Computational Studies of UV-B Induced Signalling Pathways in Plants

Doctoral Thesis

UNIVERSITY OF GOTHENBURG

Min Wu
Department of Chemistry and Molecular Biology
Faculty of Science
Sweden
2017

Cover illustration: Min Wu

ISBN: 978-91-629-0224-7, (PRINT)
ISBN: 978-91-629-0223-0, (PDF)
Available at http://hdl.handle.net/2077/52128

Printed in Gothenburg, Sweden 2017
By Ineko AB

Science never solves *a problem* without creating *ten* more.

-George Bernard Shaw

Abstract

This thesis presents theoretical studies of UV-B induced signaling pathways in plants using molecular dynamic methods combined with (TD)-DFT calculations.

Ultraviolet B (UV-B) radiation is a component of sunlight covering wavelengths 280-300nm. Despite it being a minor component of sunlight, it has a major influence not only on formation of reactive oxygen species (ROS) leading to oxidative stress, but also on regulation of plant growth and development. Albeit low levels of ROS production are required to maintain physiological functions, high levels of ROS cause serious oxidative damage to DNA, protein and lipids. In plants, UV-B radiation has been identified to induce specific changes in gene expression resulting in a UV-B induced self-protection in plants, such as flavonoid biosynthesis, DNA repair upregulation and antioxidant activity. Expression of more than 100 genes have been identified to be regulated by the UV resistance Locus 8 (UVR8) photoreceptor which provides the initial response to UV-B stress and initialize the UV-B induced signalling pathways which also include downstream regulatory proteins such as Constitutively photomorphogenic 1 (COP1), Elongated hypocotyl5 (HY5) and Repressor of UV-B photomorphogenesis (RUP) proteins.

The UVR8 protein is a homodimer using tryptophan amino acids W285 and W233 as intrinsic chromophores to absorb UV-B radiation, followed by monomerization by the dimer. The tryptophan amino acids are located at the dimer interface and are essential for the photoreception. Residues R286 and R338 are identified to be involved in salt bridge interactions at the interface, stabilizing the dimer structure. The UVR8 monomers are able to interact with the WD40 repeat domain of the downstream protein COP1, through the UVR8 C-terminus including 27 amino acids (sequence 397-423). This interaction is necessary and essential for regulation of the signalling pathways. The UVR8 dimer is localized in the cytoplasm in plants but rapidly accumulates in the nucleus in the presence of UV-B. It is found that the nuclear localization signal (NLS) domain of COP1 is required for the nuclear addressing of UVR8. The negative regulator RUP proteins stimulated by UV-B exposure via UVR8 signalling prevent hyper-activation of the responses by constraining UVR8 action through a combination of COP1 displacement and reversion of the signalling active monomers to the dimeric form.

In order to study the mechanisms of UV-B induced signalling pathways in plants, several computational methods were used. TD-DFT calculations were performed to identify the key

tryptophan residues in response to UV-B radiation at the wavelength 300nm (Paper IV). The steered molecular dynamics (SMD) and umbrella sampling simulations of wild type and mutant systems were performed to explore the stability of the dimer and to identify the key salt bridges at the dimer interface that stabilize the dimer structure (Paper V). A new generalized AMBER force field for neutral arginine radicals was obtained using the *ab initio* HF/6-31G* method in Paper III. The mechanisms of UVR8 monomerization induced by UV-B radiation was studied in Paper VI using (S)MD with the new generalized AMBER force field for neutral arginine radicals, and (TD)-DFT calculations. The inverted free energy landscape of the intrinsically disordered C-terminus was obtained in Paper VII. In Paper VIII, the interaction between the UVR8 C-terminus and the COP1 protein was studied.

List of Publication

Paper I

M.Wu, Q. Xu, Å.Strid, J.M.Martell, L.A.Eriksson, Theoretical study of pyridoxine (Vitamin B6) *J. Phys. Chem. A,* **2011**, 115(46), pp 13556-13563.

Paper II

G.Czégény, **M.Wu**, A.Dér, L.A.Eriksson, Å.Strid, É.Hideg, Hydrogen peroxide contributes to the ultraviolet-B (280-315 nm) induced oxidative stress of plant leaves through multiple pathways *FEBS Lett.*, **2014**, 588, pp 2255-2261.

Paper III

M.Wu, Å.Strid, L.A.Eriksson, Development of non-standard arginine residues parameters for use with the Amber force field *Chem. Phys. Lett.*, **2013**, 584, pp 188-194.

