C	-2.004341	2.044783	0.000026
C	-0.745110	1.457117	0.000033
C	-0.612638	0.054944	0.000013
C	0.669150	-0.626807	0.000019
C	-1.774466	-0.731069	-0.000014
C	-3.037302	-0.133378	-0.000020
C	-3.155698	1.262783	0.000000
C	3.131168	-0.975476	0.000050
C	3.598382	1.220328	0.000106
C	1.917459	-0.105674	0.000048
C	-4.212836	-0.951833	-0.000046
N	4.161024	-0.033060	0.000089
N	2.304977	1.238482	0.000083
N	-5.167801	-1.613413	-0.000067
H	0.151245	2.067632	0.000054
H	-1.697179	-1.814411	-0.000029
H	0.639826	-1.715793	-0.000004
H	4.217340	2.110997	0.000135
H	5.142484	-0.261640	0.000099
H	-4.140390	1.718765	-0.000006
H	-2.091697	3.126949	0.000041
O	3.254725	-2.183013	0.000025
Model Number 1b			
C	-2.669828	0.350148	-0.000023
C	-1.274237	0.319087	-0.000006
C	-0.602774	-0.915019	-0.000028
C	0.844609	-1.027956	-0.000011
C	-1.364114	-2.097405	-0.000066

C	-2.753961	-2.059019	-0.000082
C	-3.417113	-0.837907	-0.000060
C	3.241065	-0.355063	0.000048
C	2.779450	1.842203	0.000110
C	1.779265	-0.050192	0.000032
C	-3.350363	1.611677	0.000000
N	3.801391	0.923178	0.000100
N	1.589696	1.335019	0.000072
N	-3.910944	2.629119	0.000019
H	-0.702029	1.240155	0.000024
H	-0.851822	-3.056129	-0.000084
H	-3.323106	-2.983389	-0.000112
H	1.251251	-2.038599	-0.000035
H	2.984492	2.907228	0.000148
H	4.791234	1.112035	0.000124
H	-4.501319	-0.792525	-0.000073
O	3.841875	-1.410056	0.000023
<hr/>			
Model Number 2a			
C	-2.382542	0.051571	0.000094
C	-0.991872	-0.031375	0.000047
C	-0.205649	1.135977	0.000012
C	1.244558	1.118170	-0.000038
C	-0.854918	2.382560	0.000024
C	-2.243750	2.467023	0.000070
C	-3.007798	1.307005	0.000105
C	3.574994	0.250403	-0.000119
C	2.939021	-1.903626	-0.000098
C	2.096025	0.065595	-0.000067
C	-3.249785	-1.170363	0.000133
C	-2.594480	-2.528806	0.000123
N	4.030969	-1.069393	-0.000135
N	1.794334	-1.301497	-0.000057
H	-0.487289	-0.990847	0.000038
H	-0.256619	3.290393	-0.000004
H	1.738040	2.089581	-0.000056
H	3.058997	-2.981567	-0.000102
H	5.002631	-1.336190	-0.000169
H	-4.093013	1.340395	0.000141
H	-1.956509	-2.655953	-0.881072
H	-1.956458	-2.655943	0.881284
H	-3.373475	-3.291652	0.000150
H	-2.727314	3.439333	0.000079
O	4.262559	1.252160	-0.000142
O	-4.463720	-1.060596	0.000173
<hr/>			
Model Number 2b			
C	-1.594319	2.245517	0.054927
C	-0.359102	1.612099	0.039753
C	-0.285121	0.204213	-0.001070
C	0.965875	-0.531420	-0.019660
C	-1.479094	-0.528822	-0.025012

C	-2.722921	0.104843	-0.009683
C	-2.774339	1.502254	0.030404
C	3.409547	-0.993011	-0.027753
C	3.977557	1.178966	0.036158
C	2.238125	-0.070178	-0.003598
C	-3.951646	-0.750123	-0.037273
C	-5.305287	-0.082975	-0.021817
N	4.481824	-0.099020	0.000090
N	2.686701	1.255346	0.035688
H	-1.640487	3.330194	0.086460
H	0.560244	2.187434	0.059063
H	-1.462264	-1.615130	-0.056561
H	0.883940	-1.617325	-0.051350
H	4.637148	2.039566	0.062573
H	5.451205	-0.373810	-0.005080
H	-3.730499	2.016361	0.042573
H	-5.428482	0.573113	-0.890401
H	-5.429176	0.533021	0.875550
H	-6.074486	-0.855371	-0.039214
O	3.481814	-2.205101	-0.062224
O	-3.855508	-1.964365	-0.071999
<hr/>			
Model Number 3			
C	-1.170720	-0.877053	-0.543579
C	0.192925	-0.693181	-0.363255
C	0.726344	0.601754	-0.212399
C	2.140312	0.868165	-0.032125
C	-0.152117	1.699465	-0.246914
C	-1.517360	1.523531	-0.425218
C	-2.009616	0.232593	-0.565547
C	4.583936	0.477726	0.209600
C	4.402642	-1.752992	0.026354
C	3.181939	0.005645	0.020789
C	-4.173226	-0.259826	0.209465
N	5.296048	-0.722760	0.199603
N	3.166337	-1.389509	-0.083228
H	-1.589262	-1.869638	-0.669973
H	0.863485	-1.545192	-0.344050
H	0.246531	2.704298	-0.135756
H	-2.198923	2.366686	-0.459185
H	2.423487	1.914751	0.075740
H	4.738767	-2.783490	-0.011809
H	6.295743	-0.792788	0.304566
O	5.047936	1.592659	0.342640
O	-3.371880	0.065860	-0.823905
F	-4.155392	0.659125	1.188817
F	-3.844059	-1.434137	0.774276
F	-5.414531	-0.353563	-0.255981
<hr/>			
Model Number 4			
C	-1.850017	-1.008352	0.000038
C	-0.479634	-0.791376	0.000015

C	0.028156	0.522247	0.000046
C	1.442511	0.840590	0.000050
C	-0.884880	1.592234	0.000091
C	-2.252389	1.358400	0.000103
C	-2.769692	0.054788	0.000078
C	3.911215	0.549125	0.000049
C	3.803954	-1.693714	-0.000038
C	2.517764	0.017483	0.000019
N	4.667839	-0.623538	0.000010
N	2.549414	-1.381283	-0.000035
H	-2.231277	-2.025721	0.000028
H	0.215182	-1.624074	-0.000018
H	-0.507502	2.611969	0.000126
H	-2.916859	2.221520	0.000166
H	1.694943	1.900592	0.000091
H	4.176353	-2.712412	-0.000075
H	5.674938	-0.652087	0.000017
H	-4.954527	1.622812	-0.000635
H	-5.676045	-1.623235	0.000495
O	4.339427	1.686556	0.000098
O	-5.283516	0.721102	-0.000253
O	-4.716291	-1.546364	0.000542
B	-4.309587	-0.244123	0.000113
<hr/>			
Model Number 5			
C	2.556015	-1.211939	0.000033
C	1.202508	-0.906217	0.000054
C	0.768063	0.434352	0.000020
C	-0.628394	0.817470	0.000040
C	1.738081	1.454029	-0.000036
C	3.095123	1.160997	-0.000058
C	3.477513	-0.172399	-0.000023
C	-3.107585	0.628305	0.000104
C	-3.092582	-1.618008	0.000170
C	-1.738785	0.041502	0.000089
N	-3.911293	-0.513559	0.000157
N	-1.826531	-1.355684	0.000133
H	2.905915	-2.238886	0.000059
H	0.461720	-1.698224	0.000097
H	1.417985	2.492692	-0.000063
H	3.849427	1.940536	-0.000102
H	-0.835365	1.887312	0.000013
H	-3.505466	-2.620920	0.000210
H	-4.918709	-0.500774	0.000182
O	-3.491584	1.781849	0.000079
F	4.789792	-0.466502	-0.000044
<hr/>			
Model Number 6			
C	-0.816519	-0.519193	0.000507
C	0.569334	-0.491706	0.000271
C	1.255863	0.738138	-0.000270
C	2.701324	0.844889	-0.000530

C	0.503282	1.927974	-0.000570
C	-0.881620	1.896984	-0.000321
C	-1.555902	0.670675	0.000218
C	5.098300	0.180362	-0.000561
C	4.649835	-2.019309	0.000545
C	3.638924	-0.132255	-0.000215
C	-3.045721	0.708497	0.000485
C	-5.594086	-0.282195	0.001151
C	-6.548255	-1.464394	-0.000164
N	5.666664	-1.094665	-0.000031
N	3.457183	-1.518842	0.000462
H	-1.323793	-1.479728	0.000942
H	1.139579	-1.413752	0.000512
H	1.020640	2.884087	-0.000996
H	-1.465417	2.811904	-0.000537
H	3.110910	1.854595	-0.001038
H	4.861493	-3.083072	0.001029
H	6.657624	-1.276965	-0.000067
H	-5.730808	0.349374	0.882785
H	-5.730471	0.351217	-0.879225
H	-6.415172	-2.091824	-0.886850
H	-6.415326	-2.093634	0.885281
H	-7.579906	-1.098662	0.000103
O	5.695224	1.238162	-0.001140
O	-3.671033	1.750982	0.000468
S	-3.873311	-0.870320	0.000792

---

Model Number 7

C	-0.486788	-0.513583	-0.013210
C	0.899096	-0.486516	-0.002924
C	1.586985	0.739154	-0.092247
C	3.032514	0.845173	-0.083714
C	0.835617	1.925637	-0.192120
C	-0.549233	1.895534	-0.199073
C	-1.225033	0.673049	-0.111435
C	5.428300	0.183812	0.013845
C	4.977193	-2.005936	0.217929
C	3.968705	-0.127761	0.021174
C	-2.715401	0.712486	-0.119274
C	-5.282238	-0.327287	-0.109726
C	-5.689103	-0.012137	1.322699
C	-6.142283	-1.417677	-0.737037
N	5.995044	-1.085233	0.145115
N	3.785317	-1.508195	0.150613
H	-0.996334	-1.469424	0.065844
H	1.468117	-1.405747	0.078143
H	1.354141	2.878468	-0.262806
H	-1.131624	2.808438	-0.271811
H	3.443284	1.850643	-0.170533
H	5.187579	-3.064701	0.324262
H	6.985617	-1.266336	0.179683

H	-5.327101	0.586113	-0.711956
H	-5.608137	-0.900790	1.957315
H	-5.068051	0.781280	1.745081
H	-6.731420	0.327829	1.342212
H	-5.864489	-1.604489	-1.778000
H	-6.060467	-2.360101	-0.184344
H	-7.192909	-1.109378	-0.711817
O	6.026304	1.237061	-0.077860
O	-3.335613	1.757909	-0.086444
S	-3.534490	-0.869236	-0.201942

---

Model Number 8

C	2.301627	-1.104416	0.000052
C	0.942094	-0.839654	0.000048
C	0.473653	0.488701	0.000001
C	-0.934591	0.836107	0.000000
C	1.414168	1.535647	-0.000045
C	2.775586	1.277652	-0.000042
C	3.226883	-0.049042	0.000007
C	-3.408851	0.570648	0.000036
C	-3.322852	-1.672156	0.000105
C	-2.018358	0.025065	0.000036
C	4.630370	-0.325955	0.000013
N	-4.176156	-0.595440	0.000082
N	-2.064388	-1.372205	0.000081
N	5.770902	-0.549425	0.000017
H	2.660057	-2.128737	0.000089
H	0.223531	-1.651522	0.000083
H	1.064666	2.564599	-0.000082
H	3.493119	2.091581	-0.000077
H	-1.169800	1.899701	-0.000032
H	-3.703823	-2.687688	0.000141
H	-5.183654	-0.615993	0.000095
O	-3.823146	1.711515	0.000003

---

Model Number 9

C	1.726088	-1.194310	-0.000061
C	0.367394	-0.909306	-0.000080
C	-0.076342	0.425919	-0.000050
C	-1.472195	0.808823	-0.000052
C	0.894485	1.445168	-0.000010
C	2.249080	1.147717	0.000009
C	2.692816	-0.181046	-0.000017
C	-3.951882	0.628373	-0.000239
C	-3.945854	-1.618255	0.000039
C	-2.586067	0.036573	-0.000096
C	4.168339	-0.518600	-0.000001
C	4.866325	-0.003229	1.262754
C	4.866374	-0.003140	-1.262692
N	-4.760354	-0.510057	0.000202
N	-2.679139	-1.360525	-0.000165
H	-0.365388	-1.708905	-0.000115

H	0.571307	2.483677	0.000012
H	-1.677165	1.879271	-0.000050
H	-4.363358	-2.619291	0.000236
H	-5.767601	-0.492483	0.000829
H	2.970196	1.961288	0.000044
H	4.245474	-1.614089	-0.000038
H	4.387668	-0.389198	2.168003
H	5.917571	-0.311057	1.272445
H	4.840178	1.091101	1.310864
H	4.387752	-0.389044	-2.167988
H	4.840229	1.091193	-1.310726
H	5.917620	-0.310967	-1.272365
H	2.050061	-2.232865	-0.000083
O	-4.333682	1.783416	0.000207
<hr/>			
Model Number 10			
C	1.458448	-0.996653	-0.000049
C	0.089605	-0.769091	-0.000011
C	-0.419834	0.543763	0.000020
C	-1.835284	0.853307	0.000044
C	0.492949	1.613746	0.000019
C	1.864608	1.396937	-0.000019
C	2.336655	0.086949	-0.000057
C	-4.301596	0.530779	0.000057
C	-4.164109	-1.710879	0.000010
C	-2.901864	0.018721	0.000035
N	-5.041919	-0.652876	0.000039
N	-2.913863	-1.380578	0.000006
H	1.845422	-2.010178	-0.000075
H	-0.602867	-1.603868	-0.000009
H	0.116983	2.633801	0.000045
H	2.556085	2.232676	-0.000024
H	-2.096056	1.911092	0.000068
H	-4.522178	-2.734678	-0.000009
H	-6.048566	-0.695350	0.000046
O	-4.745879	1.661923	0.000082
BR	4.202032	-0.230176	-0.000128
<hr/>			
Model Number 11			
C	-1.940372	-1.077351	-0.001470
C	-0.575118	-0.828797	0.000937
C	-0.096791	0.494895	-0.011569
C	1.312487	0.835729	-0.008171
C	-1.032222	1.548364	-0.027093
C	-2.393053	1.294003	-0.028703
C	-2.865040	-0.024026	-0.015692
C	3.785823	0.572682	0.012676
C	3.703648	-1.670507	0.043535
C	2.397327	0.025412	0.010067
C	-4.344201	-0.246155	-0.017251
C	-4.871927	-1.659639	-0.007970
N	4.555325	-0.591938	0.034786

N	2.445155	-1.372405	0.029948
H	-2.285620	-2.106798	0.008416
H	0.137031	-1.646327	0.012748
H	-0.674026	2.574778	-0.037383
H	-3.118591	2.101484	-0.039751
H	1.549829	1.898949	-0.020902
H	4.086725	-2.685111	0.060671
H	5.562728	-0.609855	0.042848
H	-4.533923	-2.198238	0.884194
H	-5.961690	-1.625636	-0.017888
H	-4.517076	-2.217199	-0.881660
O	4.200611	1.714054	-0.000156
O	-5.108454	0.703857	-0.025592
<hr/>			
Model Number 12			
C	2.039037	-0.688645	0.000057
C	0.659508	-0.570497	0.000036
C	0.043882	0.697438	0.000007
C	-1.388930	0.888425	-0.000018
C	0.872259	1.836231	0.000001
C	2.251896	1.726630	0.000022
C	2.843602	0.459642	0.000050
C	-3.820250	0.366913	-0.000037
C	-3.503484	-1.857043	0.000027
C	-2.385601	-0.030504	-0.000008
N	-4.463043	-0.872891	-0.000011
N	-2.283992	-1.425968	0.000030
N	4.810566	-0.665569	0.000089
N	4.250140	0.431443	0.000069
N	5.448749	-1.604624	0.000109
H	2.492269	-1.676604	0.000078
H	0.035581	-1.457755	0.000042
H	0.416813	2.823188	-0.000022
H	2.887330	2.606208	0.000016
H	-1.739453	1.920324	-0.000048
H	-3.777967	-2.906275	0.000053
H	-5.463138	-0.994883	-0.000021
O	-4.357616	1.457808	-0.000074
<hr/>			
Model Number 13			
C	2.229640	-0.886854	0.000030
C	0.853906	-0.697061	0.000015
C	0.298332	0.595103	-0.000093
C	-1.120718	0.861182	-0.000116
C	1.185158	1.692275	-0.000187
C	2.554605	1.515373	-0.000170
C	3.089054	0.219504	-0.000061
C	-3.577230	0.481221	-0.000075
C	-3.390676	-1.758485	0.000168
C	-2.170807	0.002209	-0.000023
C	5.036680	-1.140234	0.000014
N	-4.291378	-0.719261	0.000055

N	-2.148895	-1.397880	0.000128
H	2.623784	-1.896892	0.000112
H	0.188468	-1.553856	0.000089
H	0.778643	2.700701	-0.000273
H	3.235774	2.360204	-0.000241
H	-1.410813	1.911741	-0.000225
H	-3.726099	-2.789865	0.000279
H	-5.296777	-0.782158	0.000062
H	6.113041	-0.966055	-0.000004
H	4.762006	-1.710840	0.895720
H	4.761992	-1.710944	-0.895624
O	-4.052586	1.601977	-0.000199
O	4.437838	0.139479	-0.000054
<hr/>			
Model Number 14			
C	-1.521775	-0.707662	0.000010
C	-0.138518	-0.583302	0.000012
C	0.477601	0.680497	-0.000009
C	1.909178	0.871876	-0.000008
C	-0.355447	1.818948	-0.000035
C	-1.732224	1.707453	-0.000037
C	-2.326045	0.438646	-0.000015
C	4.341181	0.355341	0.000019
C	4.029534	-1.870235	0.000076
C	2.909420	-0.044311	0.000020
C	-4.330115	-0.833708	-0.000010
C	-5.770133	-0.626608	-0.000023
C	-6.970187	-0.495048	-0.000035
N	4.986852	-0.883442	0.000057
N	2.809754	-1.440849	0.000056
H	-1.958326	-1.700114	0.000027
H	0.484831	-1.471060	0.000031
H	0.097826	2.807164	-0.000052
H	-2.373617	2.582715	-0.000056
H	2.258579	1.904367	-0.000031
H	4.306385	-2.918851	0.000106
H	5.987187	-1.002545	0.000066
H	-4.039338	-1.411215	-0.889314
H	-4.039352	-1.411197	0.889311
H	-8.030007	-0.368386	-0.000045
O	4.879061	1.447056	-0.000007
O	-3.681037	0.427592	-0.000018
<hr/>			
Model Number 15			
C	-0.922225	-0.611953	0.011280
C	0.463593	-0.555251	0.015213
C	1.122723	0.685386	-0.074400
C	2.566085	0.822384	-0.077096
C	0.343797	1.855362	-0.168622
C	-1.039090	1.793407	-0.171533
C	-1.691036	0.557020	-0.080234
C	4.975621	0.207406	-0.014948

C	4.569932	-1.993021	0.162609
C	3.521720	-0.133128	0.006030
C	-3.183831	0.552097	-0.077159
C	-3.893584	-0.789975	0.014194
C	-5.411525	-0.697537	-0.042171
C	-6.111223	-0.012419	1.097921
N	5.568416	-1.052141	0.090305
N	3.367328	-1.518143	0.118333
H	-1.402612	-1.583253	0.078319
H	1.053745	-1.462032	0.086155
H	0.840864	2.819507	-0.239854
H	-1.642691	2.692624	-0.244209
H	2.955829	1.836675	-0.157428
H	4.802148	-3.048897	0.249940
H	6.562714	-1.214665	0.108050
H	-3.591294	-1.278671	0.950976
H	-3.565054	-1.431086	-0.810832
H	-5.877918	1.056740	1.056521
H	-7.188359	-0.162040	1.014278
H	-5.751639	-0.387905	2.062233
O	5.550875	1.273221	-0.098770
O	-3.808106	1.596200	-0.150621
O	-6.012226	-1.186878	-0.976301
<hr/>			
Model Number 16a			
C	-2.128383	-1.539240	0.000096
C	-0.797600	-1.143725	0.000072
C	-0.469768	0.224954	0.000070
C	0.889643	0.718414	0.000042
C	-1.510733	1.173594	0.000092
C	-2.837303	0.779506	0.000116
C	-3.145056	-0.588895	0.000118
C	3.374446	0.740466	-0.000007
C	3.553082	-1.499461	-0.000004
C	2.062996	0.039194	0.000019
N	4.273582	-0.328358	-0.000020
N	2.269191	-1.345585	0.000019
H	-2.382070	-2.597483	0.000098
H	-0.003201	-1.881018	0.000056
H	-1.288952	2.236955	0.000090
H	1.008625	1.801666	0.000039
H	4.051431	-2.462738	-0.000011
H	5.275889	-0.227149	-0.000041
H	-4.615610	-1.833772	0.000133
H	-4.675170	1.251658	0.000150
O	3.662284	1.922757	-0.000017
O	-4.477165	-0.881285	0.000141
O	-3.824001	1.708450	0.000136
<hr/>			
Model Number 16b			
C	2.388879	0.829189	0.000137
C	1.018223	0.642222	0.000096

C	0.478574	-0.659677	0.000030
C	-0.942016	-0.930477	-0.000008
C	1.354086	-1.759280	0.000005
C	2.730774	-1.572898	0.000045
C	3.248191	-0.283456	0.000112
C	-3.399603	-0.556890	-0.000029
C	-3.216621	1.682743	0.000085
C	-1.992593	-0.074061	0.000015
N	-4.115863	0.641276	0.000021
N	-1.974380	1.325453	0.000085
H	0.361734	1.505389	0.000117
H	0.948948	-2.766994	-0.000047
H	-1.227538	-1.981954	-0.000063
H	-3.554097	2.713451	0.000131
H	-5.121368	0.702650	0.000010
H	5.108200	-0.788084	0.000142
H	3.403120	-2.428043	0.000025
H	3.865607	2.021401	0.000220
O	-3.870359	-1.679219	-0.000093
O	4.578685	0.015551	0.000156
O	2.901647	2.083116	0.000200

---

Model Number 17

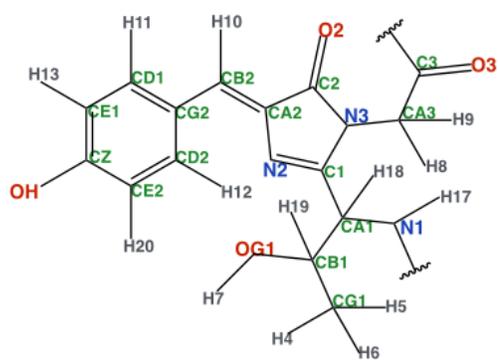
C	-1.443721	-0.621977	-0.109757
C	-0.058642	-0.531950	-0.072378
C	0.588225	0.714875	-0.003500
C	2.022775	0.868677	0.043203
C	-0.217224	1.872954	0.025397
C	-1.595720	1.795360	-0.013291
C	-2.221720	0.543108	-0.081547
C	4.440733	0.291362	0.080727
C	4.077017	-1.922546	-0.049727
C	3.001093	-0.071415	0.021093
C	-4.271925	-0.673375	-0.185982
C	-5.732062	-0.372439	-0.258394
C	-6.644124	-0.934021	0.533012
N	5.057016	-0.961344	0.029337
N	2.868370	-1.463095	-0.057566
H	-1.906181	-1.601325	-0.157873
H	0.543581	-1.433924	-0.094545
H	0.259299	2.848780	0.079320
H	-2.215366	2.686007	0.009864
H	2.396964	1.890471	0.106978
H	4.328715	-2.976210	-0.100495
H	6.053867	-1.105171	0.048764
H	-6.030690	0.332951	-1.032452
H	-3.951268	-1.217349	-1.087705
H	-4.043508	-1.293207	0.691291
H	-7.702067	-0.717168	0.420764
H	-6.365122	-1.631928	1.319455
O	5.004391	1.367735	0.155316

O	-3.572877	0.563101	-0.111857
<hr/>			
Model Number 18			
C	2.542622	-1.189107	-0.003976
C	1.193177	-0.892866	-0.003170
C	0.742111	0.444221	-0.001559
C	-0.648408	0.818073	-0.000208
C	1.718325	1.460083	-0.001021
C	3.070757	1.170319	-0.001334
C	3.508816	-0.164987	-0.002366
C	-3.127153	0.621202	0.000452
C	-3.109335	-1.626909	0.001688
C	-1.762430	0.039832	0.000963
N	-3.929463	-0.522939	0.000797
N	-1.844556	-1.358402	0.001964
N	4.850791	-0.465106	-0.049828
H	2.866453	-2.227267	-0.010277
H	0.458732	-1.691176	-0.002768
H	1.399245	2.499665	0.000659
H	3.800681	1.976113	-0.005803
H	5.132379	-1.396645	0.210128
H	-0.859826	1.887434	-0.000868
H	-3.519538	-2.630856	0.002193
H	-4.936626	-0.510310	0.000000
H	5.502754	0.256075	0.213851
O	-3.519861	1.774741	-0.000130
<hr/>			
Model Number 19			
C	1.390565	0.702759	0.000410
C	3.872932	0.809039	0.000396
C	4.128517	-1.422606	-0.000246
C	2.584466	0.061422	0.000168
C	0.047037	0.165103	0.000241
C	-1.013231	1.061962	-0.000079
C	-0.235989	-1.235055	0.000372
C	-2.356852	0.627531	-0.000251
C	-1.531056	-1.679341	0.000192
C	-2.627125	-0.774851	-0.000117
C	-3.443914	1.539430	-0.000563
C	-3.973020	-1.212087	-0.000305
C	-4.740760	1.084189	-0.000736
C	-5.007343	-0.304135	-0.000605
N	4.808340	-0.226829	0.000320
N	2.840276	-1.314850	-0.000369
H	1.473239	1.789378	0.000834
H	4.660003	-2.368016	-0.000471
H	5.806726	-0.091930	0.000523
H	-0.810521	2.131107	-0.000190
H	0.594073	-1.932505	0.000619
H	-1.737555	-2.746928	0.000287
H	-3.233802	2.606191	-0.000665
H	-4.177996	-2.279805	-0.000206