Paper IV

M.Wu, E. Grahn, Å.Strid, L.A.Eriksson, Computational evidence for the role of *Arabidopsis thaliana* UVR as UV-B photoreceptor and identification of its chromophore amino acids J*. Chem. Inf. Model*, **2011**, 51(6), pp 1287-1295.

Paper V

M.Wu, Å.Strid, L.A.Eriksson, Interactions and stabilities of the UV RESISTANCE LOCUS8 (UVR8) protein dimer and its key mutants *J. Chem. Inf. Model*, **2013**, 53, pp 1736-1746.

Paper VI

M.Wu, Å.Strid, L.A.Eriksson, The photochemical reaction mechanism of UV-B induced monomerization of UVR8 dimers as the first signaling event in UV-B-regulated gene expression in plants *J. Phys. Chem. B*, **2014**, 118, pp 951-965.

Paper VII

M.Wu, D. Farkas, Å. Strid, L.A.Eriksson, The intrinsically disordered domain of UVR8. Its inverted free energy landscape and predicted interactions with COP1 *manuscript.*

Paper VIII

M.Wu, Å.Strid, L.A.Eriksson, Theoretical prediction of the protein–protein interaction between Arabidopsis thaliana COP1 and UVR8 Theor. Chem. Acc. **2013**, 132(7), pp 1371-1377.

Related Publication Not Included

Paper IX

M.Wu, L.A.Eriksson, Absorption spectra of riboflavin—a difficult case for computational chemistry, *J. Phys. Chem. A*, **2010**, 114 (37), pp 10234–10242

Paper X

W.Q.Li, W.Shi, Z.H.Wu, J.M.Wang, **M.Wu**, W.H.Zhu, Unsymmetrical donor–acceptor–donor–acceptor type indoline based organic semiconductors with benzothiadiazole cores for solution-processed bulk heterojunction solar cells *Green Energy & Environment*, **2016**, pp 1-8

Paper XI

J.J.Chen, I.Masakazu, **M.Wu**, H.Q.Yu, K.K.Niyogi, G.R.Fleming, Protonation of major light-harvesting complex of photosystem II facilitates nonphotochemical quenching *submitted.*

Contribution report

For Paper I, III, IV, V, VI, VII, VIII and IX, I was involved in planning the project and updating the progress of the project with Prof. Åke Strid and Prof. Leif A. Eriksson. I was responsible for designing performing calculations and analyzing data and writing the papers.

For Paper II and X, I was involved in planning the project with collaborators. I was responsible for performing the calculations, analyzing data and writing parts of the papers.

For Paper XI, I was involved in part of experiments. I was responsible for sample preparations (protein purification, proteoliposome preparation) and spectroscopy Time-Correlated Single Photon Counting (TCSPC) measurements.

List of Abbreviations

UV - Ultraviolet
ROS – reactive oxygen species
SOD - superoxide dismutase
PDB – Protein data bank
FMN - Flavin mononucleotide
MPO - myeloperoxidase
QM - Quinone methide
KS - Kohn-Shan
TG - Transient grating
CD - Circular dichroism
MO - Molecule orbitals
HY5 - Enlongated hypocotyl 5
TRIB1 - Tribbles homologs 1
CV - Collective variable
AA – Amino acid
DFT - Density function theory
FF - Force field
HF - Hartree-Fock
MD - Molecular dynamic
PES - potential energy surface
IEF-PCM - Integral equation formalism polarizable continuum model
FAD - Flavin adenine dinucleotide
SPA1 – Suppressor of phytA-105
TD-DFT - Time-dependent DFT
MM - Molecular mechanics
GAFF - Generalized AMBER force field
DNA - Deoxyribonucleic acid
LDA - Local density approximation
PME - Particle mesh Ewarld
SMD - Steered molecular dynamic
WHAM - Weighted histogram analysis method
RMSD - Root means square derivation
GGA - Gradient approximation
ROS - Reactive oxygen species
UVR8 - UV Resistance Locus 8
ECD - Electron-capture dissociation
ETD - Electron-transfer dissociation
PMF - Potential mean force
COP1 - Constitutively photomorphogentic 1
BO- Born-Oppenheimer