H	-5.566149	1.790183	-0.000976
H	-6.036133	-0.652928	-0.000744
O	4.115306	2.000856	0.001007
<hr/>			
Model Number 20			
C	-0.951254	0.385524	-0.058413
C	-3.391493	0.885348	-0.073680
C	-3.990238	-1.274647	0.112890
C	-2.242940	-0.041810	-0.007198
C	0.240059	-0.401838	-0.006253
C	1.622720	0.044696	0.004058
C	0.286806	-1.790970	-0.015331
C	2.430160	-1.113357	-0.022591
C	2.258236	1.305588	0.052117
C	3.825202	-1.093195	-0.053788
C	3.654539	1.331971	0.010784
C	4.419396	0.159705	-0.048538
N	-4.476828	0.006090	0.007114
N	-2.698421	-1.359440	0.110870
N	1.525623	2.486017	0.086500
N	1.578393	-2.200728	-0.027096
H	-0.851380	1.460349	-0.185990
H	-4.658758	-2.124897	0.190985
H	-5.441888	0.294571	-0.003779
H	-0.543579	-2.482261	-0.027042
H	1.875507	-3.162262	-0.050316
H	4.412760	-2.005425	-0.081766
H	4.159299	2.294792	0.036747
H	5.502105	0.240075	-0.075350
H	0.677464	2.462072	0.633195
H	2.073074	3.308444	0.294833
O	-3.457058	2.099991	-0.166379
<hr/>			
Model Number 21			
C	-1.709938	0.781640	-0.000170
C	-4.192548	0.668746	-0.000389
C	-4.246741	-1.576278	-0.000369
C	-2.841014	0.037478	-0.000266
C	-0.324662	0.364402	-0.000041
C	0.667774	1.330106	0.000055
C	0.095330	-1.004454	-0.000005
C	2.028711	0.963833	0.000183
C	2.317290	-0.432837	0.000207
C	3.077450	1.916856	0.000284
C	3.684141	-0.842417	0.000338
C	4.381145	1.485942	0.000407
C	4.686511	0.108901	0.000435
N	-5.031119	-0.446152	-0.000448
N	-2.973174	-1.354717	-0.000263
N	1.348718	-1.383857	0.000113
H	-1.887651	1.857065	-0.000192
H	-4.691176	-2.565559	-0.000396

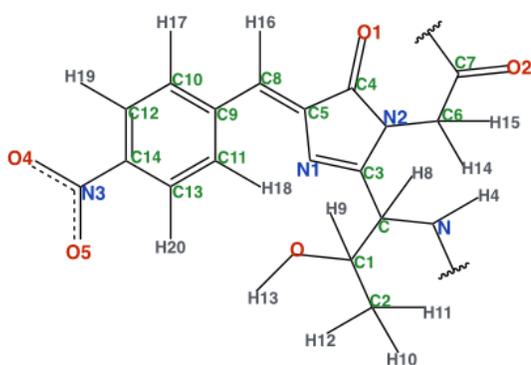
H	-6.037664	-0.401365	-0.000538
H	0.403985	2.386014	0.000032
H	-0.659550	-1.785644	-0.000078
H	2.836392	2.976096	0.000264
H	5.195247	2.204746	0.000485
H	5.727196	-0.208774	0.000536
H	4.878492	-2.319640	0.000450
O	-4.538142	1.834084	-0.000431
O	3.927013	-2.169781	0.000364
<hr/>			
Model Number 22			
C	-1.893561	-1.060417	-0.000137
C	-0.530182	-0.809610	-0.000113
C	-0.048098	0.514483	-0.000001
C	1.363734	0.849715	0.000025
C	-0.975364	1.573941	0.000086
C	-2.340453	1.333755	0.000061
C	-2.779601	0.014080	-0.000050
C	3.835722	0.565418	-0.000015
C	3.732448	-1.676104	-0.000218
C	2.440134	0.029644	-0.000053
N	4.594076	-0.606236	-0.000127
N	2.475869	-1.367189	-0.000180
N	-4.218699	-0.250316	-0.000076
H	-2.283213	-2.072024	-0.000223
H	0.179108	-1.629436	-0.000182
H	-0.613284	2.598304	0.000174
H	-3.062961	2.141453	0.000126
H	1.607538	1.911251	0.000117
H	4.105854	-2.694443	-0.000312
H	5.601452	-0.634641	-0.000139
O	4.257018	1.703183	0.000086
O	-4.971051	0.712822	-0.000019
O	-4.579678	-1.417712	-0.000151

**Table B2:** Atom Types and charges for the control model mapped by the atom names. The lower-case letters refer to atom types as in the GAFF. The charges are computed by R.E.D. Server Development online tools as explained in the main text.



Atom Name	Atom type	Charge
N1	n	-0.373500
H17	hn	0.237600
CA1	c3	-0.005700
H18	h1	0.169400
CB1	c3	0.104500
H19	h1	0.179200
CG1	c3	-0.326300
H4	hc	0.089800
H5	hc	0.089800
H6	hc	0.089800
OG1	oh	-0.605100
H7	ho	0.385500
C1	cc	0.012200
N2	nd	-0.307100
N3	n	0.011000
C2	c	0.355800
O2	o	-0.629500
CA2	cd	-0.039300
CA3	c3	-0.047000
H8	h1	0.046500
H9	h1	0.046500
C3	c	0.589400
O3	o	-0.605400
CB2	ce	-0.072700
H10	ha	0.129100
CG2	ca	-0.077800
CD1	ca	-0.103300
H11	ha	0.129900
CD2	ca	-0.103300
H12	ha	0.129900
CE1	ca	-0.369800
H13	ha	0.127800
CE2	ca	-0.369800
H20	ha	0.127800
CZ	c	0.657000
OH	o	-0.672900

**Table B3:** Atom Types and charges for the nitro model mapped by the atom names. The lower-case letters refer to atom types as in the GAFF. The charges are computed by R.E.D. Server Development online tools as explained in the main text.



Atom Name	Atom type	Charge
C7	c	0.567500
O2	o	-0.541800
C6	c3	-0.113900
H14	h1	0.090300
H15	h1	0.090300
N2	n	0.029900
C4	c	0.410300
O1	o	-0.529000
C3	cc	0.226600
C	c3	-0.047700
N	n	-0.363100
H4	hn	0.312800
C1	c3	0.027100
C2	c3	-0.054400
H10	hc	0.030700
H11	hc	0.030700
H12	hc	0.030700
O	oh	-0.679800
H13	ho	0.473700
H9	h1	0.119200
H8	h1	0.167600
N1	nd	-0.377400
C5	cd	0.102100
C8	ce	-0.154600
H16	ha	0.139600
C9	ca	0.052700
C11	ca	-0.083300
H18	ha	0.129900
C13	ca	-0.191500
H20	ha	0.181200
C14	ca	0.032100
N3	no	0.773600
O4	o	-0.459200
O5	o	-0.459200
C12	ca	-0.191500
H19	ha	0.181200
C10	ca	-0.083300
H17	ha	0.129900

Additional Parameters for the interface atoms between the chromophore model parameterized with GAFF and the rest of the protein parameterized with FF12SB. The parameters are given in the frcmod format. Some are produced by ANTECHAMBER and the rest are assigned via comparison to parameters for similar atom types.

#### BOND

C-n	490.0	1.335
c-N	490.0	1.335

#### ANGLE

ce-cd-c	63.460	120.890
o -c -N	80.0	122.90
O -C -n	80.0	122.90
c -N -H	50.0	120.00
C -n -hn	50.0	120.00
C -n -c3	63.92	121.35
c3-c -N	67.86	115.15
CX-C -n	70.0	116.60
c -N -CX	50.0	121.90

#### DIHE

ca-ca-ce-ha	1	6.650	180.000	2.000
ca-ca-ce-cd	1	6.650	180.000	2.000
ca-ce-cd-c	1	1.000	180.000	2.000
ca-ce-cd-nd	1	1.000	180.000	2.000
ha-ce-cd-c	1	1.000	180.000	2.000
ha-ce-cd-nd	1	1.000	180.000	2.000
O -C -n -hn	1	2.50	180.000	-2.000
H -N -c -o	1	2.50	180.000	-2.000
H -N -c -o	1	2.00	0.000	1.000
CX-C -n -hn	1	2.5	180.000	2.000
O -C -n -c3	1	2.5	180.000	2.000
CX-C -n -hn	1	2.5	180.000	2.000
CX-C -n -c3	1	2.5	180.000	2.000
c3-c -N -H	1	2.5	180.000	2.000
c3-c -N -CX	1	2.5	180.000	2.000
o -c -N -CX	1	2.5	180.000	2.000
o -c -N -H	1	2.5	180.000	2.000

#### IMPROPER

ca-o -no-o	1.1	180.0	2.0
ca-ca-ca-no	1.1	180.0	2.0
ca-ca-ca-ha	1.1	180.0	2.0

ca-ca-ca-ce	1.1	180.0	2.0
ca-cd-ce-ha	1.1	180.0	2.0
c -ce-cd-nd	1.1	180.0	2.0
cd-n -c -o	10.5	180.0	2.0
c3-n -cc-nd	1.1	180.0	2.0
c3-n -c -o	10.5	180.0	2.0
c -c3-n -cc	1.1	180.0	2.0

Amber Input files used in the MD simulations. Minimization was done in two steps, heating in one step and density equilibration in 4 steps. Production was carried out in identical steps of 5 ns.

```

Minimization1
&cntrl
imin=1,
maxcyc=1000,
ncyc=100,
cut=10.0,
ntpr=100,
ntr=1,
restraintmask=':1-226',
restraint_wt=500,
/
Minimization2
&cntrl
imin=1,
maxcyc=1000,
ncyc=200,
cut=10.0,
ntpr=100,
/
Heating
&cntrl
imin=0,
irest=0,
ntx=1,
ntpr=1000,
ntwx=1000,
ntr=1,
restraint_wt=10.0,
restraintmask='?:1-226?',
nstlim=10000,dt=0.002,

```

```

ntc=2,ntf=2,
cut=8.0,
ntt=3,
gamma_ln=2.0,
tempi=0.0,
temp0=300.0,
ig=-1,
/
density1
&cntrl
imin=0,
irest=1,
ntx=7,
nstlim=25000,dt=0.002,
ntc=2,ntf=2,
cut=8.0,
ntp=1,
pres0=1.0,
taup=2.0,
ntpr=1000,
ntwx=1000,
ntt=3,
gamma_ln=2.0,
ioutfm=1,
temp0=300.0,
tempi=300.0,
ig=-1,
ntr=1,
restraintmask=':1-226',
restraint_wt=5.0,
/
density2
&cntrl
imin=0,
irest=1,
ntx=7,
nstlim=25000,dt=0.002,
ntc=2,ntf=2,
cut=8.0,
ntp=1,
pres0=1.0,
taup=2.0,

```

```

ntpr=1000,
ntwx=1000,
ntt=3,
gamma_ln=2.0,
ioutfm=1,
temp0=300.0,
tempi=300.0,
ig=-1,
ntr=1,
restraintmask=':1-226',
restraint_wt=2.0,
/
density3
&cntrl
imin=0,
irest=1,
ntx=7,
nstlim=25000,dt=0.002,
ntc=2,ntf=2,
cut=8.0,
ntp=1,
pres0=1.0,
taup=2.0,
ntpr=1000,
ntwx=1000,
ntt=3,
gamma_ln=2.0,
ioutfm=1,
temp0=300.0,
tempi=300.0,
ig=-1,
ntr=1,
restraintmask=':1-226',
restraint_wt=0.5,
/
density4
&cntrl
imin=0,
irest=1,
ntx=7,
nstlim=25000,dt=0.002,
ntc=2,ntf=2,

```

```
cut=8.0,  
ntp=1,  
pres0=1.0,  
taup=2.0,  
ntpr=1000,  
ntwx=1000,  
ntt=3,  
gamma_ln=2.0,  
ioutfm=1,  
temp0=300.0,  
tempi=300.0,  
ig=-1,  
ntr=1,  
restraintmask=':1-226',  
restraint_wt=0.1,  
/  
Production  
&cntrl  
imin=0,  
irest=1,  
ntx=7,  
nstlim=2500000,dt=0.002,  
ntc=2,ntf=2,  
cut=8.0,  
ntp=1,  
pres0=1.0,  
taup=2.0,  
ntpr=1000,  
ntwx=1000,  
ioutfm=1,  
ntt=3,  
gamma_ln=2.0,  
temp0=300.0,  
tempi=300.0,  
ig=-1,  
ntr=0,  
/  
/
```

**Table B4:** *The  $\delta^{TPA}$  values for the chromophore models as computed at the TDDFT/6-31+(d,p) Level of Theory in PCM (water)*

Model	$\delta^{TPA}$ (a.u.)
1a	42
1b	13
2a	15
2b	20
3	24
4	11
5	180
6	342
7	339
8	370
9	609
10	763
11	1049
12	1358
13	1751
14	1744
15	1879
16a	2046
16b	1882
17	2067
18	2540
19	2908
20	5738
21	5258
22	8834

**Table B5:** Comparison of average bond lengths (in Å) of the control chromophore over the dynamics trajectory (Dynamic Average) to the corresponding bond lengths of the chromophore in the crystal structure (PDB ID: 2Y0G) and to those of the same chromophore as optimized at the PBE0/6-31+(d,p) level of theory in the gas phase (DFT-optimized Structure). Bond lengths differences are given in absolute values. Atom names match the figure in Table B2

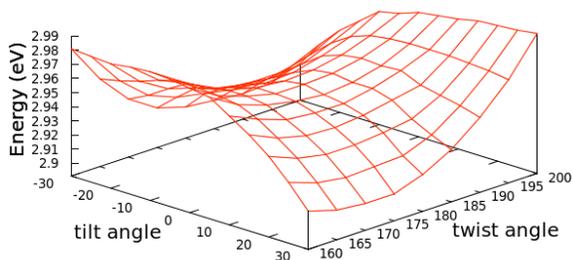
Bond	Dynamics Average (DA)	Crystal Structure (CS)	DFT-optimized Structure (DF)	DA-CS	DA-DF	DF-CS
C1-N3	1.39	1.35	1.38	0.04	0.01	0.03
C2-CA2	1.46	1.46	1.46	0.00	0.01	0.00
C2-O2	1.22	1.30	1.24	0.08	0.02	0.06
CA2-CB2	1.37	1.47	1.39	0.10	0.01	0.08
CA2-N2	1.38	1.37	1.40	0.01	0.01	0.03
CB2-CG2	1.49	1.47	1.41	0.02	0.08	0.06
CD1-CE1	1.39	1.40	1.37	0.01	0.02	0.03
CD2-CG2	1.40	1.40	1.43	0.00	0.03	0.03
CE1-CZ	1.48	1.39	1.45	0.09	0.03	0.06
CE2-CD2	1.39	1.43	1.37	0.04	0.03	0.06
CG2-CD1	1.40	1.43	1.43	0.03	0.03	0.00
CZ-CE2	1.48	1.39	1.46	0.09	0.03	0.07
CZ-OH	1.21	1.39	1.25	0.18	0.04	0.14
N2-C1	1.34	1.38	1.30	0.04	0.05	0.08
N3-C2	1.34	1.34	1.40	0.00	0.07	0.06
maximum difference				0.18	0.08	0.14

**Table B6:** Comparison of average bond lengths (in Å) of the nitro chromophore over the dynamics trajectory (Dynamic Average) to the corresponding bond lengths of the same chromophore (Model 22) as optimized at the PBE0/6-31+(d,p) level of theory in the gas phase (DFT-optimized Structure). Bond lengths differences are given in absolute values. Atom names match the figure in Table B3

Bond	Dynamics Average (DA)	DFT-optimized Structure (DF)	DA-DF
C5-N1	1.38	1.40	0.02
C5-C8	1.38	1.35	0.02
N1-C3	1.34	1.29	0.05
C3-N2	1.39	1.37	0.02
N2-C4	1.34	1.40	0.06
C4-C5	1.47	1.49	0.03
C4-O1	1.22	1.21	0.01
C8-C9	1.49	1.45	0.03
C9-C10	1.39	1.41	0.01
C10-C12	1.40	1.39	0.01
C12-C14	1.40	1.39	0.01
C14-C13	1.40	1.39	0.01
C14-N3	1.50	1.46	0.03
C13-C11	1.40	1.39	0.01
C11-C9	1.40	1.41	0.01
N3-O4	1.22	1.22	0.00
N3-O5	1.22	1.22	0.00
maximum difference			0.06

**Table B7:** Comparison of PCM and Gas phase computations at the TDDFT/6-31+(d,p) Level of Theory for selected chromophore models.

Model	PCM			Gas-phase		
	Energy(eV)	OS	TPA (GM)	Energy(eV)	OS	TPA (GM)
6	3.197	0.91	1	3.336	0.81	0
7	3.200	0.93	1	3.331	0.81	0
12	3.210	0.98	5	3.351	0.86	5
13	3.289	0.85	7	3.444	0.74	7
14	3.310	0.89	7	3.451	0.79	7
16a	3.218	0.58	8	3.348	0.42	8
18	3.137	0.93	9	3.352	0.79	10
19	3.138	0.55	11	3.250	0.33	11
20	2.689	0.36	15	2.826	0.28	17
22	2.965	0.64	29	3.257	0.61	22



**Figure B1:** Variation of Energy for various conformers of the nitro chromophore with the tilt and twist angles while fixing the methine bridge at  $134^\circ$ .

**Table B8:** *The z-matrix used in the conformational analysis of the nitro-chromophore. The entries: meth, tilt and twist refer to the methyl bridge, tilting and twisting angles referred to in the main text.*

C						
N	1	1.39729				
C	2	1.29399	1	105.277		
N	3	1.37369	2	115.035	1	0
C	4	1.39566	3	108.2405	2	0
H	4	1.00778	3	127.2314	5	180
O	5	1.21326	4	126.7683	3	180
H	3	1.08464	2	123.9485	4	180
C	1	1.3532	2	128.768	5	180
H	9	1.08917	1	114.3677	2	180
C	9	1.45109	1	meth	2	tilt
C	11	1.40793	9	117.8361	1	twist
C	12	1.38606	11	121.2143	9	180
C	13	1.39083	12	118.3849	11	0
C	14	1.3927	13	122.0847	12	0
C	15	1.38626	14	119.0858	13	0
H	15	1.08406	14	119.4249	13	180
H	16	1.08407	15	120.4418	14	180
H	12	1.08647	13	119.4459	14	180
H	13	1.08369	14	119.7805	15	180
N	14	1.46318	13	118.8164	12	180
O	21	1.22216	14	117.5846	13	0
O	21	1.22193	14	117.593	13	180

**Table B9:** *Energy, OS, TPA (in GM) and the square of the x-component of the difference between permanent moments of excited and ground states (in atomic units) of various conformers of the nitro chromophore computed at the TDDFT/6-31+(d,p) Level of Theory in the gas phase (while fixing the methine bridge at 134°).*

Tilt	Twist	Energy(eV)	OS	TPA	$(\Delta\mu_x)^2$
-30	160	2.981	0.509	26.62	6.452
-30	165	2.953	0.534	26.25	6.239
-30	170	2.932	0.556	25.78	6.011
-30	175	2.919	0.575	25.21	5.798
-30	180	2.913	0.593	24.55	5.613
-30	185	2.911	0.608	24.12	5.465

-30	190	2.909	0.620	24.16	5.358
-30	195	2.910	0.627	24.32	5.295
-30	200	2.917	0.630	24.56	5.273
-25	160	2.980	0.523	28.63	7.195
-25	165	2.956	0.548	28.12	6.890
-25	170	2.939	0.569	27.55	6.606
-25	175	2.931	0.589	26.82	6.355
-25	180	2.927	0.606	26.21	6.149
-25	185	2.923	0.619	26.07	5.994
-25	190	2.922	0.628	26.15	5.894
-25	195	2.926	0.633	26.34	5.844
-25	200	2.935	0.634	26.57	5.851
-20	160	2.979	0.539	30.11	7.796
-20	165	2.960	0.562	29.53	7.423
-20	170	2.948	0.583	28.78	7.095
-20	175	2.942	0.602	28.07	6.825
-20	180	2.935	0.617	27.78	6.615
-20	185	2.933	0.627	27.74	6.471
-20	190	2.933	0.633	27.86	6.394
-20	195	2.939	0.636	28.01	6.386
-20	200	2.949	0.635	28.32	6.444
-15	160	2.981	0.555	31.12	8.255
-15	165	2.965	0.577	30.36	7.840
-15	170	2.956	0.598	29.65	7.496
-15	175	2.946	0.614	29.25	7.228
-15	180	2.942	0.626	29.06	7.037
-15	185	2.940	0.633	29.07	6.916
-15	190	2.943	0.637	29.17	6.888
-15	195	2.951	0.637	29.40	6.947
-15	200	2.960	0.632	29.79	7.100
-10	160	2.982	0.572	31.65	8.578
-10	165	2.969	0.593	30.96	8.149
-10	170	2.957	0.610	30.46	7.809
-10	175	2.951	0.623	30.16	7.564
-10	180	2.947	0.632	30.02	7.414
-10	185	2.948	0.637	30.00	7.364
-10	190	2.953	0.638	30.09	7.418
-10	195	2.960	0.634	30.40	7.604

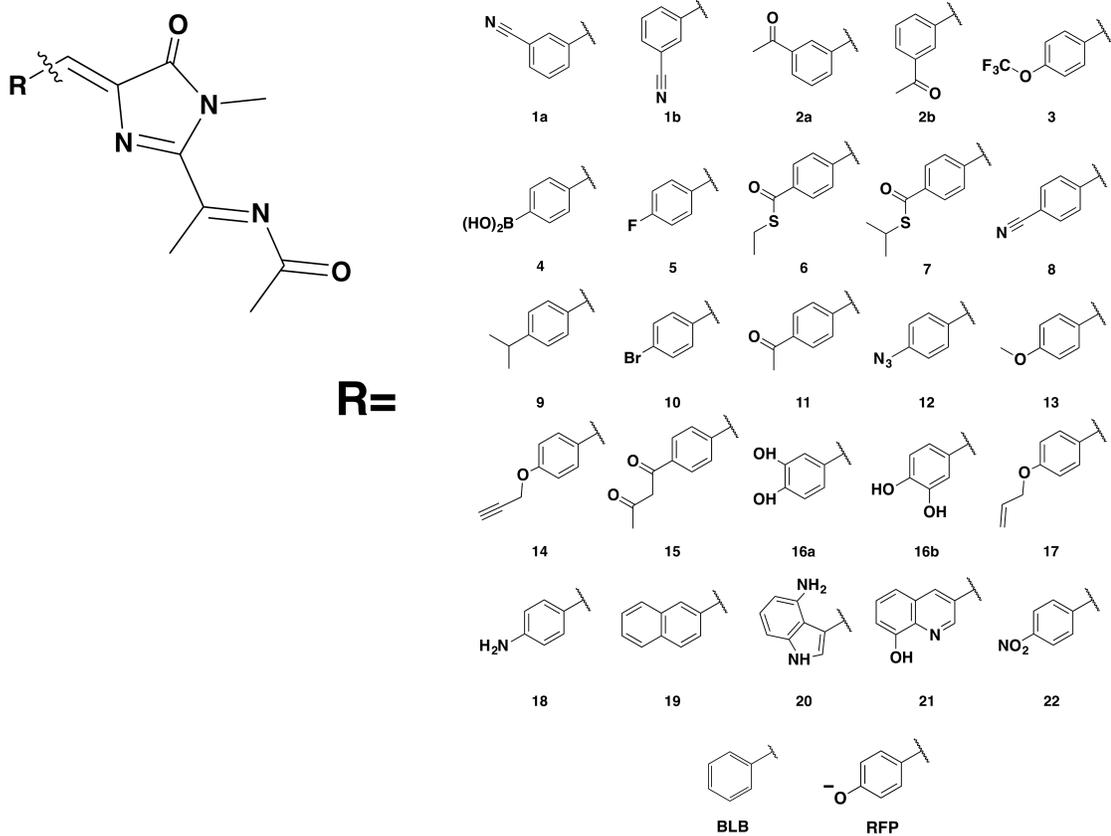
-10	200	2.971	0.626	30.73	7.921
-5	160	2.981	0.587	32.02	8.790
-5	165	2.968	0.606	31.35	8.378
-5	170	2.960	0.622	30.84	8.067
-5	175	2.955	0.632	30.55	7.870
-5	180	2.953	0.638	30.44	7.791
-5	185	2.954	0.639	30.48	7.842
-5	190	2.958	0.635	30.75	8.019
-5	195	2.965	0.627	31.15	8.339
-5	200	2.974	0.614	31.64	8.731
0	160	2.979	0.602	31.92	8.914
0	165	2.967	0.619	31.31	8.516
0	170	2.960	0.630	30.95	8.253
0	175	2.955	0.636	30.78	8.095
0	180	2.955	0.639	30.65	8.038
0	185	2.956	0.637	30.72	8.095
0	190	2.960	0.629	31.03	8.253
0	195	2.967	0.617	31.44	8.516
0	200	2.979	0.602	31.88	8.914
5	160	2.977	0.616	31.42	8.731
5	165	2.965	0.628	31.01	8.339
5	170	2.958	0.635	30.69	8.019
5	175	2.954	0.639	30.52	7.842
5	180	2.952	0.637	30.52	7.791
5	185	2.954	0.631	30.65	7.870
5	190	2.959	0.620	30.94	8.067
5	195	2.968	0.606	31.32	8.378
5	200	2.981	0.588	31.84	8.790
10	160	2.968	0.624	30.99	7.921
10	165	2.959	0.633	30.51	7.604
10	170	2.952	0.638	30.15	7.418
10	175	2.950	0.638	29.93	7.364
10	180	2.950	0.634	29.81	7.414
10	185	2.954	0.625	29.94	7.564
10	190	2.960	0.611	30.34	7.809
10	195	2.968	0.594	30.87	8.149
10	200	2.983	0.574	31.39	8.578

15	160	2.959	0.631	29.93	7.100
15	165	2.950	0.637	29.44	6.947
15	170	2.945	0.638	29.09	6.888
15	175	2.944	0.635	28.85	6.916
15	180	2.946	0.627	28.85	7.037
15	185	2.950	0.614	29.11	7.228
15	190	2.954	0.600	29.58	7.496
15	195	2.968	0.582	30.01	7.840
15	200	2.986	0.560	30.59	8.255
20	160	2.948	0.635	28.42	6.444
20	165	2.940	0.637	27.97	6.386
20	170	2.936	0.635	27.65	6.394
20	175	2.936	0.628	27.56	6.471
20	180	2.938	0.617	27.65	6.615
20	185	2.941	0.604	28.02	6.825
20	190	2.951	0.588	28.36	7.095
20	195	2.967	0.568	28.87	7.423
20	200	2.987	0.544	29.58	7.796
25	160	2.934	0.635	26.58	5.851
25	165	2.927	0.634	26.20	5.844
25	170	2.924	0.628	26.01	5.894
25	175	2.925	0.619	25.96	5.994
25	180	2.926	0.608	26.20	6.149
25	185	2.934	0.593	26.46	6.355
25	190	2.948	0.575	26.90	6.606
25	195	2.966	0.554	27.47	6.890
25	200	2.988	0.528	28.08	7.195
30	160	2.918	0.631	24.44	5.273
30	165	2.912	0.627	24.15	5.295
30	170	2.910	0.619	24.01	5.358
30	175	2.910	0.610	24.15	5.465
30	180	2.917	0.597	24.28	5.613
30	185	2.928	0.581	24.55	5.798
30	190	2.944	0.562	24.99	6.011
30	195	2.964	0.539	25.46	6.239
30	200	2.990	0.513	25.98	6.452

---

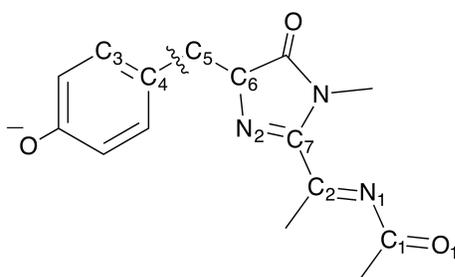
# Appendix C

## Appendix to Chapter 4



**Figure C1:** Chromophore models built from the RFP-parent structure. RFP and BLB refer to the DsRed and mBlueberry chromophore models, respectively. The rest are built via the replacement of the phenol ring of Tyr-66 with the corresponding moiety in a non-canonical amino acid. The numbering matches the corresponding previously studies models.<sup>75</sup>

**Table C1:** *Tilt, twist and acylimine ( $\theta_{\text{acylimine}}$ ) dihedral angles for red fluorescent chromophores in their native protein crystal structures. Also included are the two other dihedral angles ( $\theta_1$  and  $\theta_2$ ) defining the geometry of the extended conjugation chain.*



Protein Name (PDB ID)	Tilt Angle (in degree) (C <sub>4</sub> - C <sub>5</sub> - C <sub>6</sub> - N <sub>2</sub> )	Twist Angle (in degree) (C <sub>3</sub> - C <sub>4</sub> - C <sub>5</sub> - C <sub>6</sub> )	$\theta_{\text{acylimine}}$ (in degree) (O <sub>1</sub> - C <sub>1</sub> - N <sub>1</sub> - C <sub>2</sub> )	$\theta_1$ (in degree) (C <sub>1</sub> - N <sub>1</sub> - C <sub>2</sub> - C <sub>7</sub> )	$\theta_2$ (in degree) (N <sub>1</sub> - C <sub>2</sub> - C <sub>7</sub> - N <sub>2</sub> )
FP583 (1GGX)	2	182	145.6	114.5	191.5
DsRED (1G7K)	0.03	181.8	176.4	91.3	183.5
K83M (2H8Q)	-3.4	186.9	181.4	104.6	164.8
Rtms5 H146S (2P4M)	-2	212.6	133.5	111.9	178.0
DsRED.M1(2VAD)	7.4	180	210.7	110.2	175.8
KillerRed (2WIQ)	27.9	168.2	37.5	204.1	229.8
mKeima(pH=8) (2WHU)	-19.4	217.4	121.2	139.1	180.8
mKeima(pH=3.8) (2WHS)	-1.3	179.4	144.4	116.4	180.2
MKEIMA(pH=7)-3IR8	1.3	173.2	168.6	108.2	173.1
KillerRed (3A8S)	28.4	156.4	30.2	181.6	252.5
KillerRed (3GB3)	24.8	164.3	45.2	197.1	222.5
Azami-Green (3ADF)	31	152.6	15.7	228.4	249.5
SuperNova (3WCK)	3.9	172.5	44.6	184.7	224.8
Favina Proteins Ancestor (4DXI)	2.6	178.7	349.6	240.8	253.0
eCGP123 (4TZG)	7.9	167.4	358.9	249.4	235.5
Blue Chromoprotein sgBP (4ZB1)	-10.2	220.6	52.7	199.9	218.4
mPlum (2QLG)	-3.2	185.8	3.5	222.1	256.1
mCHERRY (2H5Q)	26.2	164.3	119.7	140.4	179.1
TagBFP (3M24)	5.2	178.8	168.4	133.8	177.2
Neptune (3IP2)	27.9	168.2	37.5	284.3	229.8
EQFP650 (4EDO)	5.2	178.8	168.4	133.8	177.2
Average	7.7	179.5	129.2	166.5	206.3

**Table C2:** *Coordinates of the optimized models at the PBE0/6-31+G(d,p) level of theory in the gas phase. All chromophores are capped with methyl groups (see Figure C1) except where noted.*

Coordinates for BLB Model			
C	-5.40977	0.15799	0.05968
C	-4.31665	0.97182	0.04862
C	-5.28268	-1.29241	0.04480
C	-3.91326	-1.79368	0.01531
C	-2.97609	0.46139	0.01704
C	-2.82982	-0.96666	0.00656
C	-1.91112	1.36417	0.00275
C	-0.52636	1.17007	-0.02172
C	1.44199	0.28769	-0.03840
C	0.41421	2.28938	-0.02118
C	2.46489	-0.73065	-0.05701
C	4.75710	-1.28481	-0.05084
C	1.98998	-2.15043	-0.21078
C	5.61207	-1.40930	1.19122
C	2.89133	2.40652	-0.06476
N	0.15647	-0.01265	-0.02929
N	1.65951	1.65890	-0.03568
N	3.70984	-0.40288	0.08555
H	-6.27139	-2.05291	0.06118
H	-6.41872	0.56307	0.06585
H	-4.44864	2.05380	0.04444
H	-3.79286	-2.87505	0.00942
H	-1.82169	-1.37362	-0.01024
H	-2.17809	2.42210	0.01818
H	1.05738	-2.29372	0.33679
H	1.78473	-2.36344	-1.26435
H	2.75272	-2.85485	0.12453
H	5.01525	-1.82646	2.01032
H	6.45912	-2.06780	0.98887
H	5.96338	-0.42665	1.51645
H	3.49391	2.23148	0.82853
H	3.48910	2.13282	-0.93934
H	2.60194	3.45933	-0.12559
O	0.24183	3.51189	-0.01863
O	5.03740	-1.85747	-1.09128
Coordinates for RFP model			
C	-5.40977	0.15799	0.05968
C	-4.31665	0.97182	0.04862
C	-5.28268	-1.29241	0.04480
C	-3.91326	-1.79368	0.01531
C	-2.97609	0.46139	0.01704
C	-2.82982	-0.96666	0.00656
C	-1.91112	1.36417	0.00275
C	-0.52636	1.17007	-0.02172
N	0.15647	-0.01265	-0.02929

C	1.44199	0.28769	-0.03840
C	0.41422	2.28938	-0.02118
N	1.65951	1.65890	-0.03568
O	-6.27139	-2.05291	0.06118
O	0.24183	3.51189	-0.01863
H	-6.41872	0.56307	0.06585
H	-4.44864	2.05380	0.04444
H	-3.79286	-2.87505	0.00942
H	-1.82169	-1.37362	-0.01024
H	-2.17809	2.42210	0.01818
C	2.46489	-0.73065	-0.05701
N	3.70984	-0.40288	0.08555
C	4.75710	-1.28481	-0.05084
O	5.03740	-1.85747	-1.09128
C	1.98998	-2.15043	-0.21078
H	1.05738	-2.29372	0.33679
H	1.78473	-2.36344	-1.26435

---

Coordinates for RFP model with H-capping

C	4.72554	0.26170	0.06860
C	3.60206	1.03307	0.08040
C	4.65429	-1.19095	-0.00045
C	3.30520	-1.74290	-0.05274
C	2.28205	0.47291	0.03084
C	2.19095	-0.95790	-0.04215
C	1.18311	1.33387	0.05182
C	-0.19303	1.08811	0.01061
C	-2.12492	0.13468	-0.09900
C	-1.17559	2.17051	0.03351
C	-3.10755	-0.91989	-0.17647
C	-5.37470	-1.55642	-0.30392
N	-0.82930	-0.11739	-0.07329
N	-2.39517	1.49501	-0.03583
N	-4.35801	-0.63049	-0.35040
H	5.71755	0.70318	0.12460
H	3.69175	2.11723	0.14864
H	3.22668	-2.82659	-0.11044
H	1.19945	-1.40181	-0.08760
H	1.40926	2.40008	0.10455
H	-2.72537	-1.94774	-0.10702
H	-5.96769	-1.62015	-1.25026
H	-3.27727	1.98448	-0.02137
O	5.67166	-1.91263	-0.01882
O	-1.05076	3.39675	0.10428
O	-5.67160	-2.19616	0.69188

---

Coordinates for Model Number 1a

C	-3.39536	-2.24899	0.00832
C	-2.38476	-1.29604	0.00383
C	-2.69784	0.07763	0.01984
C	-1.69458	1.12506	0.01365
C	-4.04789	0.45939	0.03967

C	-5.05922	-0.50410	0.04348
C	-4.73367	-1.86726	0.02798
C	0.54156	2.21786	-0.02553
C	1.67851	0.29258	-0.05799
C	-0.34308	1.02125	-0.00984
C	3.01021	2.49545	-0.12976
C	-6.43229	-0.09646	0.06307
C	2.79654	-0.67344	-0.08351
C	5.12373	-1.02576	-0.02268
C	2.41615	-2.11169	-0.27472
C	5.90687	-1.12135	1.25818
N	1.82478	1.66619	-0.05440
N	0.44018	-0.12246	-0.03291
N	-7.54670	0.23112	0.07907
N	3.98215	-0.22171	0.06579
H	-1.34210	-1.59223	-0.01399
H	-4.31192	1.51275	0.05140
H	-2.06392	2.14989	0.02544
H	-5.52553	-2.60920	0.03069
H	-3.14042	-3.30428	-0.00475
H	2.00870	-2.25115	-1.28109
H	3.27967	-2.76704	-0.15972
H	1.62689	-2.38803	0.42822
H	5.27581	-1.50218	2.06812
H	6.76558	-1.77696	1.11039
H	6.24813	-0.12530	1.55754
H	3.63104	2.37615	0.75953
H	2.65294	3.52534	-0.19963
H	3.60407	2.24485	-1.01029
O	0.27342	3.40375	-0.02222
O	5.48338	-1.51150	-1.07628

---

Coordinates for Model Number 1b

C	3.96281	1.34801	0.01766
C	2.81669	0.55193	-0.00328
C	2.93512	-0.84748	0.03093
C	1.79153	-1.73899	0.01684
C	4.22180	-1.41385	0.08487
C	5.35946	-0.61522	0.10469
C	5.24045	0.76903	0.07156
C	-0.57814	-2.49865	-0.02334
C	-1.42468	-0.42990	-0.08356
C	0.46984	-1.44205	-0.02393
C	-3.06143	-2.41890	-0.09436
C	3.83063	2.77479	-0.01394
C	-2.39191	0.68730	-0.11730
C	-4.65147	1.35995	-0.03735
C	-1.81032	2.05578	-0.31145
C	-5.36869	1.59240	1.26469
N	-1.76769	-1.76813	-0.05878
N	-0.13954	-0.19779	-0.06707

N	3.72564	3.93148	-0.03766
N	-3.62853	0.40934	0.03938
H	1.83237	1.00414	-0.04439
H	4.32153	-2.49577	0.11202
H	6.34326	-1.07210	0.14603
H	2.00822	-2.80599	0.04539
H	6.11964	1.40487	0.08712
H	-1.26766	2.09711	-1.26012
H	-2.59110	2.81608	-0.32045
H	-1.08072	2.26835	0.47459
H	-4.65851	1.85103	2.05717
H	-6.10187	2.38941	1.13817
H	-5.87450	0.67159	1.57330
H	-3.63244	-2.10607	-0.97010
H	-3.63868	-2.18971	0.80265
H	-2.85809	-3.49104	-0.14331
O	-0.48445	-3.71096	-0.00087
O	-4.98176	1.86052	-1.09336

---

Coordinates for Model Number 2a

C	-4.64457	-0.44420	0.02937
C	-3.61598	0.49966	0.03303
C	-2.26877	0.11193	0.02470
C	-1.25696	1.15034	0.02193
C	-1.96407	-1.26548	0.01366
C	-2.98353	-2.20649	0.01121
C	-4.31921	-1.80520	0.01841
C	0.99099	2.22340	-0.01475
C	2.10962	0.28583	-0.03569
C	0.09510	1.03721	0.00629
C	3.46230	2.47552	-0.11688
C	-6.05816	0.04979	0.03611
C	-7.18109	-0.95800	0.02106
C	3.21495	-0.69300	-0.05970
C	5.53806	-1.07511	-0.06113
C	2.81287	-2.13221	-0.19433
C	6.36038	-1.14699	1.19684
N	2.26925	1.65844	-0.04135
N	0.86716	-0.11485	-0.00725
N	4.41105	-0.25440	0.04031
H	-3.88768	1.55192	0.04096
H	-0.92408	-1.57139	0.00465
H	-5.10135	-2.55823	0.01491
H	-7.12543	-1.59318	-0.86971
H	-7.13057	-1.61352	0.89732
H	-8.13080	-0.42267	0.02478
H	-2.73916	-3.26471	0.00201
H	2.03590	-2.37512	0.53400
H	2.38211	-2.30006	-1.18634
H	3.67117	-2.79382	-0.07588
H	5.75479	-1.51838	2.03033

H	7.21644	-1.80234	1.03362
H	6.70543	-0.14529	1.47124
H	-1.62089	2.17718	0.02703
H	4.08774	2.34250	0.76740
H	3.11477	3.50953	-0.17549
H	4.04877	2.22701	-1.00306
O	0.73763	3.41323	-0.01649
O	-6.29347	1.24506	0.05287
O	5.86015	-1.59514	-1.11098

---

Coordinates for Model Number 2b			
C	-4.99835	1.31226	0.09638
C	-3.79492	2.00518	0.08603
C	-2.56456	1.32074	0.05636
C	-1.34359	2.10372	0.04179
C	-2.58097	-0.08329	0.04246
C	-3.78700	-0.78360	0.05338
C	-4.99877	-0.08062	0.07804
C	1.09435	2.62753	-0.01479
C	1.72382	0.48214	-0.09169
C	-0.05550	1.68346	-0.00875
C	3.55570	2.28973	-0.13345
C	-3.72859	-2.28150	0.03625
C	-5.02073	-3.05969	-0.01742
C	2.57136	-0.72728	-0.13952
C	4.74083	-1.63666	-0.00423
C	1.86253	-2.01511	-0.43191
C	5.35363	-2.01689	1.31674
N	2.20259	1.77915	-0.06364
N	0.42339	0.38424	-0.06362
N	3.82262	-0.58640	0.07689
H	-5.93790	1.85611	0.11832
H	-3.79673	3.09237	0.09914
H	-1.65090	-0.63982	0.02044
H	-1.46151	3.18636	0.07275
H	-5.94471	-0.61330	0.08580
H	-5.60042	-2.79182	-0.90735
H	-5.64456	-2.84586	0.85745
H	-4.78748	-4.12438	-0.04229
H	1.51917	-2.01259	-1.47163
H	2.52373	-2.87079	-0.29213
H	0.96943	-2.11120	0.18974
H	4.57465	-2.26711	2.04433
H	6.02586	-2.86369	1.17569
H	5.91062	-1.16579	1.72146
H	4.07998	1.88521	-1.00075
H	4.11860	2.03646	0.76663
H	3.46539	3.37478	-0.22195
O	1.12774	3.84387	0.00971
O	-2.65546	-2.85773	0.06060
O	5.07677	-2.11402	-1.06969

---

---

Coordinates for Model Number 3			
C	-2.78881	-0.72597	-0.58073
C	-1.56040	-0.10071	-0.42703
C	-1.48609	1.29319	-0.23526
C	-0.23980	2.01209	-0.07092
C	-2.68321	2.03217	-0.20614
C	-3.91651	1.41345	-0.35689
C	-3.95210	0.03641	-0.53577
C	2.20032	2.45019	0.13579
C	2.78621	0.30255	-0.08088
C	1.03408	1.54668	-0.04853
C	4.65312	2.05038	0.21692
C	-5.83575	-1.09221	0.29766
C	3.60899	-0.92146	-0.16139
C	5.74686	-1.89498	0.02720
C	2.89082	-2.17030	-0.57745
C	6.33898	-2.32785	1.34091
N	3.28921	1.57575	0.10710
N	1.48518	0.24201	-0.17854
N	4.84855	-0.82924	0.13462
H	-2.85798	-1.79700	-0.73839
H	-0.64313	-0.67782	-0.45884
H	-2.64032	3.10889	-0.06537
H	-4.84105	1.98038	-0.34007
H	-0.31840	3.09114	0.05707
H	1.97897	-2.29423	0.01134
H	2.58274	-2.08082	-1.62424
H	3.53285	-3.04617	-0.48265
H	5.54945	-2.56763	2.06041
H	6.98327	-3.19300	1.18244
H	6.92345	-1.50601	1.76695
H	5.13128	1.66144	1.11701
H	4.59052	3.13985	0.26705
H	5.24198	1.75004	-0.65165
O	2.25950	3.65725	0.27739
O	-5.18039	-0.58468	-0.76626
O	6.08337	-2.34564	-1.04972
F	-6.17038	-0.15066	1.19547
F	-5.11830	-2.01848	0.95478
F	-6.94999	-1.66157	-0.14824

---

Coordinates for Model Number 4			
C	-3.40483	-1.09461	-0.00796
C	-2.20535	-0.39886	-0.01843
C	-2.20298	1.00968	0.01897
C	-0.99879	1.81291	0.01297
C	-3.43981	1.67890	0.06604
C	-4.63131	0.96853	0.07540
C	-4.64466	-0.43357	0.03860
C	1.41230	2.42239	-0.01945
C	2.13256	0.30602	-0.08302

C	0.30354	1.43241	-0.02373
C	3.88553	2.19365	-0.09625
C	3.03190	-0.86445	-0.11972
C	5.24075	-1.67889	-0.02186
C	2.37496	-2.19119	-0.35847
C	5.92978	-1.97551	1.28274
N	2.55605	1.62189	-0.05480
N	0.83581	0.15323	-0.06793
N	4.28116	-0.66644	0.06092
H	-3.39601	-2.18059	-0.03721
H	-1.25784	-0.92453	-0.05625
H	-3.45381	2.76571	0.09599
H	-5.56159	1.53362	0.11448
H	-1.14177	2.89283	0.04250
H	-7.25382	0.23509	0.07351
H	-6.74280	-3.05228	0.03475
H	1.50999	-2.30633	0.29852
H	1.99957	-2.23255	-1.38611
H	3.07853	-3.01191	-0.21712
H	5.19643	-2.22790	2.05595
H	6.63219	-2.79803	1.14537
H	6.46504	-1.08468	1.62657
H	4.44943	1.94282	0.80389
H	3.74518	3.27531	-0.15782
H	4.43809	1.83794	-0.96767
O	1.39343	3.63931	0.00416
O	-7.22999	-0.72445	0.06844
O	-5.87730	-2.63045	0.03046
O	5.54771	-2.18809	-1.08148
B	-5.97283	-1.26971	0.04676

---

Coordinates for Model Number 5

C	3.41705	-0.98794	-0.01261
C	2.15538	-0.41733	-0.08848
C	2.00652	0.98162	-0.01604
C	0.72533	1.65212	-0.07920
C	3.16132	1.77311	0.13490
C	4.41863	1.19055	0.20777
C	4.57210	-0.20190	0.13603
C	-1.73119	1.99533	-0.26273
C	-2.19776	-0.18123	-0.49263
C	-0.51964	1.13787	-0.24658
C	-4.14528	1.48583	-0.55626
C	-2.93190	-1.46501	-0.67154
C	-4.58537	-1.10906	0.99740
C	-2.29044	-2.42405	-1.62250
C	-6.08698	-1.03324	0.92449
N	-2.77019	1.07222	-0.41401
N	-0.89678	-0.18694	-0.42684
N	-4.01014	-1.79929	-0.07330
H	3.51869	-2.06844	-0.06565

H	1.27141	-1.03516	-0.19790
H	3.05821	2.85399	0.19503
H	5.29889	1.81656	0.32425
H	0.74288	2.73669	0.02554
H	7.93103	-0.46526	0.42108
H	6.93863	-2.61291	0.21218
H	-1.23120	-2.52880	-1.36987
H	-2.79943	-3.38709	-1.57593
H	-2.33166	-2.03292	-2.64577
H	-6.45774	-0.30663	1.64829
H	-6.42890	-0.78809	-0.08498
H	-6.49364	-2.02026	1.17024
H	-4.65933	0.86637	-1.29511
H	-4.13367	2.52353	-0.89769
H	-4.67265	1.44209	0.40058
O	-1.86697	3.20297	-0.19206
O	7.07292	-0.03751	0.36495
O	6.06020	-2.22860	0.15671
O	-3.94127	-0.73492	1.95918
B	5.98675	-0.86170	0.22446

---

Coordinates for Model Number 6

C	-2.53351	-0.16433	0.03024
C	-1.19615	0.19421	0.08198
C	-0.81888	1.54848	-0.00822
C	0.55646	1.99760	0.04050
C	-1.82825	2.51888	-0.15773
C	-3.16440	2.15745	-0.20804
C	-3.53019	0.80996	-0.11170
C	3.03333	1.93407	0.22510
C	3.13715	-0.28428	0.48810
C	1.69647	1.28442	0.21575
C	5.33010	1.04192	0.54399
C	-4.98411	0.48718	-0.16779
C	-7.22062	-1.09423	-0.12343
C	-7.85029	-2.47390	-0.01855
C	3.66471	-1.66344	0.69397
C	5.31481	-1.59201	-1.01070
C	2.90214	-2.48498	1.68257
C	6.81180	-1.73648	-0.96898
N	3.90642	0.85678	0.39659
N	1.85206	-0.08149	0.41518
N	4.66637	-2.17161	0.08612
H	-2.80121	-1.21486	0.09734
H	-0.42663	-0.56171	0.18811
H	-1.55009	3.56696	-0.23333
H	-3.94553	2.90223	-0.32178
H	0.72044	3.06823	-0.07817
H	-7.57168	-0.42831	0.66910
H	-7.44477	-0.61586	-1.08013
H	-7.49865	-3.13920	-0.81284