Contents

Chapter 1. UV-B light

1.1 light

Sunlight also called the sunshine, is ordinarily divided into three main elements: first, visible light with wavelengths between 400 to 800nm, second ultraviolet radiation, with wavelengths shorter than 400nm and third infrared radiation with wavelengths longer than 800nm. The visible light can be divided into different colors basing on their respective wavelengths: red covers 620-800nm, orange 590-620nm, yellow 570-590nm, green 495-570nm, blue 450-495nm and purple 400-450. UV radiation has been divided into three subregions which together cover the solar wavelengths <400 nm, where UV-A corresponds to 315-400 nm, UV-B covers the range 280-315 nm and UV-C <280 nm. The amount of sunlight can reach the earth depending on the levels of the daytime cloud cover. About 29% of the solar energy that reaches the top of the atmosphere is reflected back to the universe because of the clouds, shown in figure 1. The rest of the solar energy is absorbed by the atmosphere, oceans, lands and living things[1]. The atmosphere is transparent to most wavelengths, but 23% of the solar spectrum are absorbed by certain "greenhouse gasses" in the atmosphere including water vapor, carbon dioxide, ozone, methane and other gasses[2]. Shorter wavelengths (UV and blue light) are diffused in the air and make the sky blue. Longer wavelengths are less diffused and make sunsets and sunrise red. The light reaches earth containing 10% UV light, 40% visible light and 50% near infrared light at the top of Earth's atmosphere. Up to sea level, approximately 6% of the radiation is UV, 50 % is visible radiation and about 40% is infrared radiation.

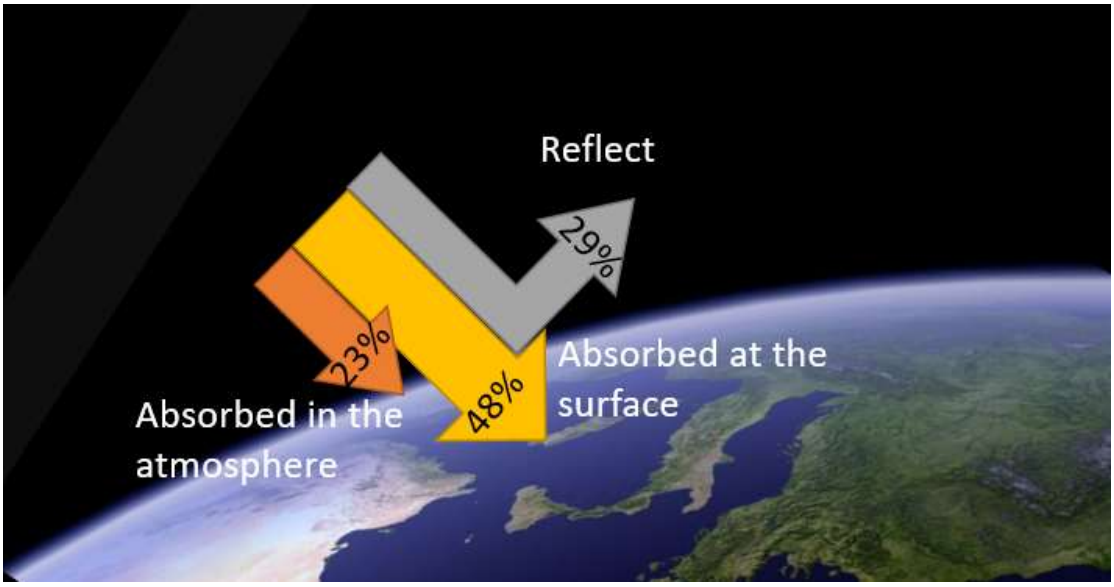


Figure 1. The Earth's energy balance diagram

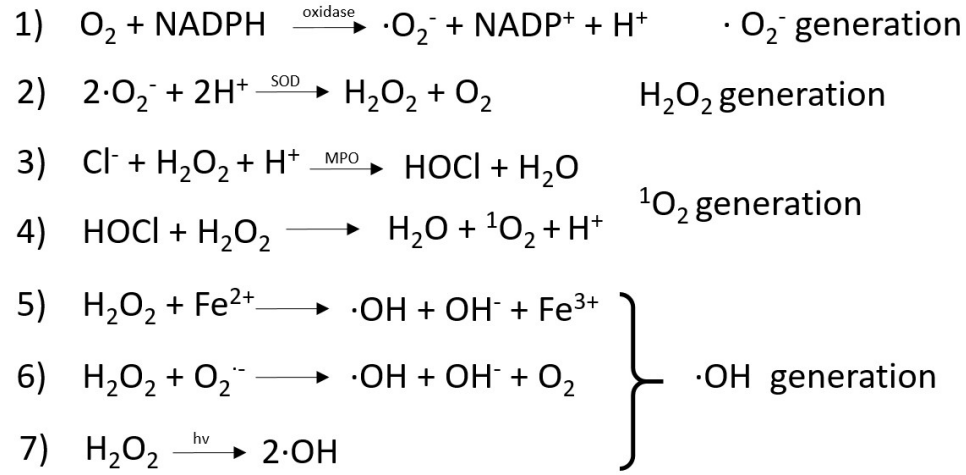
1.2 Evaluating consequences of UV-B exposure