H	-7.63041	-2.94519	0.94420
H	-8.93752	-2.38679	-0.10921
H	1.83832	-2.45991	1.42761
H	3.27651	-3.50884	1.67958
H	2.99350	-2.05856	2.68806
H	7.06541	-2.77789	-1.19465
H	7.26759	-1.08916	-1.71891
H	7.20903	-1.51986	0.02679
H	5.73621	0.33020	1.26606
H	5.48723	2.05942	0.90926
H	5.84485	0.93692	-0.41504
O	3.36354	3.10133	0.13717
O	-5.83686	1.33855	-0.32701
O	4.70918	-1.14427	-1.96555
S	-5.41370	-1.23272	0.02012

---

Coordinates for Model Number 7

C	-2.25135	-0.05028	-0.05208
C	-0.93671	0.38749	-0.04951
C	-0.64571	1.76608	-0.02152
C	0.69835	2.30314	-0.01653
C	-1.71871	2.67778	0.00365
C	-3.03144	2.23747	0.00121
C	-3.31114	0.86611	-0.02787
C	3.18388	2.39836	-0.02692
C	3.44903	0.17858	-0.06615
C	1.89292	1.65963	-0.03388
C	5.55635	1.66024	-0.08711
C	-4.74601	0.45921	-0.03127
C	-6.89080	-1.29366	-0.09431
C	-7.39584	-1.18545	1.33740
C	-7.38043	-2.55459	-0.79589
C	4.08686	-1.15340	-0.08322
C	6.08542	-2.39966	0.01984
C	3.16789	-2.32000	-0.29012
C	6.74434	-2.78834	1.31541
N	4.13658	1.37736	-0.04566
N	2.14813	0.29798	-0.06154
N	5.35127	-1.21242	0.09057
H	-2.45306	-1.11753	-0.07029
H	-0.11864	-0.32347	-0.06880
H	-1.50761	3.74394	0.02498
H	-3.86039	2.93764	0.02106
H	0.78140	3.38928	0.00351
H	-7.19168	-0.40297	-0.65522
H	-7.06133	-2.03805	1.93753
H	-7.04893	-0.26372	1.81069
H	-8.49223	-1.17511	1.33965
H	-7.02930	-2.60564	-1.83035
H	-7.04732	-3.45846	-0.27403
H	-8.47564	-2.56227	-0.80469

H	2.72324	-2.26365	-1.28864
H	3.70511	-3.26462	-0.20374
H	2.34256	-2.27946	0.42496
H	6.00802	-2.84581	2.12362
H	7.24837	-3.74737	1.19340
H	7.47359	-2.02237	1.59855
H	6.05674	1.28701	0.80766
H	5.64435	2.74780	-0.13969
H	6.01924	1.20519	-0.96458
O	3.41749	3.59224	-0.01323
O	-5.64845	1.27117	0.03652
O	6.25111	-2.99107	-1.02853
S	-5.05956	-1.29154	-0.15704

---

Coordinates for Model Number 8

C	-3.72742	-1.43414	-0.02592
C	-2.58476	-0.65225	-0.03692
C	-2.67939	0.75271	0.01418
C	-1.52672	1.62953	0.00900
C	-3.95731	1.34027	0.07520
C	-5.10461	0.56393	0.08739
C	-4.99419	-0.83230	0.03708
C	0.84785	2.37110	-0.02036
C	1.68285	0.29867	-0.08382
C	-0.20656	1.32040	-0.02798
C	3.33114	2.27800	-0.08429
C	-6.17154	-1.64452	0.05049
C	2.64323	-0.82395	-0.11602
C	4.89701	-1.51427	-0.03665
C	2.05271	-2.18913	-0.30869
C	5.62715	-1.73639	1.25984
N	2.03375	1.63414	-0.05427
N	0.39580	0.07340	-0.07256
N	-7.12733	-2.30558	0.06219
N	3.88186	-0.55513	0.04054
H	-3.65251	-2.51602	-0.06598
H	-1.60386	-1.11126	-0.08409
H	-4.04247	2.42277	0.11477
H	-6.08534	1.02558	0.13482
H	-1.73101	2.69898	0.04176
H	1.32155	-2.39407	0.47791
H	1.51110	-2.22674	-1.25813
H	2.82879	-2.95428	-0.31650
H	4.92413	-1.97817	2.06394
H	6.35185	-2.54144	1.13543
H	6.14451	-0.81622	1.55048
H	3.90503	2.04275	0.81336
H	3.13361	3.35137	-0.12984
H	3.90250	1.96548	-0.95994
O	0.76082	3.58361	0.00547
O	5.21115	-2.03051	-1.09019

---

---

Coordinates for Model Number 9			
C	-3.27777	-1.21481	-0.06882
C	-2.09569	-0.48861	-0.07750
C	-2.12625	0.91757	-0.02110
C	-0.94250	1.74632	-0.02360
C	-3.38330	1.54910	0.04302
C	-4.55787	0.81228	0.05365
C	-4.52863	-0.58745	-0.00255
C	1.45689	2.40532	-0.04345
C	2.22061	0.30246	-0.09612
C	0.36988	1.39534	-0.05486
C	3.93493	2.22385	-0.09815
C	-5.80101	-1.40568	0.01736
C	-6.72511	-1.06077	-1.15467
C	-6.53579	-1.26690	1.35552
C	3.14194	-0.84956	-0.11778
C	5.36025	-1.62808	0.01825
C	2.51616	-2.18866	-0.37292
C	6.04595	-1.90222	1.32981
N	2.61684	1.62688	-0.06662
N	0.92671	0.12604	-0.09299
N	4.38474	-0.63093	0.08645
H	-1.13656	-0.99158	-0.12882
H	-3.42698	2.63481	0.08690
H	-1.10897	2.82294	0.00714
H	-5.51025	1.33339	0.10695
H	-5.50580	-2.45792	-0.09121
H	-6.20829	-1.16465	-2.11369
H	-7.59586	-1.72479	-1.16037
H	-7.09345	-0.03148	-1.08219
H	-5.88644	-1.53262	2.19548
H	-6.88209	-0.23915	1.51170
H	-7.41287	-1.92240	1.37959
H	-3.23498	-2.30078	-0.11540
H	1.61192	-2.30346	0.22820
H	2.21236	-2.25267	-1.42310
H	3.21889	-2.99844	-0.17321
H	5.31143	-2.15075	2.10307
H	6.75648	-2.71992	1.20622
H	6.57162	-1.00220	1.66425
H	4.50208	1.87670	-0.96378
H	4.49598	1.98615	0.80732
H	3.77433	3.30244	-0.16418
O	1.41621	3.62271	-0.02305
O	5.68286	-2.14374	-1.03402

---

Coordinates for Model Number 10			
C	-3.02655	-0.94193	-0.01198
C	-1.81556	-0.26551	-0.01829
C	-1.77777	1.14227	0.02083
C	-0.55244	1.91084	0.01689

C	-2.99830	1.84054	0.06552
C	-4.21543	1.17356	0.07233
C	-4.21915	-0.21900	0.03352
C	1.87709	2.44398	-0.01818
C	2.52929	0.30526	-0.08040
C	0.73766	1.49059	-0.01883
C	4.34158	2.13470	-0.10732
C	3.39062	-0.89337	-0.11566
C	5.57675	-1.77345	-0.03083
C	2.68848	-2.20122	-0.33067
C	6.26338	-2.08150	1.27237
N	2.99447	1.60641	-0.05616
N	1.22798	0.19451	-0.06172
N	4.64714	-0.73338	0.05039
H	-3.04996	-2.02604	-0.04305
H	-0.88261	-0.81707	-0.05498
H	-2.98886	2.92716	0.09557
H	-5.14862	1.72514	0.10753
H	-0.66368	2.99435	0.04596
H	1.87117	-2.30859	0.38674
H	2.23688	-2.21465	-1.32738
H	3.38162	-3.03890	-0.25180
H	5.52668	-2.33008	2.04386
H	6.95648	-2.91138	1.13206
H	6.80738	-1.19792	1.62095
H	4.90048	1.87570	0.79376
H	4.23639	3.21969	-0.17963
H	4.87883	1.75144	-0.97634
O	1.89915	3.66082	0.00404
O	5.86305	-2.29778	-1.08882
BR	-5.86775	-1.14565	0.04214

---

Coordinates for Model Number 11

C	-3.41353	-1.10410	-0.01126
C	-2.21928	-0.39868	-0.01668
C	-2.22733	1.00925	0.01322
C	-1.02520	1.81597	0.01149
C	-3.46961	1.67371	0.04764
C	-4.65814	0.96450	0.05203
C	-4.64515	-0.43557	0.02291
C	1.38629	2.42634	-0.01432
C	2.10731	0.31104	-0.07956
C	0.27707	1.43549	-0.02136
C	3.85996	2.19898	-0.08705
C	-5.95592	-1.15644	0.02987
C	-5.95995	-2.66511	0.00074
C	3.00554	-0.86069	-0.11643
C	5.21943	-1.67196	-0.04315
C	2.34257	-2.19206	-0.31142
C	5.92575	-1.94831	1.25644
N	2.53065	1.62601	-0.04949

N	0.81021	0.15691	-0.06615
N	4.25764	-0.66088	0.03776
H	-3.38203	-2.18923	-0.03485
H	-1.26834	-0.91843	-0.04443
H	-3.48752	2.76043	0.07102
H	-5.61829	1.47027	0.07808
H	-1.17074	2.89537	0.04025
H	-5.42639	-3.07624	0.86471
H	-6.99326	-3.01280	0.01569
H	-5.46417	-3.04139	-0.90087
H	1.61004	-2.36376	0.48184
H	1.78992	-2.19538	-1.25514
H	3.07847	-2.99599	-0.33173
H	5.20356	-2.19513	2.04184
H	6.63036	-2.76911	1.12030
H	6.46098	-1.05091	1.58252
H	4.42470	1.94035	0.81027
H	3.71877	3.28098	-0.13847
H	4.41173	1.85170	-0.96220
O	1.36630	3.64241	0.01114
O	-7.00168	-0.53083	0.05849
O	5.51396	-2.19432	-1.09968

---

Coordinates for Model Number 12

C	-3.55823	-0.91678	0.00984
C	-2.35054	-0.24070	-0.00255
C	-2.31023	1.16826	0.03134
C	-1.08600	1.93244	0.02019
C	-3.53443	1.86432	0.07721
C	-4.74593	1.19641	0.08981
C	-4.76402	-0.20188	0.05627
C	1.34605	2.45793	-0.02623
C	1.99196	0.31635	-0.08001
C	0.20481	1.50919	-0.01804
C	3.80933	2.13941	-0.12416
C	2.84905	-0.88441	-0.11268
C	5.03000	-1.77427	-0.02993
C	2.14459	-2.19072	-0.33036
C	5.72321	-2.08236	1.26991
N	2.46078	1.61620	-0.06391
N	0.68985	0.21145	-0.05599
N	-6.09298	-2.03637	0.04296
N	-6.03319	-0.80581	0.07188
N	-6.28925	-3.15412	0.02023
N	4.10662	-0.72936	0.05366
H	-3.56536	-2.00345	-0.01759
H	-1.41711	-0.79132	-0.04027
H	-3.52540	2.95091	0.10323
H	-5.68685	1.73558	0.12512
H	-1.19429	3.01645	0.04476
H	1.28664	-2.27065	0.34082

H	1.75375	-2.22857	-1.35223
H	2.82238	-3.03323	-0.19083
H	4.98997	-2.32868	2.04546
H	6.41371	-2.91397	1.12703
H	6.27059	-1.19952	1.61497
H	4.37160	1.88315	0.77557
H	3.70772	3.22435	-0.20205
H	4.34115	1.74931	-0.99351
O	1.37358	3.67537	-0.01043
O	5.30720	-2.30429	-1.08776

---

Coordinates for Model Number 13

C	-3.67279	-1.22060	0.01317
C	-2.50219	-0.47529	-0.00455
C	-2.53290	0.93128	0.01928
C	-1.35301	1.75789	0.00546
C	-3.79582	1.56094	0.06101
C	-4.96664	0.83012	0.07830
C	-4.91413	-0.57103	0.05464
C	1.04394	2.42257	-0.03188
C	1.81501	0.32141	-0.08345
C	-0.03787	1.41056	-0.03057
C	3.52171	2.24942	-0.10060
C	-6.12197	-2.61626	0.05318
C	2.73859	-0.82693	-0.11531
C	4.96263	-1.59795	-0.02944
C	2.11181	-2.17070	-0.34485
C	5.69244	-1.85091	1.26269
N	2.20640	1.64738	-0.06206
N	0.52074	0.14265	-0.06735
N	3.98598	-0.60446	0.05751
H	-3.61223	-2.30282	-0.00625
H	-1.54048	-0.97530	-0.03779
H	-3.84405	2.64684	0.08041
H	-5.93757	1.31375	0.10987
H	-1.52054	2.83463	0.02787
H	-7.17365	-2.90323	0.07315
H	-5.65721	-3.00730	-0.86001
H	-5.61587	-3.03432	0.93184
H	1.22084	-2.28250	0.27634
H	1.78660	-2.24593	-1.38780
H	2.82119	-2.97614	-0.15201
H	4.98486	-2.09264	2.06289
H	6.40313	-2.66658	1.12725
H	6.22350	-0.94360	1.56762
H	4.09119	2.00694	0.79832
H	3.35601	3.32781	-0.15709
H	4.08331	1.91129	-0.97343
O	1.00068	3.64076	-0.01484
O	-6.10632	-1.20312	0.07440
O	5.25470	-2.12703	-1.08441

---

---

Coordinates for Model Number 14			
C	-3.18833	-0.58267	-0.01951
C	-1.93308	0.01112	-0.01825
C	-1.78810	1.40886	0.03659
C	-0.51267	2.08070	0.03291
C	-2.96228	2.19060	0.09099
C	-4.21565	1.61159	0.09206
C	-4.33757	0.21682	0.03589
C	1.95070	2.43085	-0.01602
C	2.44302	0.24749	-0.05626
C	0.74633	1.56707	0.00049
C	4.38414	1.93683	-0.12598
C	-5.78887	-1.66339	-0.01573
C	-7.21374	-1.95476	0.00026
C	-8.38812	-2.23388	0.01139
C	3.20700	-1.01322	-0.07882
C	5.31573	-2.06529	-0.04673
C	2.39827	-2.26821	-0.22924
C	6.09246	-2.34617	1.21173
N	3.00269	1.51133	-0.05329
N	1.13691	0.23747	-0.02747
N	4.47879	-0.95140	0.04215
H	-3.25732	-1.66365	-0.06570
H	-1.04230	-0.60592	-0.06275
H	-2.87653	3.27347	0.13279
H	-5.11875	2.21167	0.13420
H	-0.54022	3.16992	0.05597
H	-5.33711	-2.06942	-0.93201
H	-5.30171	-2.14184	0.84576
H	-9.42890	-2.47072	0.02138
H	1.93903	-2.28868	-1.22255
H	3.02221	-3.15515	-0.11932
H	1.58185	-2.27698	0.49664
H	5.41818	-2.44933	2.06833
H	6.68087	-3.25491	1.08142
H	6.75791	-1.50352	1.42604
H	4.93574	1.64348	0.76920
H	4.36005	3.02612	-0.20604
H	4.87881	1.50757	-0.99875
O	2.06745	3.64395	-0.00760
O	-5.60522	-0.25645	0.03860
O	5.47029	-2.67777	-1.08560

---

Coordinates for Model Number 15			
C	-2.55953	-0.27567	-0.04361
C	-1.26796	0.22908	-0.04840
C	-1.04850	1.61914	0.00009
C	0.26909	2.22048	-0.00713
C	-2.16711	2.47439	0.05357
C	-3.45428	1.96574	0.06227
C	-3.66714	0.58219	0.01378

C	2.74954	2.42193	-0.04647
C	3.10806	0.21557	-0.07244
C	1.49003	1.62916	-0.03303
C	5.15036	1.78277	-0.13889
C	-5.07678	0.09053	0.03231
C	-5.31425	-1.40943	-0.05522
C	-6.78567	-1.78910	-0.14385
C	-7.63150	-1.58435	1.08134
C	3.79916	-1.09045	-0.08273
C	5.84145	-2.25994	0.02695
C	2.92555	-2.29353	-0.27944
C	6.51130	-2.61799	1.32544
N	3.74434	1.44153	-0.07030
N	1.80318	0.27949	-0.05382
N	5.06489	-1.09837	0.08921
H	-2.70008	-1.35138	-0.08640
H	-0.41375	-0.43681	-0.09261
H	-2.01023	3.54944	0.08830
H	-4.31887	2.62063	0.10458
H	0.30242	3.30917	0.00919
H	-4.85678	-1.88165	0.82529
H	-4.81195	-1.80800	-0.94310
H	-7.65849	-0.51617	1.32125
H	-8.64187	-1.95304	0.89992
H	-7.19343	-2.10414	1.94141
H	2.10067	-2.27821	0.43698
H	2.47789	-2.26188	-1.27765
H	3.49819	-3.21639	-0.18656
H	5.77943	-2.67960	2.13743
H	7.03531	-3.56776	1.21385
H	7.22622	-1.83364	1.59515
H	5.67788	1.46145	0.76068
H	5.19166	2.87079	-0.22635
H	5.62117	1.31961	-1.00760
O	2.93109	3.62427	-0.04498
O	-6.00617	0.87421	0.11366
O	-7.22671	-2.25705	-1.17265
O	6.02884	-2.85044	-1.01790

---

Coordinates for Model Number 16a

C	-3.49974	-1.82217	-0.02308
C	-2.40535	-0.96886	-0.03463
C	-2.59700	0.42512	0.00615
C	-1.51232	1.37754	-0.00468
C	-3.91209	0.92925	0.05731
C	-5.00371	0.07910	0.07060
C	-4.79212	-1.30735	0.03025
C	0.79672	2.29989	-0.04255
C	1.79390	0.29608	-0.05976
C	-0.16833	1.17405	-0.02828
C	3.27917	2.39920	-0.11248

C	2.83880	-0.74400	-0.07123
C	5.14028	-1.25193	-0.03975
C	2.35949	-2.15807	-0.21646
C	5.95488	-1.35125	1.22211
N	2.03743	1.65689	-0.06210
N	0.52765	-0.02472	-0.04182
N	4.05750	-0.37579	0.04985
H	-3.35183	-2.89973	-0.05581
H	-1.39802	-1.36619	-0.07616
H	-4.09297	1.99992	0.08668
H	-1.79474	2.42980	0.00809
H	-5.70933	-3.00181	0.01126
H	-6.88692	-0.15039	0.11894
H	1.54460	-2.35268	0.48440
H	1.95539	-2.30461	-1.22313
H	3.17431	-2.86703	-0.06719
H	5.33340	-1.70487	2.05188
H	6.78970	-2.03416	1.06184
H	6.33003	-0.36155	1.50058
H	3.87834	2.22300	0.78236
H	2.99697	3.45304	-0.17124
H	3.86642	2.11980	-0.98905
O	0.62202	3.50574	-0.04199
O	-5.92483	-2.06436	0.04637
O	-6.25930	0.58399	0.12062
O	5.44271	-1.79930	-1.08249

---

Coordinates for Model Number 16b

C	-3.72932	-1.36345	0.00468
C	-2.59078	-0.57926	-0.01054
C	-2.69235	0.82621	0.01866
C	-1.54347	1.69993	0.00885
C	-3.96891	1.41500	0.06247
C	-5.11396	0.62867	0.07721
C	-4.99698	-0.75581	0.04866
C	0.83005	2.44770	-0.02229
C	1.66986	0.37381	-0.07899
C	-0.21854	1.39884	-0.02596
C	3.31286	2.35657	-0.08533
C	2.62739	-0.74738	-0.10972
C	4.88093	-1.44330	-0.04893
C	2.03255	-2.11416	-0.28082
C	5.62830	-1.65039	1.24164
N	2.01734	1.71227	-0.05361
N	0.38347	0.15103	-0.06577
N	3.87047	-0.48386	0.03047
H	-1.61636	-1.05284	-0.04456
H	-4.05983	2.49693	0.08556
H	-1.75076	2.76921	0.03492
H	-6.88468	-1.14354	0.08893
H	-6.09886	1.08908	0.11063

H	-4.50908	-3.09458	-0.00883
H	1.30903	-2.30916	0.51522
H	1.47880	-2.16126	-1.22259
H	2.80761	-2.88051	-0.28907
H	4.93540	-1.87946	2.05785
H	6.34977	-2.45851	1.11873
H	6.15123	-0.72761	1.51281
H	3.88323	2.04772	-0.96328
H	3.89084	2.11943	0.80960
H	3.11396	3.43006	-0.12690
O	0.74587	3.66338	-0.00049
O	-6.04932	-1.62089	0.06030
O	-3.62166	-2.71306	-0.02206
O	5.18667	-1.97532	-1.09832

---

Coordinates for Model Number 17

C	-3.12807	-0.45900	-0.14519
C	-1.86062	0.10567	-0.10763
C	-1.68460	1.49743	-0.00103
C	-0.39532	2.13913	0.03963
C	-2.84199	2.30332	0.06478
C	-4.10747	1.75287	0.02991
C	-4.26165	0.36307	-0.07492
C	2.07584	2.43502	0.05204
C	2.52121	0.24328	-0.04924
C	0.85336	1.59890	0.01376
C	4.49998	1.88986	-0.01301
C	-5.76616	-1.48570	-0.19244
C	-7.24168	-1.70930	-0.20348
C	-7.86395	-2.56701	0.60307
C	3.25610	-1.03350	-0.09602
C	5.33715	-2.13823	-0.06118
C	2.42017	-2.26433	-0.29280
C	6.08676	-2.47301	1.20074
N	3.10816	1.49366	0.00868
N	1.21486	0.26257	-0.04993
N	4.52708	-1.00660	0.04586
H	-3.22394	-1.53551	-0.22897
H	-0.98289	-0.52906	-0.16181
H	-2.73200	3.38180	0.14620
H	-4.99707	2.37206	0.08358
H	-0.39955	3.22733	0.09942
H	-7.80164	-1.13971	-0.94359
H	-5.31616	-1.86258	-1.12364
H	-5.29952	-2.00692	0.65394
H	-8.93505	-2.73156	0.53777
H	-7.32223	-3.13665	1.35490
H	1.97447	-2.24664	-1.29220
H	3.02134	-3.16878	-0.19927
H	1.59394	-2.27329	0.42177
H	5.39812	-2.57439	2.04599

H	6.64850	-3.39640	1.05648
H	6.77543	-1.65709	1.44339
H	5.02375	1.53887	0.87802
H	4.50119	2.98208	-0.03979
H	5.00541	1.49279	-0.89507
O	2.21848	3.64463	0.10312
O	-5.53484	-0.08436	-0.09674
O	5.49369	-2.72639	-1.11386

---

Coordinates for Model Number 18

C	-3.86979	-1.71358	-0.02142
C	-2.76840	-0.88132	-0.03410
C	-2.91548	0.52245	0.01938
C	-1.81595	1.44595	0.00718
C	-4.22709	1.03618	0.08675
C	-5.33399	0.20943	0.10250
C	-5.17528	-1.18683	0.04696
C	0.51281	2.32042	-0.03570
C	1.46924	0.29570	-0.07205
C	-0.47279	1.21717	-0.02850
C	2.99673	2.36743	-0.10426
C	2.48776	-0.76710	-0.08599
C	4.77943	-1.32990	-0.04876
C	1.97141	-2.17071	-0.21060
C	5.60451	-1.40866	1.20886
N	1.73992	1.65178	-0.06270
N	0.19528	0.00414	-0.05555
N	-6.26385	-2.02113	0.10228
N	3.71711	-0.43089	0.02929
H	-3.73220	-2.79165	-0.06067
H	-1.76904	-1.29995	-0.08695
H	-4.37060	2.11332	0.12664
H	-6.33237	0.63603	0.15924
H	-6.14499	-2.97966	-0.18250
H	-2.07476	2.50445	0.03246
H	-7.17277	-1.63148	-0.08802
H	1.24773	-2.37325	0.58314
H	1.43702	-2.28293	-1.15850
H	2.78631	-2.89377	-0.18108
H	4.96809	-1.56017	2.08637
H	6.33014	-2.21756	1.11828
H	6.13262	-0.46025	1.35344
H	3.59213	2.16741	0.78831
H	2.73749	3.42790	-0.15011
H	3.57876	2.08632	-0.98395
O	0.36536	3.53174	-0.02458
O	5.06987	-1.90732	-1.07933