UV-B is 10 to 100 times less abundant than UV-A, whereas UV-C, absorbed in the atmosphere, is practically absent. Due to differences in the scattering and absorption, the ratio of UV-B to UV-A depends on several factors such as latitude, zenith angle, cloud cover and thickness of the ozone layer[3]. Small quantities of UV-B radiation acting as a catalyst in the generation vitamin D are essential to human health[4]. In the late 1980s, the ozone layer depletion triggered concerns about the potentially harmful effects of increased UV-B radiation[5]. Many studies show that large amounts of UV-B induce increased levels of reactive oxygen species (ROS) and cause damage to DNAs, proteins and membranes leading to diseases such as skin cancers, cataracts and immunosuppression[6]. However, the natural level of UV-B light is not harmful to plant growth. UV-B radiation has been identified to induce specific changes in gene expressions, leading to accumulation increases of UV-screening pigments[7] and phytochemical content changes resulting in a UV-B tolerance increase[8], leading to the optimization of plants' growth and developments in response to their ambient environment.

Chapter 2. UV-B induced reactive oxygen species (ROS)

2.1 ROS generation

ROS contain chemical reactive species: hydrogen peroxides(H_2O_2), superoxide anion radical ($\cdot O_2^-$), hydroxyl radical($\cdot OH$) and singlet oxygen($^1O_2$). Up to UV-B exposure, the increased levels of ROS may be formed as a result of disruption of metabolic activities or due to an increased activity of membrane-localized NADPH-oxidase[9]. The generations of ROS are listed in Scheme 1[10]. The superoxide anion radical $\cdot O_2^-$, are initially rapid generated with the uptake of oxygen and activation of NADPH oxidase. Then the superoxide anion radical $\cdot O_2^-$ is then rapidly converted to H_2O_2 by superoxide dismutase (SOD). The singlet oxygen $^1O_2$ is produced by the H_2O_2, which is firstly converted to hypochlorous acid (HOCl) by myeloperoxidase (MPO) in neutrophils and then followed by the reaction of HOCl with H_2O_2. The hydroxyl radical can be generated from $\cdot O_2^-$, $Fe^{2+}$and H_2O_2[10]. From our group, we find that the H_2O_2 production is like to directly converse to $\cdot OH$ radical in response to UV-B irradiation in plants[11]. The studies are in Paper II. Low levels of ROS production are required to maintain physiological functions, including proliferations, host defenses, signal transductions and gene expressions. However, in a high UV-B stress, the excessive ROS production results in the oxidative stress. The excessiveness of ROS causes oxidative damage to the DNAs, proteins and lipids. For example, ROS can react with the nucleic acids attaching the nitrogenous bases and the sugar phosphate backbones, resulting in evoking single or double stranded DNA breaks. A number of accumulated damages in DNAs might lead to increase of the living cells' ages[12].



Scheme 1. The generation of ROS containing: hydrogen peroxides(H_2O_2) , superoxide anion ($O_2^{\cdot -}$), hydroxyl radical (·OH) and singlet oxygen ($^1O_2$).

2.2 ROS scavenging

Antioxidants capable of scavenging ROS have attracted much attention in photochemistry. Two kinds of antioxidants: enzymatic and non-enzymatic antioxidants can be used to define endogenous compounds in cells[6]. The non-enzymatic antioxidants are divided into metabolic antioxidants which are produced by metabolism in the body such as coenzyme Q10[6] and nutrient antioxidants belonging to exogenous antioxidants which cannot be produced in the body and must be provided through foods or supplements such as Vitamin C[6], E[6] and B_6, shown in figure 2. Vitamin C is a water-soluble vitamin and it works together with Vitamin E to quench free radicals. Vitamin B_6, known as pyridoxine, is widely distributed in plants such as bananas and potatoes. It plays a crucial role in protecting cells against ROS because of its antioxidant capacities[13]. The antioxidant activity of Vitamin B_6 has been shown to exceed that of Vitamins C and E[6]. The possible reactions of Vitamin B_6 induced by UV-radiation were studied in Paper I. It is concluded that neither aromatic ring-opening nor photodegradation reactions leading to loss of ring-bound functional groups (hydroxymethyl groups, a methyl group and a hydroxyl group) are likely to occur upon exposure to UV-radiation. However, the generation of the *o*-QM compound is selectively produced by Vitamin B_6 pyridoxine after UV-irradiation.