---

Coordinates for Model Number 19

C	-0.98632	1.47295	-0.05706
C	1.35088	2.31935	-0.07752
C	2.28293	0.28543	-0.10045

C	0.34887	1.22337	-0.08502
C	3.83584	2.33980	-0.11329
C	-2.10393	0.55776	-0.05465
C	-3.38177	1.09234	0.05679
C	-1.95863	-0.85942	-0.16084
C	-4.53242	0.27461	0.07540
C	-3.06074	-1.67132	-0.15414
C	-4.37391	-1.14005	-0.03490
C	-5.83832	0.81550	0.19996
C	-5.52644	-1.96063	-0.02017
C	-6.93887	-0.00771	0.21412
C	-6.78114	-1.40781	0.10226
C	3.29531	-0.78766	-0.09164
C	5.57488	-1.36959	0.04158
C	2.77798	-2.18259	-0.28146
C	6.32863	-1.49601	1.33852
N	2.56976	1.63799	-0.08667
N	1.00774	0.00364	-0.10509
N	4.51767	-0.45869	0.08848
H	-1.23195	2.53398	-0.02488
H	-3.50583	2.17036	0.13632
H	-0.96068	-1.27369	-0.24704
H	-2.94183	-2.74865	-0.23996
H	-5.95368	1.89310	0.28486
H	-5.40576	-3.03757	-0.10608
H	-7.93453	0.41518	0.31044
H	-7.65790	-2.04901	0.11347
H	1.92137	-2.35795	0.37320
H	2.42170	-2.30206	-1.30979
H	3.55959	-2.92113	-0.10195
H	5.64444	-1.69980	2.16861
H	7.07051	-2.29030	1.25288
H	6.83016	-0.54813	1.56066
H	4.43249	2.03628	-0.97529
H	4.40884	2.14849	0.79538
H	3.59030	3.40206	-0.18313
O	1.21254	3.52949	-0.06864
O	5.90829	-1.91993	-0.98972

---

Coordinates for Model Number 20

C	-1.27512	0.89162	0.09447
C	0.90741	2.09412	0.07778
C	2.14054	0.22529	-0.01028
C	0.09091	0.86673	0.06930
C	3.35748	2.49222	-0.03930
C	-2.17968	-0.20893	0.06684
C	-3.63235	-0.19194	0.01995
C	-1.81482	-1.55103	0.12855
C	-4.06146	-1.53510	0.07714
C	-4.61003	0.82267	-0.08235
C	-5.39939	-1.92961	0.08323

C	-5.95253	0.43412	-0.06647
C	-6.33723	-0.90954	0.02097
C	3.29666	-0.68162	-0.07428
C	5.63985	-0.92531	-0.12224
C	2.98173	-2.13938	-0.24892
C	6.53574	-0.90496	1.08756
N	2.21699	1.60489	0.02857
N	0.91710	-0.24046	0.01250
N	-4.25874	2.16452	-0.14365
N	-2.92643	-2.32205	0.13669
N	4.47075	-0.18021	0.02772
H	-1.68631	1.89513	0.16621
H	-0.81901	-1.96698	0.17739
H	-2.92875	-3.32751	0.19061
H	-5.69059	-2.97387	0.13384
H	-6.71930	1.20203	-0.13574
H	-7.39565	-1.15265	0.02425
H	-3.42931	2.38196	-0.67584
H	-5.01982	2.78507	-0.37813
H	2.20863	-2.44331	0.46004
H	2.57986	-2.30677	-1.25303
H	3.87497	-2.75301	-0.13137
H	5.99331	-1.24897	1.97446
H	7.40619	-1.53669	0.90828
H	6.85824	0.12189	1.28740
H	4.00581	2.37105	0.83044
H	2.94849	3.50522	-0.06134
H	3.94388	2.30541	-0.94091
O	0.58946	3.27328	0.10442
O	5.94432	-1.45338	-1.17531

---

Coordinates for Model Number 21

C	-0.66422	1.71469	-0.04093
C	1.72316	2.41036	-0.05921
C	2.51837	0.32054	-0.10654
C	0.65065	1.38093	-0.07262
C	4.20349	2.26922	-0.10415
C	-1.83629	0.86977	-0.03792
C	-3.08801	1.45664	0.05500
C	-1.79006	-0.55871	-0.12443
C	-4.25429	0.66576	0.06849
C	-4.07366	-0.74567	-0.02478
C	-5.55425	1.22114	0.17049
C	-5.23125	-1.57950	-0.01627
C	-6.64554	0.38734	0.17792
C	-6.48680	-1.01049	0.08388
C	3.45593	-0.81998	-0.11698
C	5.69362	-1.55500	0.01617
C	2.84127	-2.17414	-0.31176
C	6.37181	-1.79865	1.33763
N	2.89440	1.65124	-0.07849

N	1.22824	0.12156	-0.10822
N	-2.85004	-1.32652	-0.12012
N	4.69724	-0.57676	0.06256
H	-0.84700	2.78854	-0.00663
H	-3.17904	2.53898	0.12256
H	-0.82313	-1.04779	-0.19915
H	-5.67014	2.29875	0.24213
H	-7.64763	0.79857	0.25613
H	-7.36614	-1.65143	0.09062
H	-5.87995	-3.36290	-0.09151
H	2.10375	-2.36625	0.47238
H	3.60413	-2.95290	-0.31726
H	2.29950	-2.20240	-1.26135
H	5.63551	-2.04808	2.10874
H	7.09624	-2.60643	1.23091
H	6.88135	-0.88584	1.66250
H	4.76529	2.03612	0.80192
H	4.02675	3.34548	-0.16605
H	4.77765	1.93394	-0.96961
O	1.65936	3.62593	-0.03970
O	-5.02980	-2.91041	-0.10755
O	6.03721	-2.07425	-1.02712

---

Coordinates for Model Number 21-H (with H-capping)

C	-0.60369	-1.69151	0.01384
C	1.79348	-2.38307	0.03666
C	2.56321	-0.27831	-0.00029
C	0.71147	-1.34981	0.01244
C	-1.78157	-0.85884	-0.00551
C	-3.02821	-1.46591	0.00137
C	-1.75077	0.57325	-0.03161
C	-4.20468	-0.69174	-0.01673
C	-4.03967	0.72458	-0.04293
C	-5.49964	-1.26792	-0.00939
C	-5.20786	1.54305	-0.06176
C	-6.60171	-0.44857	-0.02778
C	-6.45843	0.95363	-0.05406
C	3.52994	0.80331	-0.01556
C	5.68117	1.61210	0.03120
N	2.94146	-1.59881	0.02681
N	1.27412	-0.08162	-0.01133
N	-2.82150	1.32476	-0.04984
N	4.78210	0.53851	-0.03313
H	-0.77809	-2.76742	0.03291
H	3.89524	-1.92827	0.03497
H	-3.10580	-2.55121	0.02136
H	-0.78966	1.07911	-0.03645
H	-5.60323	-2.34891	0.01092
H	-7.60037	-0.87518	-0.02230
H	-7.34622	1.58261	-0.06835
H	-5.87623	3.32097	-0.09445

H	3.12186	1.81874	-0.02276
H	6.54063	1.48404	-0.64787
O	1.73254	-3.59551	0.05860
O	-5.02091	2.87803	-0.08637
O	5.57818	2.55111	0.78522

---

Coordinates for Model Number 22			
C	-3.38903	-1.19183	-0.01122
C	-2.21672	-0.45346	-0.01895
C	-2.25823	0.95522	0.01461
C	-1.07085	1.78611	0.01189
C	-3.51223	1.59522	0.05453
C	-4.69087	0.86584	0.06199
C	-4.60940	-0.52239	0.02949
C	1.33442	2.42612	-0.01473
C	2.07859	0.32003	-0.08232
C	0.23475	1.42128	-0.02282
C	3.81128	2.22615	-0.08732
C	2.99042	-0.84282	-0.12075
C	5.21161	-1.62958	-0.03426
C	2.34368	-2.17929	-0.33235
C	5.90048	-1.91842	1.27135
N	2.48741	1.63873	-0.05051
N	0.78304	0.15006	-0.06954
N	-5.84748	-1.30356	0.03872
N	4.23793	-0.62775	0.04558
H	-3.37686	-2.27545	-0.03628
H	-1.25412	-0.95041	-0.05095
H	-3.55442	2.68056	0.08026
H	-5.66204	1.34584	0.09213
H	-1.23301	2.86276	0.04249
H	1.59797	-2.36146	0.44616
H	1.80947	-2.18545	-1.28686
H	3.08686	-2.97644	-0.34150
H	5.16895	-2.17567	2.04468
H	6.60915	-2.73574	1.13508
H	6.42959	-1.02333	1.61390
H	4.37680	1.97439	0.81132
H	3.65943	3.30655	-0.14096
H	4.36709	1.88256	-0.96130
O	1.29806	3.64069	0.01200
O	-6.90331	-0.68947	0.07397
O	-5.74831	-2.52113	0.01099
O	5.52469	-2.13209	-1.09443

---

**Table C3:** One-photon excitation energies [in eV] and the corresponding wavelengths [in nm], OPA oscillator strengths, TPA transition moments ( $\delta^{TPA}$ ) and TPA cross-sections [in GM] for the transition to  $S_1$  as determined at the B3LYP/6-31G+(d,p) Level of Theory and PCM with parameters for  $H_2O$ . The properties for RFP-derived chromophores (Figure C1) are computed in Chapter 4, while the GFP-derived ones were previously reported (see Chapter 3).<sup>75</sup>

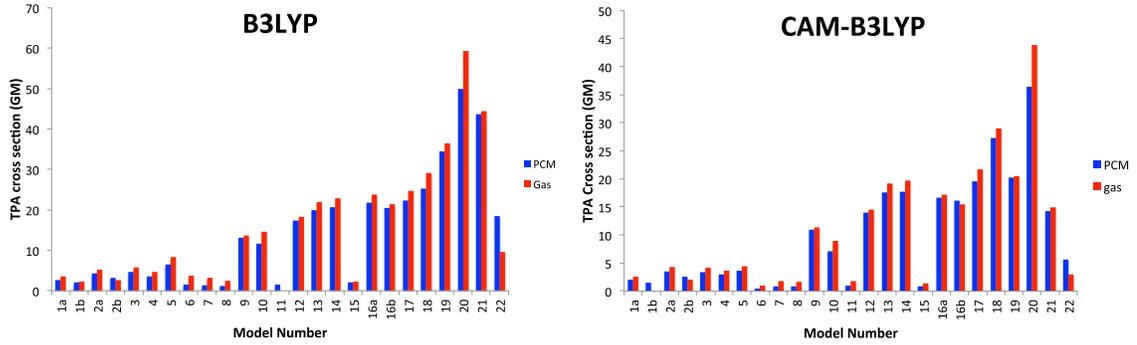
Model	RFP-derived Chromophores					GFP-derived Chromophores				
	Energy [eV]	Wavelength (nm)	Oscillator Strength	$\delta^{TPA}$	TPA [GM]	Energy [eV]	Wavelength (nm)	Oscillator Strength	$\delta^{TPA}$	TPA [GM]
1a	2.977	417	0.64	820	2.66	3.455	359	0.640	42	0.19
1b	2.987	415	0.61	618	2.02	3.468	358	0.679	15	0.07
2a	2.968	418	0.67	1335	4.31	3.473	357	0.652	13	0.06
2b	2.995	414	0.63	983	3.23	3.452	359	0.557	20	0.09
3	2.968	418	0.68	1425	4.60	3.463	358	0.705	24	0.11
4	2.923	424	0.75	1162	3.64	3.387	366	0.787	11	0.05
5	2.945	421	0.70	2062	6.56	3.467	358	0.704	180	0.79
6	2.881	430	0.83	527	1.60	3.197	388	0.912	342	1.28
7	2.819	440	0.88	450	1.31	3.200	388	0.927	339	1.27
8	2.883	430	0.75	395	1.20	3.304	375	0.796	370	1.48
9	2.897	428	0.82	4254	13.09	3.403	364	0.841	609	2.59
10	2.898	428	0.81	3776	11.62	3.369	368	0.848	763	3.18
11	2.859	434	0.78	498	1.49	3.256	381	0.812	1049	4.08
12	2.772	447	0.92	6135	17.28	3.210	386	0.984	1358	5.13
13	2.785	445	0.83	7024	19.98	3.289	377	0.853	1751	6.94
14	2.801	443	0.85	7205	20.72	3.310	375	0.894	1744	7.00
15	2.840	437	0.80	724	2.14	3.222	385	0.837	1879	7.15
16a	2.733	454	0.64	7938	21.75	3.218	385	0.577	2046	7.77
16b	2.783	446	0.77	7197	20.44	3.260	380	0.711	1882	7.33
17	2.775	447	0.86	7918	22.36	3.284	378	0.901	2067	8.17
18	2.630	471	0.91	9944	25.23	3.137	395	0.927	2540	9.16
19	2.732	454	0.80	12564	34.39	3.138	395	0.553	2908	10.50
20	2.239	554	0.48	27128	49.88	2.689	461	0.363	5738	15.21
21	2.654	467	0.46	16899	43.64	2.985	415	0.298	5258	17.19
22	2.710	458	0.69	6876	18.52	2.965	418	0.638	8834	28.47

**Table C4:** Differences between the properties computed for the RFP-derived chromophore and those previously computed for the corresponding GFP-derived chromophore for the transition to  $S_1$  as in Table C3.

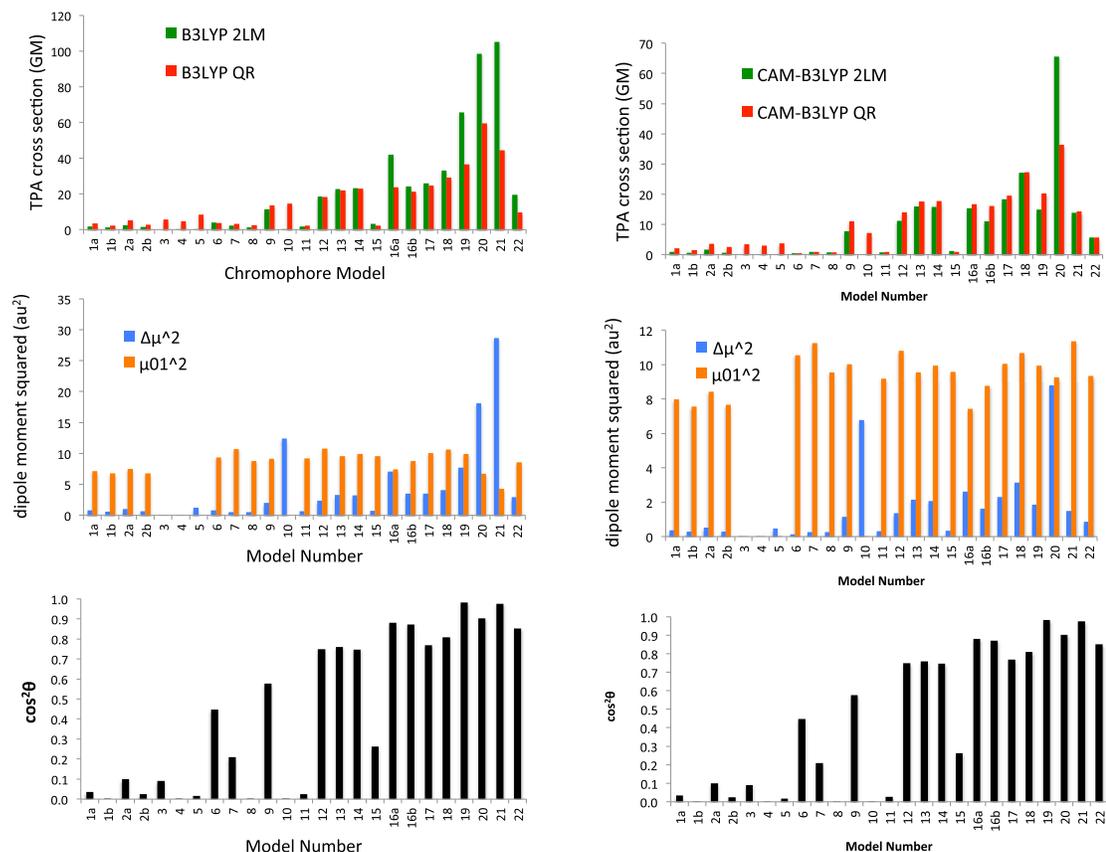
Model	$\Delta$ Energy [eV]	$\Delta$ Wavelength (nm)	$\Delta$ Oscillator Strength	$\Delta\delta^{TPA}$	$\Delta$ TPA [GM]	%increase in OPA OS	%increase in TPA
1a	-0.478	58	0.001	778	2.48	0.2	1339.8
1b	-0.481	58	-0.070	603	1.96	-10.3	2984.8
2a	-0.505	61	0.020	1322	4.25	3.1	7356.3
2b	-0.457	55	0.077	963	3.14	13.8	3630.8
3	-0.495	60	-0.028	1401	4.50	-4.0	4243.4
4	-0.464	58	-0.034	1151	3.60	-4.3	7951.8
5	-0.522	63	-0.003	1882	5.77	-0.4	728.5
6	-0.316	43	-0.084	184	0.32	-9.2	24.9
7	-0.381	52	-0.052	111	0.04	-5.6	3.0
8	-0.421	55	-0.049	25	-0.28	-6.2	-18.7
9	-0.506	64	-0.026	3645	10.51	-3.1	406.5
10	-0.471	60	-0.037	3013	8.45	-4.4	266.0
11	-0.397	53	-0.031	-551	-2.59	-3.8	-63.4
12	-0.438	61	-0.066	4777	12.15	-6.7	236.8
13	-0.504	68	-0.019	5274	13.03	-2.2	187.7
14	-0.509	68	-0.042	5461	13.72	-4.7	195.9
15	-0.382	52	-0.037	-1155	-5.01	-4.4	-70.1
16a	-0.485	68	0.063	5892	13.98	10.9	179.9
16b	-0.477	65	0.058	5315	13.10	8.2	178.7
17	-0.509	69	-0.043	5851	14.18	-4.8	173.5
18	-0.507	76	-0.016	7404	16.06	-1.7	175.3
19	-0.406	59	0.242	9655	23.89	43.8	227.5
20	-0.450	93	0.118	21391	34.67	32.5	227.9
21	-0.331	52	0.165	11641	26.46	55.4	154.0
22	-0.255	39	0.051	-1958	-9.95	8.0	-35.0

**Table C5:** One-photon excitation energies [in eV] and the corresponding wavelengths [in nm], OPA oscillator strengths, TPA transition moments ( $\delta^{TPA}$ ) and TPA cross-sections [in GM] for the transition to  $S_1$  as determined at the B3LYP and CAM-B3LYP/6-31G+(d,p) Level of Theory in the gas phase for the RFP-derived chromophores given in Figure C1.

Model	B3LYP					CAM-B3LYP				
	Energy [eV]	Wavelength (nm)	Oscillator Strength	$\delta^{TPA}$	TPA [GM]	Energy [eV]	Wavelength (nm)	Oscillator Strength	$\delta^{TPA}$	TPA [GM]
1a	3.032	409	0.53	1042	3.51	3.384	366	0.66	616	2.59
2a	3.050	407	0.51	661	2.25	3.371	368	0.70	1025	4.27
1b	3.016	411	0.55	1580	5.27	-	-	-	-	-
2b	3.059	405	0.51	779	2.67	3.400	365	0.64	474	2.01
3	3.023	410	0.56	1708	5.72	3.371	368	0.70	1017	4.23
4	2.982	416	0.63	1445	4.71	3.323	373	0.77	917	3.71
5	2.991	415	0.56	2551	8.37	3.374	368	0.76	1049	4.38
6	2.949	420	0.68	1175	3.75	3.339	371	0.86	254	1.04
7	2.890	429	0.76	1033	3.16	3.256	381	0.90	474	1.84
8	2.959	419	0.64	756	2.43	3.304	375	0.77	417	1.67
9	2.962	419	0.66	4240	13.64	3.305	375	0.81	2846	11.40
10	2.956	419	0.67	4546	14.57	3.315	374	0.80	2213	8.92
11	2.931	423	0.66	737	2.32	3.283	378	0.80	469	1.85
12	2.857	434	0.76	6134	18.36	3.206	387	0.90	3850	14.51
13	2.872	432	0.67	7243	21.91	3.219	385	0.81	5051	19.19
14	2.874	431	0.70	7581	22.96	3.225	385	0.85	5156	19.66
15	2.915	425	0.68	719	2.24	3.274	379	0.83	351	1.38
16a	2.832	438	0.52	8071	23.73	3.242	382	0.73	4465	17.21
16b	2.901	428	0.62	6914	21.33	3.277	378	0.78	3915	15.42
17	2.854	435	0.70	8271	24.70	3.204	387	0.85	5764	21.69
18	2.777	446	0.73	10315	29.18	3.119	398	0.89	8115	28.95
19	2.841	436	0.69	12303	36.42	3.251	381	0.99	5278	20.45
20	2.366	524	0.39	28912	59.35	2.869	432	0.65	14550	43.93
21	2.718	456	0.29	16415	44.46	3.296	376	0.92	3756	14.96
22	2.881	430	0.60	3171	9.65	3.276	378	0.75	754	2.97



**Figure C2:** A visual representation for the TPA cross sections computed in the gas phase for the RFP-derived chromophores as given in Table C5 with the corresponding PCM values as given in Tables C3 and C7.



**Figure C3:** Plots for the B3LYP (left) and CAM-B3LYP (right) data given in Table C6. Top figures compare the TPA cross sections computed via quadratic response (QR) to those obtained from the 2LM approximation. Middle figures compare the magnitudes of the transition dipole moment ( $\mu_{01}$ ) and the difference between permanent moments of excited and ground states ( $\Delta\mu$ ). The alignment between the  $\mu_{01}$  and  $\Delta\mu$  vectors is plotted in the bottom figures in terms of  $\cos^2\theta$  (the vectors are perfectly aligned when  $\cos^2\theta = 1$ ).

**Table C6:** *TPA cross sections via the truncated SOS approach using 2 states; i.e. the 2-level model (2LM), and via quadratic response (QR), and the contributing components to the 2LM expression ((Equation 5 in the main text)) determined at the B3LYP and CAM-B3LYP/6-31+G(d,p) levels of theory in the gas phase.*

Model	B3LYP					CAM-B3LYP				
	2LM (GM)	QR (GM)	$\Delta\mu$ (au)	$\mu_{01}$ (au)	$\cos^2\theta$	2LM (GM)	QR (GM)	$\Delta\mu$ (au)	$\mu_{01}$ (au)	$\cos^2\theta$
1a	1.70	3.51	0.88	2.67	0.03	0.85	2.03	0.60	2.83	0.01
1b	1.16	2.25	0.77	2.60	0.00	0.60	1.50	0.52	2.75	0.00
2a	2.54	5.27	0.99	2.73	0.10	1.61	3.52	0.72	2.90	0.14
2b	1.38	2.67	0.82	2.61	0.03	0.66	2.55	0.53	2.77	0.02
3	0.00	5.72	0.05	0.02	0.09	0.00	3.42	0.05	0.02	0.00
4	0.00	4.71	0.17	0.06	0.00	0.00	2.94	0.19	0.06	0.00
5	0.00	8.37	1.09	0.03	0.02	0.00	3.69	0.69	0.07	0.00
6	3.95	3.75	0.88	3.06	0.45	0.44	0.49	0.38	3.25	0.00
7	2.18	3.16	0.70	3.27	0.21	0.85	0.89	0.51	3.35	0.00
8	1.36	2.43	0.73	2.96	0.00	0.73	0.80	0.51	3.09	0.02
9	11.26	13.64	1.41	3.02	0.58	7.80	10.97	1.07	3.16	0.67
10	0.06	14.57	3.53	0.12	0.00	0.03	7.13	2.60	0.12	0.00
11	1.83	2.32	0.81	3.03	0.03	0.81	0.94	0.55	3.03	0.00
12	18.50	18.36	1.54	3.29	0.75	11.18	13.94	1.16	3.29	0.82
13	22.63	21.91	1.81	3.09	0.76	15.98	17.62	1.46	3.09	0.85
14	23.29	22.96	1.80	3.15	0.75	15.84	17.69	1.43	3.15	0.84
15	3.11	2.24	0.86	3.09	0.26	1.22	0.92	0.58	3.09	0.16
16a	41.99	23.73	2.66	2.72	0.88	15.34	16.62	1.61	2.72	0.87
16b	24.26	21.33	1.87	2.96	0.87	11.08	16.14	1.28	2.96	0.84
17	25.86	24.70	1.87	3.17	0.77	18.29	19.58	1.52	3.17	0.86
18	33.00	29.18	2.02	3.27	0.81	27.11	27.28	1.77	3.27	0.90
19	65.60	36.42	2.78	3.15	0.98	14.89	20.30	1.36	3.15	0.90
20	98.51	59.35	4.26	2.59	0.90	65.57	36.39	2.96	3.04	0.89
21	105.19	44.46	5.35	2.08	0.97	13.86	14.28	1.22	3.37	0.91
22	19.38	9.65	1.70	2.92	0.85	5.61	5.65	0.93	3.06	0.71

**Table C7:** One-photon excitation energies [in eV] and the corresponding wavelengths [in nm], OPA oscillator strengths, TPA transition moments ( $\delta^{TPA}$ ) and TPA cross-sections [in GM] for the transition to  $S_1$  as determined at the CAM-B3LYP/6-31G+(d,p) Level of Theory and PCM with parameters for  $H_2O$ . Models for the RFP-derived chromophores are given in Figure C1. Equivalent GFP-derived chromophores are given in Figure 2 of Chapter 3.<sup>75</sup>

Model	RFP-derived Chromophores					GFP-derived Chromophores				
	Energy [eV]	Wavelength (nm)	Oscillator Strength	$\delta^{TPA}$	TPA [GM]	Energy [eV]	Wavelength (nm)	Oscillator Strength	$\delta^{TPA}$	TPA [GM]
1a	3.306	375	0.78	507	2.03	3.707	334	0.730	1	0.01
1b	3.312	374	0.74	372	1.50	3.714	334	0.720	11	0.06
2a	3.295	376	0.81	884	3.52	3.702	335	0.743	76	0.38
2b	3.315	374	0.76	634	2.55	3.705	335	0.752	33	0.17
3	3.292	377	0.81	862	3.42	3.695	336	0.769	76	0.38
4	3.242	383	0.89	763	2.94	3.625	342	0.853	19	0.09
5	3.292	377	0.87	929	3.69	3.688	336	0.758	263	1.31
6	3.246	382	0.98	126	0.49	3.506	354	0.994	146	0.66
7	3.168	391	1.01	242	0.89	3.508	353	1.009	150	0.68
8	3.212	386	0.89	211	0.80	3.573	347	0.883	85	0.40
9	3.215	386	0.95	2894	10.97	3.624	342	0.889	688	3.31
10	3.235	383	0.93	1860	7.13	3.625	342	0.889	454	2.19
11	3.194	388	0.93	252	0.94	3.544	350	0.906	166	0.76
12	3.103	400	1.05	3949	13.94	3.466	358	1.029	1252	5.51
13	3.112	398	0.97	4963	17.62	3.522	352	0.900	1623	7.38
14	3.131	396	0.99	4922	17.69	3.542	350	0.938	1609	7.40
15	3.183	390	0.95	249	0.92	3.525	352	0.941	379	1.73
16a	3.127	397	0.87	4636	16.62	3.537	351	0.789	1483	6.80
16b	3.147	394	0.93	4443	16.14	3.548	349	0.852	1424	6.58
17	3.106	399	1.00	5536	19.58	3.517	353	0.948	1882	8.54
18	2.950	420	1.07	8551	27.28	3.364	369	0.988	2953	12.25
19	3.142	395	1.16	5606	20.30	3.511	353	0.999	1458	6.59
20	2.730	454	0.77	13314	36.39	3.139	395	0.630	3413	12.33
21	3.198	388	1.06	3808	14.28	3.535	351	0.871	1301	5.96
22	3.140	395	0.88	1564	5.65	3.432	361	0.862	2579	11.14

**Table C8:** Differences between the properties computed for the RFP-derived chromophore and those previously computed for the corresponding GFP-derived chromophore for the transition to  $S_1$  as in Table C7.

Model	$\Delta$ Energy [eV]	$\Delta$ Wavelength (nm)	$\Delta$ Oscillator Strength	$\Delta\delta^{TPA}$	$\Delta$ TPA [GM]	%increase in OPA OS	%increase in TPA
1a	-0.402	41	0.048	506	2.03	6.6	28419.8
1b	-0.402	41	0.019	361	1.44	2.6	2603.6
2a	-0.407	41	0.071	808	3.14	9.6	818.1
2b	-0.390	39	0.012	601	2.39	1.6	1439.5
3	-0.403	41	0.041	786	3.04	5.3	805.1
4	-0.383	40	0.036	745	2.85	4.2	3120.5
5	-0.397	41	0.116	666	2.38	15.3	181.6
6	-0.260	28	-0.013	-20	-0.17	-1.3	-26.0
7	-0.341	38	-0.002	92	0.22	-0.2	31.9
8	-0.361	39	0.008	126	0.40	0.9	101.1
9	-0.408	43	0.060	2206	7.66	6.7	231.0
10	-0.391	41	0.037	1405	4.94	4.2	225.8
11	-0.350	38	0.019	86	0.18	2.1	23.2
12	-0.363	42	0.016	2697	8.43	1.6	152.9
13	-0.410	46	0.065	3340	10.24	7.2	138.8
14	-0.411	46	0.050	3314	10.29	5.3	139.0
15	-0.342	38	0.006	-130	-0.80	0.6	-46.5
16a	-0.410	46	0.080	3153	9.82	10.1	144.3
16b	-0.401	45	0.081	3019	9.56	9.5	145.4
17	-0.412	47	0.049	3654	11.04	5.2	129.3
18	-0.415	52	0.081	5598	15.02	8.2	122.6
19	-0.369	41	0.156	4149	13.71	15.6	208.1
20	-0.408	59	0.138	9900	24.06	21.9	195.2
21	-0.338	37	0.189	2506	8.31	21.7	139.4
22	-0.292	34	0.020	-1016	-5.49	2.3	-49.2

**Table C9:** *The  $\Lambda$ -diagnostic for the first excited state computed at the B3LYP and CAM-B3LYP/6-31+G(d,p) level of theory in PCM ( $H_2O$ ) for the RFP-derived models (see Figure C1).*

Model	B3LYP	CAM-B3LYP
1a	0.804	0.801
1b	0.804	0.799
2a	0.802	0.796
2b	0.804	0.798
3	0.803	0.797
4	0.807	0.801
5	0.788	0.787
6	0.791	0.788
7	0.803	0.796
8	0.811	0.807
9	0.787	0.775
10	0.793	0.780
11	0.808	0.804
12	0.776	0.755
13	0.756	0.738
14	0.764	0.748
15	0.804	0.801
16a	0.722	0.716
16b	0.730	0.715
17	0.757	0.740
18	0.712	0.690
19	0.738	0.726
20	0.615	0.608
21	0.660	0.670
22	0.738	0.743

**Table C10:** *One-photon excitation energies [in eV] and the corresponding wavelengths [in nm], OPA oscillator strengths, TPA transition moments ( $\delta^{TPA}$ ) and TPA cross-sections [in GM] for the transition to  $S_1$  as determined at the B3LYP/6-31G+(d,p) Level of Theory (except where noted) and PCM with parameters for  $H_2O$ . The properties are computed for the RFP-derived chromophores (Figure C1) at varying Acylimine dihedral angle ( $\theta_{acylimine}$ ).*

Model	$\theta_{acylimine}$	Energy [eV]	Wavelength [nm]	Oscillator Strength	$\delta^{TPA}$	TPA [GM]
13	0°	2.526	491	0.717	15955	37.33
	10°	2.559	484	0.735	14871	35.72
	20°	2.597	477	0.750	13483	33.35
	30°	2.636	470	0.767	12077	30.77
	40°	2.678	463	0.785	10597	27.87
	50°	2.720	456	0.804	9181	24.91
	60°	2.756	450	0.820	7974	22.21
	70°	2.787	445	0.835	6957	19.81
	80°	2.809	441	0.844	6246	18.07
	90°	2.819	440	0.848	5921	17.25
	100°	2.814	441	0.843	6057	17.58
	110°	2.791	444	0.833	6782	19.37
	120°	2.750	451	0.815	8151	22.60
	130°	2.693	461	0.794	10149	26.98
	140°	2.630	471	0.775	12329	31.27
	150°	2.571	482	0.757	14133	34.25
	160°	2.539	488	0.746	15096	35.67
	170°	2.524	491	0.739	15669	36.61
	180°	2.525	491	0.735	15945	37.29
	190°	2.535	489	0.734	16150	38.06
	200°	2.565	483	0.736	15796	38.12
	210°	2.628	472	0.751	14012	35.50
	220°	2.703	459	0.779	11469	30.72
	230°	2.772	447	0.818	8936	25.17
	240°	2.828	438	0.866	6781	19.88
	250°	2.856	434	0.900	5653	16.91
	260°	2.849	435	0.899	5813	17.31
	270°	2.809	441	0.864	7244	20.96
	280°	2.750	451	0.819	9374	26.00
	290°	2.685	462	0.780	11783	31.14
	300°	2.619	473	0.746	14080	35.41
	310°	2.563	484	0.720	15802	38.05
	320°	2.523	491	0.702	16666	38.90
	330°	2.497	497	0.686	16886	38.60
	340°	2.490	498	0.677	16538	37.59
350°	2.500	496	0.690	16386	37.54	
14	0°	2.544	487	0.725	16363	38.82
	10°	2.577	481	0.746	15350	37.38
	20°	2.615	474	0.763	13884	34.81
	30°	2.655	467	0.781	12406	32.06
	40°	2.698	460	0.801	10856	28.97

50°	2.740	453	0.821	9394	25.86	
60°	2.778	446	0.840	8092	22.90	
70°	2.809	441	0.856	7031	20.34	
80°	2.831	438	0.866	6284	18.47	
90°	2.840	437	0.870	5980	17.68	
100°	2.834	438	0.867	6160	18.13	
110°	2.812	441	0.856	6869	19.92	
120°	2.774	447	0.838	8196	23.12	
130°	2.720	456	0.816	10119	27.44	
140°	2.659	466	0.794	12277	31.82	
150°	2.599	477	0.772	14205	35.18	
160°	2.566	483	0.759	15230	36.76	
170°	2.545	487	0.750	16004	38.02	
180°	2.540	488	0.746	16559	39.17	
190°	2.555	485	0.748	16481	39.44	
200°	2.600	477	0.758	15426	38.23	
210°	2.662	466	0.778	13524	35.15	
220°	2.730	454	0.806	11193	30.58	
230°	2.797	443	0.847	8679	24.89	
240°	2.849	435	0.892	6646	19.79	
250°	2.875	431	0.923	5619	17.03	
260°	2.869	432	0.922	5847	17.64	
270°	2.831	438	0.889	7208	21.18	
280°	2.775	447	0.847	9302	26.26	
290°	2.710	458	0.805	11724	31.58	
300°	2.647	468	0.770	13972	35.89	
310°	2.591	479	0.741	15747	38.76	
320°	2.549	487	0.718	16733	39.85	
330°	2.521	492	0.697	16908	39.42	
340°	2.511	494	0.674	16444	38.00	
350°	2.519	492	0.684	16258	37.84	
16a	0°	2.482	500	0.571	17136	38.71
	10°	2.513	493	0.580	15911	36.85
	20°	2.550	486	0.589	14462	34.47
	30°	2.589	479	0.599	12964	31.85
	40°	2.632	471	0.611	11389	28.92
	50°	2.674	464	0.623	9903	25.97
	60°	2.712	457	0.632	8630	23.27
	70°	2.740	453	0.639	7652	21.07
	80°	2.762	449	0.644	6944	19.42
	90°	2.773	447	0.646	6618	18.65
	100°	2.767	448	0.643	6801	19.09
	110°	2.744	452	0.636	7543	20.82
	120°	2.702	459	0.627	8868	23.73
	130°	2.652	468	0.616	10729	27.66
	140°	2.585	480	0.603	12933	31.69
	150°	2.531	490	0.592	14798	34.76
	160°	2.494	497	0.584	16020	36.55
	170°	2.476	501	0.580	16803	37.77
	180°	2.472	502	0.577	17234	38.60

190°	2.480	500	0.577	17437	39.33
200°	2.526	491	0.584	16337	38.22
210°	2.596	478	0.600	14059	34.74
220°	2.667	465	0.619	11659	30.40
230°	2.736	453	0.646	9174	25.18
240°	2.789	445	0.674	7245	20.66
250°	2.814	441	0.691	6287	18.26
260°	2.807	442	0.688	6479	18.72
270°	2.768	448	0.667	7813	21.94
280°	2.713	457	0.642	9792	26.43
290°	2.648	468	0.616	12117	31.15
300°	2.583	480	0.595	14449	35.36
310°	2.528	490	0.578	16321	38.25
320°	2.487	499	0.567	17500	39.69
330°	2.464	503	0.560	18032	40.13
340°	2.451	506	0.555	18137	39.96
350°	2.459	504	0.560	17859	39.59

16b	0°	2.515	493	0.667	16785	38.94
	10°	2.547	487	0.682	15740	37.43
	20°	2.585	480	0.696	14282	34.99
	30°	2.627	472	0.711	12710	32.17
	40°	2.670	464	0.727	11182	29.24
	50°	2.714	457	0.744	9659	26.09
	60°	2.753	450	0.759	8317	23.12
	70°	2.785	445	0.771	7246	20.61
	80°	2.805	442	0.777	6565	18.95
	90°	2.813	441	0.780	6279	18.22
	100°	2.807	442	0.777	6460	18.67
	110°	2.784	445	0.769	7197	20.46
	120°	2.744	452	0.755	8564	23.64
	130°	2.688	461	0.737	10500	27.82
	140°	2.627	472	0.721	12612	31.91
	150°	2.572	482	0.706	14369	34.84
	160°	2.539	488	0.696	15452	36.52
	170°	2.514	493	0.686	16463	38.16
	180°	2.509	494	0.681	16963	39.16
	190°	2.523	492	0.681	16992	39.65
	200°	2.566	483	0.689	16032	38.71
	210°	2.633	471	0.707	14021	35.64
	220°	2.703	459	0.731	11634	31.18
	230°	2.770	448	0.764	9172	25.81
	240°	2.824	439	0.805	7090	20.73
	250°	2.850	435	0.831	6019	17.93
	260°	2.842	436	0.827	6233	18.46
	270°	2.804	442	0.798	7570	21.83
	280°	2.747	451	0.762	9679	26.78
	290°	2.683	462	0.729	12030	31.76
	300°	2.618	474	0.700	14336	36.04
	310°	2.563	484	0.677	16122	38.82
	320°	2.522	492	0.662	17116	39.92

	330°	2.495	497	0.648	17483	39.90
	340°	2.485	499	0.637	17267	39.09
	350°	2.493	497	0.645	17060	38.88
17	0°	2.520	492	0.737	17470	40.68
	10°	2.552	486	0.753	16331	38.99
	20°	2.591	479	0.770	14797	36.41
	30°	2.631	471	0.789	13272	33.68
	40°	2.675	464	0.810	11596	30.42
	50°	2.718	456	0.831	10044	27.20
	60°	2.753	450	0.848	8798	24.45
	70°	2.785	445	0.864	7682	21.84
	80°	2.806	442	0.875	6934	20.02
	90°	2.816	440	0.879	6597	19.18
	100°	2.811	441	0.876	6754	19.57
	110°	2.789	445	0.864	7521	21.45
	120°	2.747	451	0.845	9009	24.93
	130°	2.692	461	0.823	11033	29.32
	140°	2.633	471	0.801	13211	33.58
	150°	2.574	482	0.780	15165	36.83
	160°	2.542	488	0.768	16146	38.27
	170°	2.520	492	0.758	17016	39.61
	180°	2.518	492	0.754	17381	40.40
	190°	2.530	490	0.755	17491	41.04
	200°	2.571	482	0.763	16636	40.33
	210°	2.634	471	0.781	14747	37.51
	220°	2.702	459	0.809	12316	32.98
	230°	2.769	448	0.849	9723	27.34
	240°	2.825	439	0.900	7463	21.84
	250°	2.852	435	0.934	6317	18.85
	260°	2.844	436	0.930	6547	19.42
	270°	2.806	442	0.896	8004	23.11
	280°	2.749	451	0.852	10232	28.36
	290°	2.685	462	0.809	12758	33.72
	300°	2.621	473	0.774	15075	37.99
	310°	2.567	483	0.748	16826	40.65
	320°	2.523	492	0.727	18008	42.01
	330°	2.497	497	0.711	18255	41.74
	340°	2.487	499	0.701	18060	40.96
	350°	2.496	497	0.712	17865	40.82
18	0°	2.367	524	0.789	20503	42.12
	10°	2.398	517	0.802	19229	40.55
	20°	2.436	509	0.817	17739	38.60
	30°	2.476	501	0.835	16164	36.33
	40°	2.521	492	0.854	14445	33.67
	50°	2.565	483	0.877	12632	30.48
	60°	2.605	476	0.897	10989	27.35
	70°	2.637	470	0.914	9638	24.58
	80°	2.662	466	0.927	8631	22.42
	90°	2.674	464	0.930	8169	21.41

100°	2.668	465	0.927	8387	21.89
110°	2.643	469	0.913	9376	24.02
120°	2.600	477	0.892	11158	27.65
130°	2.541	488	0.869	13523	32.02
140°	2.478	500	0.848	15865	35.72
150°	2.423	512	0.831	17555	37.80
160°	2.384	520	0.819	18704	38.98
170°	2.364	524	0.811	19463	39.89
180°	2.360	525	0.804	20081	41.00
190°	2.369	523	0.800	20614	42.41
200°	2.411	514	0.803	20205	43.07
210°	2.471	502	0.811	18802	42.09
220°	2.540	488	0.831	16345	38.66
230°	2.613	475	0.871	13024	32.60
240°	2.674	464	0.929	9942	26.07
250°	2.708	458	0.977	8151	21.92
260°	2.701	459	0.977	8371	22.39
270°	2.658	466	0.932	10389	26.92
280°	2.597	478	0.883	13351	33.01
290°	2.530	490	0.842	16374	38.42
300°	2.464	503	0.810	19024	42.35
310°	2.409	515	0.789	20816	44.30
320°	2.371	523	0.777	21708	44.73
330°	2.347	528	0.772	21959	44.35
340°	2.338	530	0.772	21838	43.78
350°	2.346	528	0.779	21337	43.07

19

0°	2.501	496	0.739	26756	61.38
10°	2.530	490	0.750	25089	58.88
20°	2.563	484	0.759	22846	55.04
30°	2.600	477	0.768	20447	50.69
40°	2.638	470	0.777	18093	46.18
50°	2.677	463	0.786	15797	41.50
60°	2.711	457	0.793	13815	37.22
70°	2.738	453	0.796	12263	33.70
80°	2.755	450	0.797	11299	31.44
90°	2.762	449	0.798	10884	30.44
100°	2.758	450	0.798	11101	30.95
110°	2.739	453	0.796	12141	33.39
120°	2.704	459	0.790	14125	37.86
130°	2.655	467	0.781	17021	43.99
140°	2.600	477	0.771	20346	50.42
150°	2.549	486	0.759	23296	55.51
160°	2.518	493	0.752	25189	58.54
170°	2.494	497	0.743	26892	61.33
180°	2.489	498	0.739	27618	62.73
190°	2.502	496	0.741	27345	62.78
200°	2.547	487	0.756	25159	59.85
210°	2.614	474	0.780	21369	53.53
220°	2.679	463	0.807	17504	46.08
230°	2.738	453	0.832	13957	38.36

240°	2.782	446	0.854	11220	31.83
250°	2.801	443	0.863	9949	28.63
260°	2.793	444	0.857	10314	29.51
270°	2.763	449	0.841	12038	33.70
280°	2.716	457	0.820	14898	40.28
290°	2.661	466	0.798	18235	47.33
300°	2.602	477	0.775	21792	54.10
310°	2.551	486	0.755	24738	59.01
320°	2.513	494	0.740	26571	61.51
330°	2.486	499	0.727	27479	62.28
340°	2.475	501	0.715	27492	61.73
350°	2.483	499	0.720	27070	61.18

20

0°	2.053	604	0.493	44252	68.40
10°	2.062	601	0.493	43311	67.55
20°	2.087	594	0.493	41174	65.78
30°	2.120	585	0.492	38446	63.33
40°	2.156	575	0.491	35288	60.16
50°	2.199	564	0.490	31543	55.91
60°	2.236	555	0.488	28105	51.53
70°	2.267	547	0.486	25186	47.47
80°	2.287	542	0.484	23282	44.63
90°	2.292	541	0.482	22621	43.58
100°	2.283	543	0.482	23366	44.66
110°	2.258	549	0.483	25641	47.93
120°	2.217	559	0.485	29371	52.92
130°	2.166	572	0.488	33998	58.51
140°	2.117	586	0.492	38377	63.08
150°	2.074	598	0.494	41894	66.10
160°	2.044	607	0.496	44178	67.67
170°	2.030	611	0.496	45104	68.17
180°	2.027	612	0.495	45298	68.24
190°	2.038	608	0.493	44419	67.64
200°	2.054	604	0.492	43071	66.63
210°	2.076	597	0.490	41342	65.33
220°	2.110	588	0.488	38743	63.23
230°	2.155	575	0.486	34972	59.54
240°	2.203	563	0.484	30689	54.61
250°	2.247	552	0.483	26694	49.42
260°	2.277	544	0.481	23944	45.54
270°	2.292	541	0.481	22765	43.83
280°	2.287	542	0.482	23299	44.68
290°	2.266	547	0.483	25372	47.76
300°	2.233	555	0.485	28571	52.25
310°	2.193	565	0.488	32424	57.19
320°	2.150	577	0.489	36433	61.77
330°	2.109	588	0.490	40037	65.32
340°	2.079	596	0.492	42515	67.39
350°	2.058	602	0.493	44047	68.42
360°	2.053	604	0.493	44252	68.40

21	0°	2.455	505	0.468	31199	68.94
	10°	2.479	500	0.468	29438	66.31
	20°	2.509	494	0.469	27071	62.48
	30°	2.541	488	0.470	24655	58.37
	40°	2.573	482	0.469	22363	54.29
	50°	2.605	476	0.469	20194	50.25
	60°	2.634	471	0.468	18309	46.56
	70°	2.655	467	0.465	16891	43.67
	80°	2.669	465	0.462	16008	41.81
	90°	2.674	464	0.460	15656	41.04
	100°	2.669	465	0.461	15911	41.56
	110°	2.653	467	0.462	16890	43.60
	120°	2.624	473	0.463	18772	47.38
	130°	2.581	480	0.463	21600	52.77
	140°	2.531	490	0.461	25029	58.80
	150°	2.486	499	0.460	28215	63.92
	160°	2.461	504	0.461	30103	66.85
	170°	2.442	508	0.459	31717	69.34
	180°	2.440	508	0.459	32197	70.28
	190°	2.454	505	0.461	31560	69.71
	200°	2.497	497	0.472	28963	66.23
	210°	2.559	485	0.486	25026	60.09
	220°	2.619	473	0.498	21162	53.23
	230°	2.669	465	0.504	17949	46.87
	240°	2.704	459	0.506	15644	41.92
	250°	2.716	456	0.503	14683	39.73
	260°	2.708	458	0.498	15057	40.49
	270°	2.683	462	0.494	16598	43.79
	280°	2.642	469	0.490	19155	49.03
	290°	2.595	478	0.486	22253	54.95
	300°	2.545	487	0.480	25671	60.95
	310°	2.499	496	0.475	28743	65.82
	320°	2.466	503	0.472	30849	68.77
	330°	2.442	508	0.469	32146	70.28
	340°	2.433	510	0.466	32312	70.13
	350°	2.438	509	0.466	32000	69.76
21-H	0°	2.185	568	0.328	40977	71.73
	10°	2.195	565	0.329	40302	71.23
	20°	2.227	557	0.333	38478	69.99
	30°	2.273	546	0.339	35865	67.94
	40°	2.329	532	0.345	32594	64.82
	50°	2.387	519	0.352	29097	60.80
	60°	2.438	509	0.355	25992	56.64
	70°	2.478	500	0.357	23551	53.01
	80°	2.500	496	0.357	22166	50.79
	90°	2.505	495	0.358	21918	50.43
	100°	2.492	498	0.358	22881	52.11
	110°	2.462	504	0.358	25008	55.57
	120°	2.418	513	0.357	28008	60.04
	130°	2.370	523	0.354	31229	64.34

140°	2.324	534	0.351	34277	67.87
150°	2.281	544	0.346	36898	70.37
160°	2.247	552	0.342	38835	71.86
170°	2.222	558	0.338	40135	72.69
180°	2.216	560	0.337	40528	72.94
190°	2.221	558	0.337	40304	72.87
200°	2.240	554	0.340	39320	72.32
210°	2.268	547	0.343	37759	71.21
220°	2.310	537	0.348	35237	68.97
230°	2.360	525	0.353	32099	65.54
240°	2.409	515	0.356	28773	61.23
250°	2.453	506	0.358	25721	56.75
260°	2.486	499	0.358	23389	53.00
270°	2.504	495	0.358	22098	50.82
280°	2.504	495	0.358	22090	50.77
290°	2.483	499	0.358	23384	52.84
300°	2.447	507	0.357	25697	56.42
310°	2.398	517	0.354	28831	60.78
320°	2.341	530	0.349	32344	64.98
330°	2.284	543	0.343	35702	68.30
340°	2.234	555	0.336	38504	70.47
350°	2.198	564	0.330	40371	71.54

22

0°	2.478	500	0.013	74	0.17
10°	2.562	484	0.012	68	0.16
20°	2.627	472	0.672	3697	9.35
30°	2.646	469	0.678	4279	10.98
40°	2.665	465	0.683	4988	12.99
50°	2.682	462	0.685	5703	15.04
60°	2.697	460	0.689	6327	16.88
70°	2.709	458	0.689	6842	18.42
80°	2.718	456	0.689	7172	19.42
90°	2.722	456	0.688	7298	19.82
100°	2.721	456	0.688	7140	19.38
110°	2.714	457	0.688	6696	18.08
120°	2.700	459	0.688	5931	15.86
130°	2.680	463	0.687	4991	13.15
140°	2.655	467	0.685	4071	10.52
150°	2.628	472	0.679	3394	8.59
160°	2.607	476	0.670	3076	7.67
170°	2.596	478	0.668	2895	7.15
180°	2.596	478	0.669	2875	7.10
190°	2.607	476	0.678	3007	7.49
200°	2.631	471	0.694	3594	9.12
210°	2.660	466	0.710	4743	12.31
220°	2.687	461	0.722	6187	16.38
230°	2.709	458	0.728	7557	20.33
240°	2.724	455	0.729	8552	23.26
250°	2.730	454	0.727	8880	24.28
260°	2.728	454	0.723	8598	23.47
270°	2.719	456	0.718	7697	20.86

	280°	2.697	460	0.709	6662	17.77
	290°	2.673	464	0.699	5457	14.30
	300°	2.649	468	0.686	4286	11.03
	310°	2.620	473	0.659	3498	8.80
	320°	2.587	479	0.573	2711	6.65
	330°	2.532	490	0.231	1118	2.63
	340°	2.465	503	0.055	312	0.69
	350°	2.443	508	0.023	136	0.30
BLB	0°	2.642	469	0.020	160	0.41
	10°	2.735	453	0.090	921	2.53
	20°	2.857	434	0.003	82	0.25
	30°	2.812	441	0.634	4517	13.10
	40°	2.852	435	0.648	3728	11.12
	50°	2.890	429	0.661	3016	9.24
	60°	2.928	424	0.674	2345	7.37
	70°	2.957	419	0.682	1860	5.96
	80°	2.977	416	0.687	1538	5.00
	90°	2.986	415	0.689	1401	4.58
	100°	2.981	416	0.686	1462	4.76
	110°	2.960	419	0.679	1785	5.73
	120°	2.920	425	0.668	2456	7.68
	130°	2.866	433	0.654	3490	10.51
	140°	2.806	442	0.639	4680	13.51
	150°	2.752	451	0.624	5745	15.96
	160°	2.718	456	0.614	6439	17.44
	170°	2.700	459	0.609	6948	18.57
	180°	2.698	460	0.608	7169	19.14
	190°	2.710	458	0.610	7165	19.30
	200°	2.752	451	0.620	6500	18.05
	210°	2.824	439	0.640	5075	14.84
	220°	2.898	428	0.666	3577	11.01
	230°	2.963	419	0.695	2302	7.41
	240°	3.010	412	0.721	1451	4.82
	250°	3.028	410	0.733	1149	3.86
	260°	3.016	411	0.728	1301	4.34
	270°	2.981	416	0.709	1864	6.07
	280°	2.929	423	0.684	2808	8.83
	290°	2.870	432	0.657	3925	11.85
	300°	2.809	441	0.629	5070	14.67
	310°	2.755	450	0.597	5894	16.40
	320°	2.710	458	0.519	5603	15.08
	330°	2.657	467	0.251	2585	6.69
	340°	2.601	477	0.065	542	1.34
	350°	2.593	478	0.028	204	0.50
RFP	0°	2.222	558	1.028	4389	7.94
	10°	2.238	554	1.031	4310	7.91
	20°	2.261	548	1.033	4299	8.05
	30°	2.287	542	1.036	4304	8.25
	40°	2.318	535	1.040	4226	8.32

50°	2.354	527	1.046	3962	8.04
60°	2.393	518	1.056	3428	7.19
70°	2.430	510	1.066	2744	5.94
80°	2.456	505	1.075	2170	4.79
90°	2.465	503	1.079	1938	4.31
100°	2.452	506	1.072	2188	4.82
110°	2.418	513	1.058	2902	6.22
120°	2.369	523	1.045	3818	7.85
130°	2.317	535	1.037	4464	8.78
140°	2.272	546	1.037	4546	8.60
150°	2.241	553	1.042	4254	7.83
160°	2.226	557	1.047	3918	7.11
170°	2.217	559	1.050	3695	6.65
180°	2.216	560	1.049	3676	6.61
190°	2.218	559	1.044	3838	6.92
200°	2.227	557	1.038	4141	7.52
210°	2.241	553	1.031	4505	8.29
220°	2.265	547	1.025	4836	9.09
230°	2.301	539	1.021	4888	9.48
240°	2.348	528	1.026	4435	8.96
250°	2.397	517	1.041	3529	7.43
260°	2.438	509	1.060	2612	5.69
270°	2.460	504	1.073	2124	4.71
280°	2.455	505	1.073	2270	5.01
290°	2.427	511	1.061	3003	6.48
300°	2.380	521	1.044	4029	8.36
310°	2.326	533	1.029	4872	9.66
320°	2.277	545	1.020	5216	9.91
330°	2.241	553	1.017	5151	9.48
340°	2.219	559	1.020	4849	8.75
350°	2.213	560	1.024	4549	8.16

RFP-H

0°	2.227	557	1.149	677	1.23
10°	2.230	556	1.148	704	1.28
20°	2.243	553	1.146	794	1.46
30°	2.262	548	1.142	926	1.74
40°	2.289	542	1.138	1074	2.06
50°	2.323	534	1.137	1157	2.29
60°	2.359	526	1.137	1093	2.23
70°	2.392	518	1.141	897	1.88
80°	2.413	514	1.143	727	1.55
90°	2.413	514	1.140	763	1.63
100°	2.392	518	1.134	1033	2.17
110°	2.360	526	1.134	1356	2.77
120°	2.328	533	1.146	1463	2.91
130°	2.303	538	1.164	1317	2.56
140°	2.289	542	1.185	1065	2.05
150°	2.281	544	1.203	837	1.60
160°	2.277	545	1.218	679	1.29
170°	2.277	545	1.227	597	1.13
180°	2.277	545	1.230	575	1.09

	190°	2.278	544	1.227	599	1.14
	200°	2.278	544	1.218	678	1.29
	210°	2.280	544	1.204	826	1.58
	220°	2.289	542	1.185	1056	2.03
	230°	2.304	538	1.165	1310	2.55
	240°	2.329	532	1.146	1454	2.89
	250°	2.360	525	1.135	1352	2.76
	260°	2.393	518	1.134	1036	2.17
	270°	2.414	514	1.141	763	1.63
	280°	2.415	513	1.144	730	1.56
	290°	2.394	518	1.141	899	1.89
	300°	2.361	525	1.137	1099	2.25
	310°	2.324	533	1.137	1168	2.31
	320°	2.291	541	1.139	1087	2.09
	330°	2.264	548	1.142	941	1.77
	340°	2.243	553	1.146	799	1.47
	350°	2.231	556	1.148	708	1.29
	0°	2.382	521	1.189	6552	13.62
RFP	10°	2.396	517	1.193	6188	13.03
(CAM-B3LYP)	20°	2.419	513	1.197	5763	12.37
	30°	2.445	507	1.202	5306	11.63
	40°	2.475	501	1.208	4753	10.67
	50°	2.506	495	1.215	4113	9.47
	60°	2.538	489	1.221	3420	8.08
	70°	2.564	484	1.224	2816	6.79
	80°	2.580	481	1.226	2411	5.89
	90°	2.585	480	1.226	2269	5.56
	100°	2.577	481	1.223	2436	5.93
	110°	2.555	485	1.219	2935	7.02
	120°	2.521	492	1.214	3744	8.72
	130°	2.478	500	1.210	4721	10.63
	140°	2.435	509	1.209	5554	12.07
	150°	2.400	517	1.209	6007	12.69
	160°	2.379	521	1.211	6100	12.66
	170°	2.366	524	1.213	6067	12.46
	180°	2.364	524	1.212	6057	12.41
	190°	2.369	523	1.209	6119	12.59
	200°	2.382	521	1.206	6164	12.83
	210°	2.402	516	1.203	6086	12.88
	220°	2.431	510	1.202	5754	12.47
	230°	2.468	502	1.203	5090	11.37
	240°	2.508	494	1.207	4195	9.68
	250°	2.545	487	1.214	3318	7.88
	260°	2.571	482	1.219	2693	6.53
	270°	2.583	480	1.223	2408	5.89
	280°	2.581	480	1.224	2499	6.10
	290°	2.564	484	1.220	2959	7.13
	300°	2.532	490	1.214	3759	8.84
	310°	2.491	498	1.205	4783	10.88
	320°	2.448	507	1.195	5774	12.68

330°	2.412	514	1.187	6497	13.86
340°	2.386	520	1.185	6820	14.24
350°	2.375	522	1.186	6804	14.07

---

**Table C11:** *TPA Cross-sections [in GM] extracted from Tables C10 and C12 as determined at the B3LYP/6-31G+(d,p) Level of Theory (except where noted) with PCM (H<sub>2</sub>O). The Cross-sections correspond to the S<sub>0</sub> to <sub>1</sub> transition except for Models 22 and BLB at the near-planar conformations ( $\theta_{acylimine} = 0, 10, 20, 340$  and  $350$ ) where the transition is to the second excited state.*

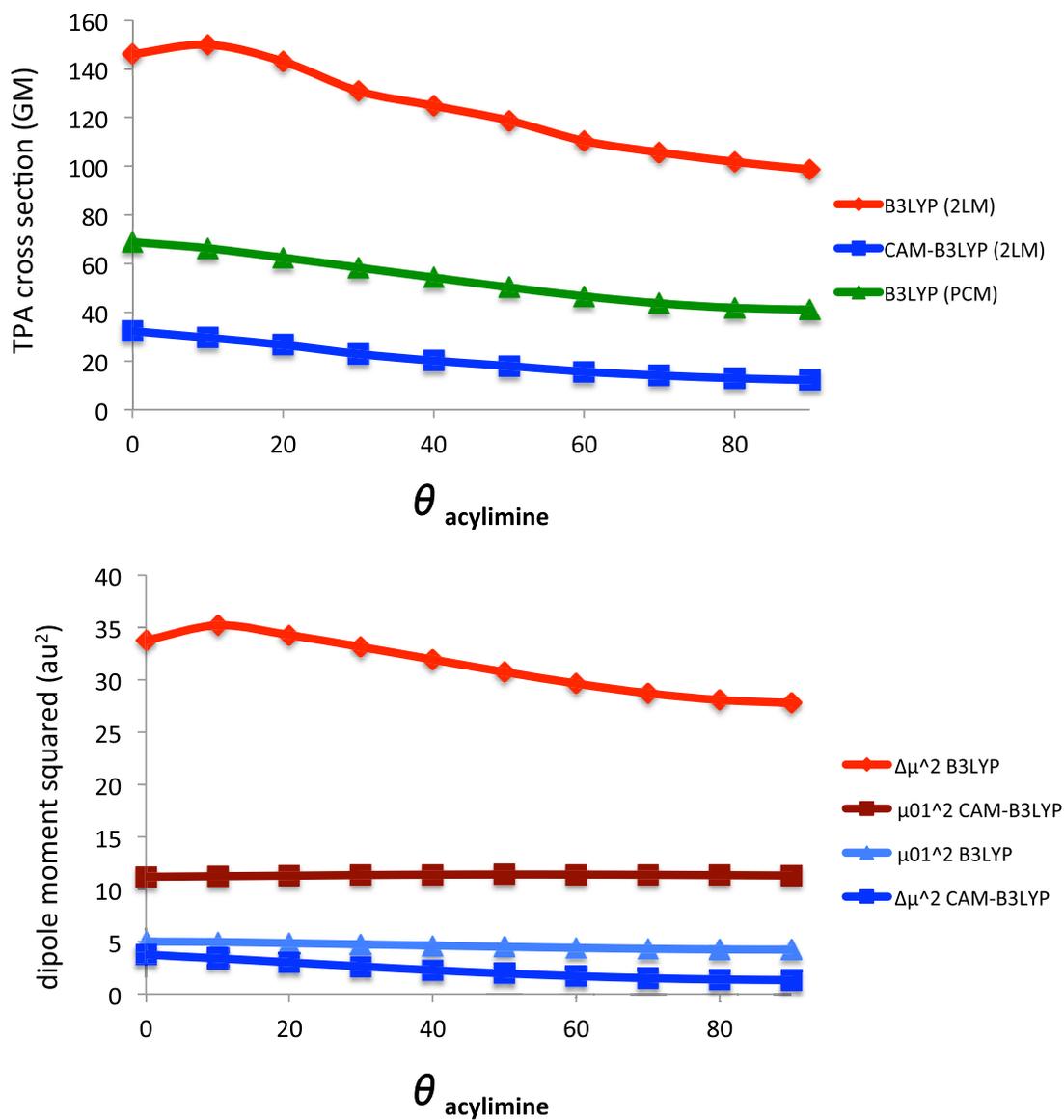
Model	13	14	16a	16b	17	18	19	20	21	21-H	22	BLB	RFP	RFP-H	RFP (CAM-B3LYP)
0°	37.33	38.82	38.71	38.94	40.68	42.12	61.38	68.40	68.94	71.73	7.25	18.11	7.94	1.23	13.62
10°	35.72	37.38	36.85	37.43	38.99	40.55	58.88	67.55	66.31	71.23	8.06	14.42	7.91	1.28	13.03
20°	33.35	34.81	34.47	34.99	36.41	38.60	55.04	65.78	62.48	69.99	9.35	15.02	8.05	1.46	12.37
30°	30.77	32.06	31.85	32.17	33.68	36.33	50.69	63.33	58.37	67.94	10.98	13.10	8.25	1.74	11.63
40°	27.87	28.97	28.92	29.24	30.42	33.67	46.18	60.16	54.29	64.82	12.99	11.12	8.32	2.06	10.67
50°	24.91	25.86	25.97	26.09	27.20	30.48	41.50	55.91	50.25	60.80	15.04	9.24	8.04	2.29	9.47
60°	22.21	22.90	23.27	23.12	24.45	27.35	37.22	51.53	46.56	56.64	16.88	7.37	7.19	2.23	8.08
70°	19.81	20.34	21.07	20.61	21.84	24.58	33.70	47.47	43.67	53.01	18.42	5.96	5.94	1.88	6.79
80°	18.07	18.47	19.42	18.95	20.02	22.42	31.44	44.63	41.81	50.79	19.42	5.00	4.79	1.55	5.89
90°	17.25	17.68	18.65	18.22	19.18	21.41	30.44	43.58	41.04	50.43	19.82	4.58	4.31	1.63	5.56
100°	17.58	18.13	19.09	18.67	19.57	21.89	30.95	44.66	41.56	52.11	19.38	4.76	4.82	2.17	5.93
110°	19.37	19.92	20.82	20.46	21.45	24.02	33.39	47.93	43.60	55.57	18.08	5.73	6.22	2.77	7.02
120°	22.60	23.12	23.73	23.64	24.93	27.65	37.86	52.92	47.38	60.04	15.86	7.68	7.85	2.91	8.72
130°	26.98	27.44	27.66	27.82	29.32	32.02	43.99	58.51	52.77	64.34	13.15	10.51	8.78	2.56	10.63
140°	31.27	31.82	31.69	31.91	33.58	35.72	50.42	63.08	58.80	67.87	10.52	13.51	8.60	2.05	12.07
150°	34.25	35.18	34.76	34.84	36.83	37.80	55.51	66.10	63.92	70.37	8.59	15.96	7.83	1.60	12.69
160°	35.67	36.76	36.55	36.52	38.27	38.98	58.54	67.67	66.85	71.86	7.67	17.44	7.11	1.29	12.66
170°	36.61	38.02	37.77	38.16	39.61	39.89	61.33	68.17	69.34	72.69	7.15	18.57	6.65	1.13	12.46
180°	37.29	39.17	38.60	39.16	40.40	41.00	62.73	68.24	70.28	72.94	7.10	19.14	6.61	1.09	12.41
190°	38.06	39.44	39.33	39.65	41.04	42.41	62.78	67.64	69.71	72.87	7.49	19.30	6.92	1.14	12.59
200°	38.12	38.23	38.22	38.71	40.33	43.07	59.85	66.63	66.23	72.32	9.12	18.05	7.52	1.29	12.83
210°	35.50	35.15	34.74	35.64	37.51	42.09	53.53	65.33	60.09	71.21	12.31	14.84	8.29	1.58	12.88
220°	30.72	30.58	30.40	31.18	32.98	38.66	46.08	63.23	53.23	68.97	16.38	11.01	9.09	2.03	12.47
230°	25.17	24.89	25.18	25.81	27.34	32.60	38.36	59.54	46.87	65.54	20.33	7.41	9.48	2.55	11.37
240°	19.88	19.79	20.66	20.73	21.84	26.07	31.83	54.61	41.92	61.23	23.26	4.82	8.96	2.89	9.68
250°	16.91	17.03	18.26	17.93	18.85	21.92	28.63	49.42	39.73	56.75	24.28	3.86	7.43	2.76	7.88
260°	17.31	17.64	18.72	18.46	19.42	22.39	29.51	45.54	40.49	53.00	23.47	4.34	5.69	2.17	6.53
270°	20.96	21.18	21.94	21.83	23.11	26.92	33.70	43.83	43.79	50.82	20.86	6.07	4.71	1.63	5.89
280°	26.00	26.26	26.43	26.78	28.36	33.01	40.28	44.68	49.03	50.77	17.77	8.83	5.01	1.56	6.10
290°	31.14	31.58	31.15	31.76	33.72	38.42	47.33	47.76	54.95	52.84	14.30	11.85	6.48	1.89	7.13
300°	35.41	35.89	35.36	36.04	37.99	42.35	54.10	52.25	60.95	56.42	11.03	14.67	8.36	2.25	8.84
310°	38.05	38.76	38.25	38.82	40.65	44.30	59.01	57.19	65.82	60.78	8.80	16.40	9.66	2.31	10.88
320°	38.90	39.85	39.69	39.92	42.01	44.73	61.51	61.77	68.77	64.98	6.65	15.08	9.91	2.09	12.68
330°	38.60	39.42	40.13	39.90	41.74	44.35	62.28	65.32	70.28	68.30	4.80	13.80	9.48	1.77	13.86
340°	37.59	38.00	39.96	39.09	40.96	43.78	61.73	67.39	70.13	70.47	6.37	19.10	8.75	1.47	14.24
350°	37.54	37.84	39.59	38.88	40.82	43.07	61.18	68.42	69.76	71.54	6.73	19.21	8.16	1.29	14.07

**Table C12:** TD-DFT data for the first 3 states of Models 22 and BLB at their near-planar conformations ( $\theta_{\text{acylimine}} = 0, 10, 20, 340$  and  $350$ ). Level of Theory is B3LYP/6-31G+(d,p) with PCM ( $\text{H}_2\text{O}$ ).

Model	$\theta_{\text{acylimine}}$	State	Energy [eV]	wavelength [nm]	Oscillator Strength	$\delta^{\text{TPA}}$	TPA [GM]	
22	0°	1	2.478	500	0.014	75	0.17	
		2	2.596	478	0.650	2938	7.25	
		3	3.055	406	0.000	4	0.01	
	10°	1	2.562	484	0.012	68	0.16	
		2	2.61	475	0.656	3231	8.06	
		3	3.074	403	0.000	4	0.01	
	20°	1	2.627	472	0.672	3699	9.35	
		2	2.675	464	0.003	71	0.19	
		3	3.098	400	0.000	4	0.01	
	330°	1	2.532	490	0.231	1119	2.63	
		2	2.622	473	0.435	1907	4.80	
		3	3.065	405	0.001	2	0.01	
	340°	1	2.465	503	0.056	313	0.70	
		2	2.593	478	0.606	2585	6.37	
		3	3.049	407	0.001	2	0.01	
	350°	1	2.443	508	0.023	137	0.30	
		2	2.588	479	0.637	2744	6.73	
		3	3.045	407	0.000	3	0.01	
	BLB	0°	1	2.642	469	0.02	160	0.41
			2	2.718	456	0.585	6693	18.11
			3	3.246	382	0.001	3	0.01
10°		1	2.735	453	0.09	919	2.52	
		2	2.745	452	0.524	5226	14.42	
		3	3.272	379	0.001	3	0.01	
20°		1	2.776	447	0.621	5319	15.02	
		2	2.857	434	0.003	75	0.23	
		3	3.303	375	0.001	5	0.02	
330°		1	2.657	467	0.251	2584	6.68	
		2	2.735	453	0.352	5035	13.80	
		3	3.248	382	0.001	4	0.02	
340°		1	2.601	477	0.065	542	1.34	
		2	2.704	459	0.534	7132	19.10	
		3	3.232	384	0.001	4	0.02	
350°	1	2.593	478	0.028	204	0.50		
	2	2.703	459	0.572	7178	19.21		
	3	3.231	384	0.001	3	0.01		

**Table C13:** Components contributing to the 2LM expression (Equation 5 in the main text) determined at the B3LYP and CAM-B3LYP/6-31+G(d,p) levels of theory in the gas phase for Model 21 at  $\theta_{\text{acylimine}}$  varying from  $0^\circ$  to  $90^\circ$ . All data is for the  $S_0$  to  $S_1$  transition except for the CAM-B3LYP data for conformers with  $\theta_{\text{acylimine}} = 0, 10$  in which data for the  $S_0$  to  $S_2$  is given because the first excited-state was dark.

$\theta_{\text{acylimine}}$	B3LYP				CAM-B3LYP			
	$\Delta\mu$ (au)	$\mu_{01}$ (au)	$\cos^2\theta$	TPA (GM)	$\Delta\mu$ (au)	$\mu_{01}$ (au)	$\cos^2\theta$	TPA (GM)
0	5.81	2.24	0.99	146.05	1.93	3.35	0.84	32.25
10	5.93	2.23	0.99	149.94	1.85	3.35	0.84	29.56
20	5.85	2.20	0.99	143.05	1.74	3.36	0.85	26.60
30	5.75	2.18	0.99	130.93	1.62	3.37	0.86	22.88
40	5.65	2.15	0.99	124.85	1.51	3.38	0.87	20.12
50	5.54	2.12	0.98	118.80	1.40	3.38	0.89	17.96
60	5.44	2.10	0.98	110.40	1.31	3.38	0.90	15.58
70	5.36	2.08	0.97	105.71	1.23	3.37	0.91	14.03
80	5.30	2.06	0.97	101.80	1.18	3.37	0.91	12.88
90	5.27	2.06	0.97	98.68	1.15	3.36	0.92	12.12



**Figure C4:** Top: TPA cross sections computed for Model 21 (at  $\theta_{\text{acylimine}}$  varying from  $0^\circ$  to  $90^\circ$ ) via the 2LM in the gas phase (as given in Table C13) versus the corresponding PCM values for B3LYP (given in Table C11). Bottom: The magnitudes of the dipole moments contributing to the 2LM expression (as given in Table C13).

**Table C14:** Variation of TD-DFT properties with tilt and twist angles (see Table C1) for Model 21 at fixed  $\theta_{\text{acylimine}}$  of  $0^\circ$ ,  $90^\circ$ ,  $180^\circ$  and  $270^\circ$ . The TD-DFT values are for the transition to  $S_1$  as determined at the B3LYP/6-31G+(d,p) level of theory and PCM with parameters for  $H_2O$ .

$\theta_{\text{acylimine}}$	tilt	twist	Energy [eV]	wavelength [nm]	Oscillator Strength	$\delta^{\text{TPA}}$	TPA [GM]
$0^\circ$	-30	160	2.401	516	0.317	30361	64.16
	-30	165	2.402	516	0.356	30743	65.04
	-30	170	2.401	516	0.391	30992	65.52
	-30	175	2.398	517	0.420	31296	65.99
	-30	180	2.392	518	0.441	31702	66.52
	-30	185	2.388	519	0.456	32059	67.03
	-30	190	2.391	519	0.471	32098	67.27
	-30	195	2.391	519	0.476	32374	67.85
	-30	200	2.389	519	0.471	32748	68.52
	-25	160	2.423	512	0.334	30060	64.69
	-25	165	2.422	512	0.371	30458	65.51
	-25	170	2.420	512	0.405	30766	66.08
	-25	175	2.414	514	0.428	31236	66.76
	-25	180	2.408	515	0.445	31741	67.46
	-25	185	2.407	515	0.461	31955	67.90
	-25	190	2.409	515	0.472	32079	68.27
	-25	195	2.406	515	0.469	32528	69.05
	-25	200	2.408	515	0.466	32560	69.21
	-20	160	2.437	509	0.351	30049	65.46
	-20	165	2.436	509	0.386	30461	66.27
	-20	170	2.431	510	0.414	30904	66.99
	-20	175	2.424	512	0.433	31461	67.78
	-20	180	2.420	512	0.450	31851	68.41
	-20	185	2.422	512	0.465	31971	68.77
	-20	190	2.421	512	0.467	32297	69.38
	-20	195	2.421	512	0.464	32433	69.68
	-20	200	2.425	511	0.459	32247	69.51
	-15	160	2.448	506	0.368	30161	66.29
	-15	165	2.445	507	0.400	30586	67.06
	-15	170	2.439	508	0.423	31096	67.83
	-15	175	2.433	510	0.440	31599	68.59
	-15	180	2.433	510	0.457	31772	68.97
	-15	185	2.434	509	0.466	31943	69.38
	-15	190	2.433	510	0.465	32149	69.77
	-15	195	2.437	509	0.462	31960	69.58
	-15	200	2.439	508	0.452	31898	69.58
	-10	160	2.457	505	0.386	30237	66.94
	-10	165	2.453	506	0.413	30662	67.65
	-10	170	2.446	507	0.432	31208	68.45
	-10	175	2.444	507	0.450	31462	68.90
	-10	180	2.445	507	0.463	31578	69.19
	-10	185	2.444	507	0.467	31750	69.54
	-10	190	2.447	507	0.466	31658	69.48
	-10	195	2.450	506	0.459	31476	69.25

-10	200	2.445	507	0.436	31687	69.44
-5	160	2.462	504	0.402	30399	67.54
-5	165	2.456	505	0.424	30855	68.22
-5	170	2.452	506	0.443	31205	68.77
-5	175	2.452	506	0.459	31283	68.94
-5	180	2.451	506	0.466	31420	69.22
-5	185	2.452	506	0.467	31410	69.25
-5	190	2.451	506	0.458	31362	69.10
-5	195	2.456	505	0.450	31156	68.92
-5	200	2.459	504	0.431	31030	68.78
0	160	2.461	504	0.415	30599	67.97
0	165	2.461	504	0.440	30787	68.36
0	170	2.459	504	0.456	30895	68.48
0	175	2.457	505	0.466	31009	68.65
0	180	2.455	505	0.468	31199	68.94
0	185	2.457	505	0.465	30996	68.62
0	190	2.458	504	0.455	30893	68.46
0	195	2.461	504	0.439	30722	68.22
0	200	2.463	503	0.417	30528	67.93
5	160	2.460	504	0.432	30916	68.59
5	165	2.457	505	0.450	31055	68.74
5	170	2.451	506	0.457	31183	68.71
5	175	2.452	506	0.466	31094	68.56
5	180	2.455	505	0.467	30880	68.24
5	185	2.457	505	0.461	30775	68.11
5	190	2.452	506	0.444	31068	68.49
5	195	2.462	504	0.429	30362	67.46
5	200	2.465	503	0.403	30101	67.05
10	160	2.447	507	0.437	31363	68.83
10	165	2.446	507	0.453	31351	68.76
10	170	2.447	507	0.465	31147	68.39
10	175	2.450	506	0.471	30888	67.95
10	180	2.452	506	0.468	30686	67.66
10	185	2.454	505	0.458	30552	67.45
10	190	2.458	505	0.443	30209	66.91
10	195	2.461	504	0.420	29997	66.60
10	200	2.463	503	0.390	29741	66.14
15	160	2.437	509	0.449	31583	68.79
15	165	2.439	509	0.463	31333	68.32
15	170	2.442	508	0.473	30939	67.63
15	175	2.445	507	0.475	30608	67.08
15	180	2.446	507	0.468	30488	66.90
15	185	2.450	506	0.456	30134	66.33
15	190	2.453	505	0.435	29900	65.98
15	195	2.456	505	0.408	29657	65.61
15	200	2.455	505	0.373	29523	65.26
20	160	2.424	512	0.458	31766	68.45
20	165	2.428	511	0.472	31237	67.52
20	170	2.431	510	0.478	30824	66.81
20	175	2.434	510	0.475	30610	66.47
20	180	2.436	509	0.464	30370	66.09

	20	185	2.440	508	0.449	30021	65.54
	20	190	2.443	508	0.425	29809	65.25
	20	195	2.444	507	0.393	29635	64.92
	20	200	2.440	508	0.355	29594	64.63
	25	160	2.410	515	0.467	31771	67.64
	25	165	2.414	514	0.478	31175	66.62
	25	170	2.417	513	0.480	30847	66.07
	25	175	2.419	513	0.472	30641	65.75
	25	180	2.423	512	0.460	30297	65.20
	25	185	2.426	511	0.439	30051	64.85
	25	190	2.428	511	0.412	29872	64.58
	25	195	2.425	511	0.376	29907	64.50
	25	200	2.420	513	0.336	29904	64.19
	30	160	2.394	518	0.475	31555	66.33
	30	165	2.397	517	0.482	31149	65.63
	30	170	2.399	517	0.478	30928	65.27
	30	175	2.401	516	0.469	30694	64.89
	30	180	2.404	516	0.452	30421	64.49
	30	185	2.407	515	0.428	30279	64.31
	30	190	2.406	515	0.396	30282	64.25
	30	195	2.398	517	0.357	30540	64.40
	30	200	2.391	519	0.314	30540	64.04
90°	-30	160	2.607	476	0.320	16778	41.80
	-30	165	2.605	476	0.356	16689	41.52
	-30	170	2.601	477	0.387	16577	41.13
	-30	175	2.596	478	0.413	16554	40.91
	-30	180	2.589	479	0.430	16649	40.92
	-30	185	2.585	480	0.444	16726	40.97
	-30	190	2.588	479	0.457	16625	40.82
	-30	195	2.588	479	0.461	16735	41.10
	-30	200	2.586	479	0.456	16948	41.57
	-25	160	2.635	471	0.337	16251	41.37
	-25	165	2.631	471	0.370	16206	41.13
	-25	170	2.627	472	0.400	16154	40.87
	-25	175	2.619	473	0.421	16276	40.93
	-25	180	2.611	475	0.435	16455	41.14
	-25	185	2.611	475	0.450	16458	41.13
	-25	190	2.613	475	0.460	16448	41.17
	-25	195	2.610	475	0.456	16704	41.71
	-25	200	2.612	475	0.452	16715	41.81
	-20	160	2.654	467	0.353	15964	41.24
	-20	165	2.650	468	0.385	15943	41.05
	-20	170	2.643	469	0.410	16016	41.03
	-20	175	2.634	471	0.426	16221	41.27
	-20	180	2.630	471	0.442	16323	41.40
	-20	185	2.632	471	0.455	16295	41.37
	-20	190	2.630	471	0.456	16432	41.69
	-20	195	2.630	471	0.453	16508	41.88
	-20	200	2.636	470	0.449	16385	41.74
	-15	160	2.668	465	0.369	15782	41.21

-15	165	2.663	466	0.398	15795	41.08
-15	170	2.655	467	0.418	15949	41.22
-15	175	2.648	468	0.434	16116	41.43
-15	180	2.648	468	0.449	16108	41.40
-15	185	2.649	468	0.458	16105	41.44
-15	190	2.648	468	0.456	16203	41.66
-15	195	2.653	467	0.454	16078	41.48
-15	200	2.657	467	0.446	16063	41.58
-10	160	2.679	463	0.385	15626	41.13
-10	165	2.673	464	0.411	15686	41.09
-10	170	2.664	465	0.428	15876	41.31
-10	175	2.661	466	0.445	15904	41.31
-10	180	2.662	466	0.457	15879	41.26
-10	185	2.662	466	0.460	15914	41.36
-10	190	2.665	465	0.458	15874	41.33
-10	195	2.669	465	0.453	15764	41.18
-10	200	2.663	466	0.430	16143	41.97
-5	160	2.685	462	0.400	15556	41.11
-5	165	2.676	463	0.421	15685	41.20
-5	170	2.671	464	0.438	15789	41.30
-5	175	2.669	465	0.452	15776	41.22
-5	180	2.669	465	0.458	15800	41.28
-5	185	2.671	464	0.459	15698	41.05
-5	190	2.674	464	0.454	15693	41.13
-5	195	2.676	463	0.443	15716	41.27
-5	200	2.681	463	0.425	15742	41.48
0	160	2.683	462	0.412	15633	41.26
0	165	2.682	462	0.436	15543	41.00
0	170	2.678	463	0.450	15544	40.88
0	175	2.676	463	0.459	15552	40.84
0	180	2.674	464	0.460	15656	41.04
0	185	2.675	464	0.455	15651	41.05
0	190	2.677	463	0.446	15606	41.02
0	195	2.681	462	0.432	15594	41.11
0	200	2.686	462	0.412	15598	41.27
5	160	2.682	462	0.428	15649	41.27
5	165	2.669	465	0.437	15900	41.54
5	170	2.669	465	0.449	15816	41.30
5	175	2.670	464	0.456	15722	41.08
5	180	2.672	464	0.458	15597	40.84
5	185	2.676	463	0.452	15524	40.76
5	190	2.679	463	0.440	15489	40.76
5	195	2.676	463	0.420	15733	41.32
5	200	2.688	461	0.401	15502	41.08
10	160	2.664	465	0.430	16056	41.79
10	165	2.662	466	0.446	15965	41.50
10	170	2.664	465	0.456	15786	41.08
10	175	2.667	465	0.461	15607	40.69
10	180	2.671	464	0.460	15465	40.44
10	185	2.673	464	0.450	15462	40.50
10	190	2.677	463	0.436	15378	40.42

10	195	2.681	462	0.415	15419	40.65	
10	200	2.685	462	0.388	15477	40.92	
15	160	2.652	468	0.441	16187	41.75	
15	165	2.653	467	0.454	15972	41.22	
15	170	2.656	467	0.463	15701	40.62	
15	175	2.660	466	0.466	15481	40.17	
15	180	2.663	466	0.459	15453	40.17	
15	185	2.666	465	0.447	15379	40.08	
15	190	2.670	464	0.428	15373	40.18	
15	195	2.674	464	0.404	15437	40.48	
15	200	2.676	463	0.372	15590	40.92	
20	160	2.637	470	0.449	16258	41.46	
20	165	2.641	470	0.462	15911	40.68	
20	170	2.645	469	0.468	15631	40.08	
20	175	2.648	468	0.466	15508	39.86	
20	180	2.650	468	0.455	15487	39.87	
20	185	2.654	467	0.441	15402	39.77	
20	190	2.657	467	0.418	15487	40.09	
20	195	2.660	466	0.390	15591	40.46	
20	200	2.660	466	0.355	15791	40.98	
25	160	2.620	473	0.456	16263	40.92	
25	165	2.624	473	0.467	15868	40.07	
25	170	2.628	472	0.470	15658	39.64	
25	175	2.629	472	0.463	15630	39.62	
25	180	2.632	471	0.451	15570	39.55	
25	185	2.635	471	0.431	15624	39.78	
25	190	2.638	470	0.406	15744	40.18	
25	195	2.639	470	0.374	15965	40.76	
25	200	2.636	470	0.337	16214	41.33	
30	160	2.599	477	0.463	16246	40.23	
30	165	2.602	477	0.470	15977	39.65	
30	170	2.603	476	0.467	15913	39.53	
30	175	2.604	476	0.457	15907	39.56	
30	180	2.606	476	0.441	15924	39.67	
30	185	2.609	475	0.419	16061	40.08	
30	190	2.610	475	0.391	16257	40.61	
30	195	2.607	476	0.355	16616	41.40	
30	200	2.603	476	0.317	16897	41.98	
180°	-30	160	2.387	519	0.314	30663	64.07
	-30	165	2.388	519	0.352	31072	64.97
	-30	170	2.387	519	0.387	31380	65.57
	-30	175	2.383	520	0.414	31780	66.15
	-30	180	2.375	522	0.432	32271	66.76
	-30	185	2.373	523	0.449	32559	67.23
	-30	190	2.375	522	0.463	32667	67.57
	-30	195	2.373	523	0.465	33125	68.39
	-30	200	2.373	523	0.464	33345	68.83
	-25	160	2.408	515	0.330	30444	64.73
	-25	165	2.408	515	0.366	30871	65.63
	-25	170	2.405	516	0.398	31251	66.30

-25	175	2.398	517	0.420	31786	67.04
-25	180	2.393	518	0.437	32263	67.73
-25	185	2.394	518	0.455	32444	68.17
-25	190	2.393	518	0.462	32774	68.82
-25	195	2.391	519	0.461	33153	69.48
-25	200	2.394	518	0.460	33086	69.53
-20	160	2.423	512	0.345	30512	65.68
-20	165	2.422	512	0.380	30944	66.54
-20	170	2.416	513	0.406	31466	67.35
-20	175	2.409	515	0.425	32057	68.20
-20	180	2.408	515	0.444	32343	68.75
-20	185	2.409	515	0.458	32537	69.22
-20	190	2.406	515	0.458	32960	69.95
-20	195	2.408	515	0.459	32935	70.03
-20	200	2.410	514	0.452	32928	70.13
-15	160	2.433	510	0.362	30743	66.73
-15	165	2.430	510	0.393	31192	67.54
-15	170	2.423	512	0.414	31787	68.43
-15	175	2.420	512	0.434	32134	69.03
-15	180	2.422	512	0.451	32293	69.43
-15	185	2.421	512	0.457	32573	69.99
-15	190	2.422	512	0.458	32625	70.17
-15	195	2.426	511	0.457	32430	69.97
-15	200	2.428	511	0.446	32464	70.16
-10	160	2.443	508	0.379	30828	67.46
-10	165	2.438	509	0.404	31311	68.25
-10	170	2.434	509	0.426	31742	68.96
-10	175	2.434	509	0.445	31888	69.29
-10	180	2.435	509	0.456	32070	69.71
-10	185	2.435	509	0.460	32201	70.00
-10	190	2.438	509	0.460	32049	69.85
-10	195	2.440	508	0.453	31972	69.80
-10	200	2.442	508	0.438	31954	69.89
-5	160	2.449	506	0.394	30945	68.07
-5	165	2.443	508	0.416	31478	68.90
-5	170	2.442	508	0.437	31712	69.33
-5	175	2.442	508	0.452	31832	69.61
-5	180	2.442	508	0.459	31968	69.89
-5	185	2.444	507	0.461	31882	69.83
-5	190	2.446	507	0.457	31735	69.63
-5	195	2.447	507	0.444	31760	69.72
-5	200	2.450	506	0.426	31614	69.56
0	160	2.448	507	0.406	31363	68.91
0	165	2.445	507	0.428	31705	69.47
0	170	2.444	507	0.446	31779	69.62
0	175	2.442	508	0.456	32024	70.04
0	180	2.440	508	0.459	32198	70.28
0	185	2.440	508	0.455	32189	70.26
0	190	2.440	508	0.445	32109	70.08
0	195	2.449	506	0.432	31552	69.40
0	200	2.452	506	0.410	31324	69.07

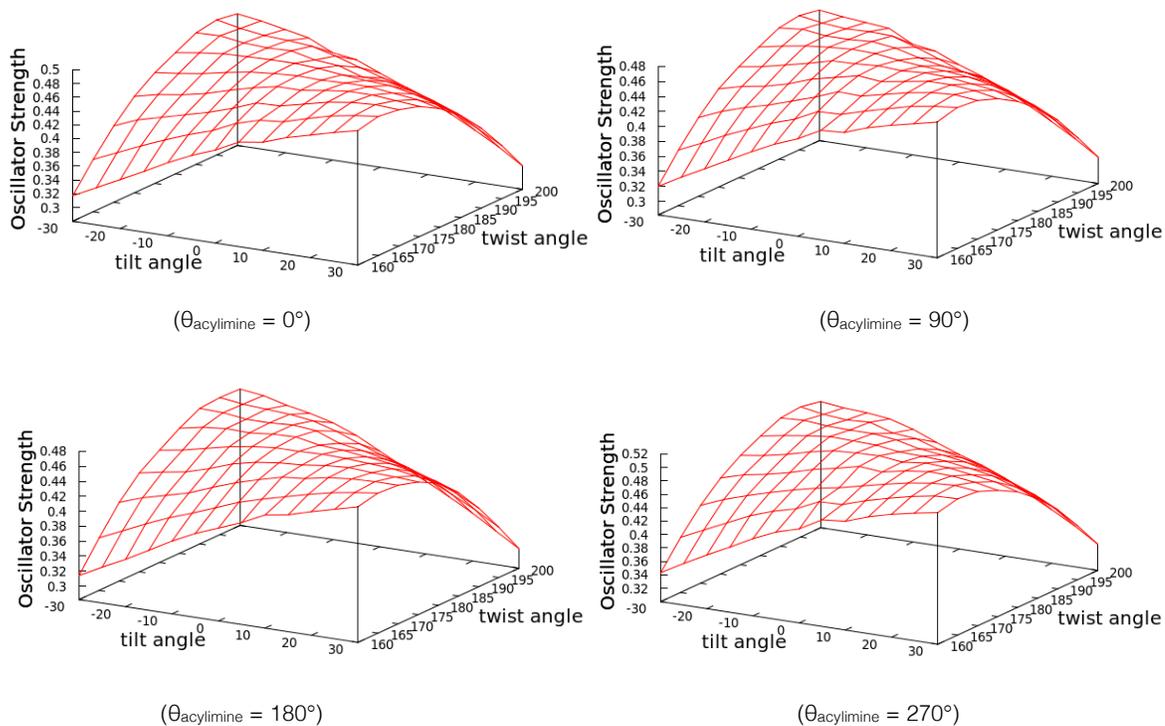
5	160	2.443	508	0.418	31810	69.60
5	165	2.441	508	0.437	31966	69.83
5	170	2.441	508	0.454	32127	70.17
5	175	2.437	509	0.458	32387	70.54
5	180	2.444	507	0.460	31755	69.54
5	185	2.445	507	0.453	31669	69.39
5	190	2.447	507	0.440	31493	69.13
5	195	2.450	506	0.421	31283	68.83
5	200	2.451	506	0.394	31083	68.48
10	160	2.435	509	0.433	32371	70.40
10	165	2.435	509	0.445	32185	69.96
10	170	2.436	509	0.457	32045	69.72
10	175	2.438	509	0.462	31847	69.41
10	180	2.439	508	0.459	31738	69.25
10	185	2.441	508	0.449	31566	68.97
10	190	2.444	507	0.433	31319	68.61
10	195	2.447	507	0.411	31087	68.27
10	200	2.447	507	0.379	30922	67.89
15	160	2.424	512	0.438	32621	70.27
15	165	2.426	511	0.453	32394	69.88
15	170	2.428	511	0.463	32108	69.42
15	175	2.431	510	0.464	31890	69.08
15	180	2.432	510	0.457	31767	68.87
15	185	2.436	509	0.445	31392	68.30
15	190	2.440	508	0.426	31124	67.93
15	195	2.442	508	0.398	30893	67.53
15	200	2.440	508	0.363	30732	67.07
20	160	2.410	514	0.447	32945	70.17
20	165	2.414	514	0.460	32527	69.48
20	170	2.417	513	0.466	32196	68.95
20	175	2.419	513	0.462	32004	68.65
20	180	2.423	512	0.455	31615	68.05
20	185	2.428	511	0.439	31247	67.54
20	190	2.431	510	0.415	30999	67.17
20	195	2.430	510	0.383	30893	66.88
20	200	2.426	511	0.346	30751	66.37
25	160	2.396	518	0.455	33012	69.49
25	165	2.400	517	0.465	32587	68.80
25	170	2.403	516	0.467	32276	68.32
25	175	2.406	515	0.461	31962	67.83
25	180	2.412	514	0.450	31500	67.17
25	185	2.416	513	0.431	31228	66.81
25	190	2.417	513	0.403	31034	66.47
25	195	2.412	514	0.367	31082	66.32
25	200	2.407	515	0.327	30989	65.83
30	160	2.380	521	0.463	32947	68.45
30	165	2.384	520	0.469	32548	67.83
30	170	2.387	519	0.467	32229	67.34
30	175	2.392	518	0.459	31798	66.72
30	180	2.397	517	0.445	31444	66.26
30	185	2.399	517	0.420	31266	66.00

	30	190	2.396	518	0.387	31360	65.99
	30	195	2.389	519	0.349	31545	66.00
	30	200	2.384	520	0.308	31376	65.38
270°	-30	160	2.617	474	0.343	17714	44.49
	-30	165	2.615	474	0.382	17618	44.17
	-30	170	2.611	475	0.416	17491	43.72
	-30	175	2.606	476	0.443	17454	43.46
	-30	180	2.600	477	0.463	17496	43.36
	-30	185	2.596	478	0.478	17534	43.33
	-30	190	2.598	477	0.491	17438	43.17
	-30	195	2.598	477	0.494	17570	43.48
	-30	200	2.596	478	0.488	17805	44.01
	-25	160	2.645	469	0.362	17198	44.12
	-25	165	2.641	470	0.398	17139	43.83
	-25	170	2.636	470	0.430	17077	43.52
	-25	175	2.629	472	0.453	17174	43.53
	-25	180	2.622	473	0.468	17335	43.69
	-25	185	2.621	473	0.484	17324	43.65
	-25	190	2.623	473	0.493	17312	43.66
	-25	195	2.620	473	0.488	17572	44.22
	-25	200	2.621	473	0.484	17598	44.34
	-20	160	2.665	465	0.380	16900	44.01
	-20	165	2.660	466	0.414	16865	43.77
	-20	170	2.654	467	0.441	16913	43.68
	-20	175	2.645	469	0.459	17091	43.85
	-20	180	2.641	470	0.475	17202	43.98
	-20	185	2.643	469	0.489	17132	43.87
	-20	190	2.641	470	0.490	17278	44.19
	-20	195	2.641	470	0.485	17373	44.42
	-20	200	2.646	469	0.482	17228	44.23
	-15	160	2.680	463	0.397	16670	43.90
	-15	165	2.675	464	0.430	16646	43.68
	-15	170	2.667	465	0.451	16789	43.78
	-15	175	2.659	466	0.467	16980	44.02
	-15	180	2.659	466	0.484	16941	43.92
	-15	185	2.660	466	0.493	16928	43.92
	-15	190	2.658	466	0.489	17078	44.25
	-15	195	2.662	466	0.487	16951	44.06
	-15	200	2.667	465	0.478	16908	44.09
	-10	160	2.690	461	0.415	16522	43.85
	-10	165	2.684	462	0.442	16582	43.78
	-10	170	2.675	464	0.460	16775	44.00
	-10	175	2.671	464	0.477	16829	44.03
	-10	180	2.673	464	0.491	16756	43.88
	-10	185	2.671	464	0.492	16853	44.10
	-10	190	2.673	464	0.489	16842	44.11
	-10	195	2.671	464	0.476	16906	44.23
	-10	200	2.680	463	0.468	16813	44.26
	-5	160	2.694	460	0.430	16539	44.00
	-5	165	2.685	462	0.451	16668	44.07

-5	170	2.678	463	0.468	16838	44.28
-5	175	2.678	463	0.484	16780	44.11
-5	180	2.680	463	0.493	16606	43.72
-5	185	2.678	463	0.491	16677	43.86
-5	190	2.681	463	0.485	16724	44.06
-5	195	2.683	462	0.473	16734	44.17
-5	200	2.687	461	0.454	16757	44.37
0	160	2.691	461	0.440	16673	44.26
0	165	2.690	461	0.467	16580	43.98
0	170	2.678	463	0.474	16891	44.42
0	175	2.684	462	0.492	16551	43.72
0	180	2.683	462	0.494	16598	43.80
0	185	2.682	462	0.486	16670	43.97
0	190	2.684	462	0.477	16621	43.92
0	195	2.689	461	0.462	16582	43.95
0	200	2.693	460	0.441	16579	44.10
5	160	2.690	461	0.459	16659	44.21
5	165	2.678	463	0.468	16910	44.47
5	170	2.678	463	0.482	16774	44.11
5	175	2.679	463	0.489	16692	43.91
5	180	2.681	462	0.490	16556	43.64
5	185	2.684	462	0.485	16477	43.53
5	190	2.687	461	0.473	16408	43.45
5	195	2.692	461	0.455	16366	43.50
5	200	2.697	460	0.430	16373	43.68
10	160	2.676	463	0.462	16911	44.42
10	165	2.674	464	0.479	16835	44.14
10	170	2.675	464	0.490	16663	43.72
10	175	2.677	463	0.496	16488	43.32
10	180	2.680	463	0.494	16350	43.06
10	185	2.682	462	0.483	16347	43.10
10	190	2.686	462	0.468	16249	42.99
10	195	2.690	461	0.445	16284	43.21
10	200	2.695	460	0.416	16315	43.44
15	160	2.665	465	0.474	17012	44.31
15	165	2.665	465	0.488	16816	43.79
15	170	2.667	465	0.497	16566	43.20
15	175	2.670	464	0.500	16343	42.72
15	180	2.672	464	0.492	16331	42.74
15	185	2.674	464	0.478	16251	42.62
15	190	2.678	463	0.459	16241	42.72
15	195	2.683	462	0.432	16294	43.00
15	200	2.685	462	0.398	16428	43.42
20	160	2.650	468	0.483	17085	43.99
20	165	2.652	468	0.496	16767	43.23
20	170	2.655	467	0.502	16471	42.57
20	175	2.657	467	0.499	16352	42.33
20	180	2.659	466	0.487	16328	42.32
20	185	2.663	466	0.472	16237	42.21
20	190	2.666	465	0.448	16322	42.54
20	195	2.670	464	0.417	16409	42.88

20	200	2.670	464	0.380	16603	43.41
25	160	2.631	471	0.490	17135	43.49
25	165	2.635	471	0.502	16716	42.55
25	170	2.637	470	0.504	16485	42.05
25	175	2.639	470	0.495	16454	42.01
25	180	2.641	470	0.482	16394	41.93
25	185	2.644	469	0.461	16432	42.13
25	190	2.648	468	0.435	16529	42.50
25	195	2.650	468	0.401	16708	43.02
25	200	2.649	468	0.361	16953	43.61
30	160	2.609	475	0.496	17147	42.78
30	165	2.611	475	0.503	16841	42.10
30	170	2.612	475	0.500	16748	41.91
30	175	2.613	474	0.488	16742	41.92
30	180	2.616	474	0.472	16744	42.00
30	185	2.619	474	0.448	16848	42.37
30	190	2.621	473	0.418	17015	42.87
30	195	2.620	473	0.380	17332	43.62
30	200	2.617	474	0.339	17599	44.20

---



**Figure C5:** Variation of OPA oscillator strengths with tilt and twist angles (see Figure 1 in the main text) for Model 21 at fixed  $\theta_{\text{acylimine}}$  of  $0^\circ$ ,  $90^\circ$ ,  $180^\circ$  and  $270^\circ$ . The OPA values are for the transition to  $S_1$  as determined at the B3LYP/6-31G+(d,p) level of theory and PCM with parameters for  $H_2O$ .