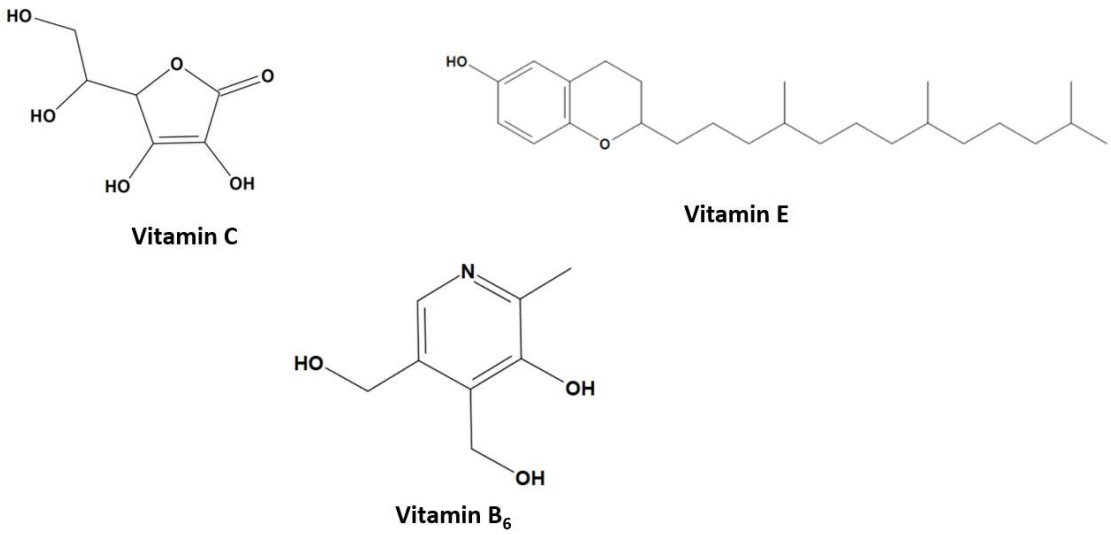


Figure 2. The molecule structures of Vitamin C, E and B_6 nutrient antioxidants

Chapter 3. Why plants can survive from UV-B stress?

The oxidative stress has been flagged as a key factor in a UV-B stress[5]. However, plants can survive from the UV-B radiation and the mechanisms of UV-B induced signaling pathways in plants have been studied several years. UV-B irradiation has no significant effects on photosynthesis. On the other hand, it has been found that the reactive oxygen species (ROS), DNA damages, membrane degradation products play a critical role in mediating UV-B protection in plants. Therefore, it has become increasingly clear that under natural UV-B exposure condition, UV-B is not harmful to plant growth. UV-B radiation has been identified to induce specific changes in gene expressions, leading to accumulation increases of UV-screening pigments[7] and phytochemical content changes resulting in a UV-B tolerance increase[8]. Those events are known as photomorphogenesis processes in plants. With photomorphogenesis processes, plants have been shown to have a high degree of developmental ability to optimize their growth and reproduction in response to their ambient environment, such as light, temperature, humidity, and salinity[14]. UV-B light is one of the critical light environment and is utilized as a source of information for plants exposure environment that influences the *A. thaliana* growth and development, shown in figure 3.

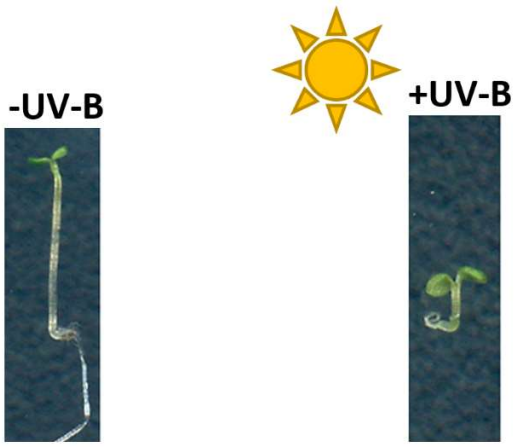


Figure 3. Phenotypes of 4-day-old wild-type *A. thaliana* seedlings grown under −UV-B and +UV-B[15].

In plants, Expression of more than 100 genes have been identified to be regulated by the UV resistance Locus 8 (UVR8) photoreceptor which provides the initial response to UV-B stress[16] and initialize the UV-B induced signalling pathways including the downstream regulatory proteins: Constitutively photomorphogenic 1 (COP1)[17], Elongated hypocotyl5 (HY5)[18] and Repressor

of UV-B photomorphogenesis (RUP) proteins[19]. The overview of possible UV-B induced signalling pathways is shown in figure 4. The detailed description of each process is described in the followed paragraphs in Chapter 3.

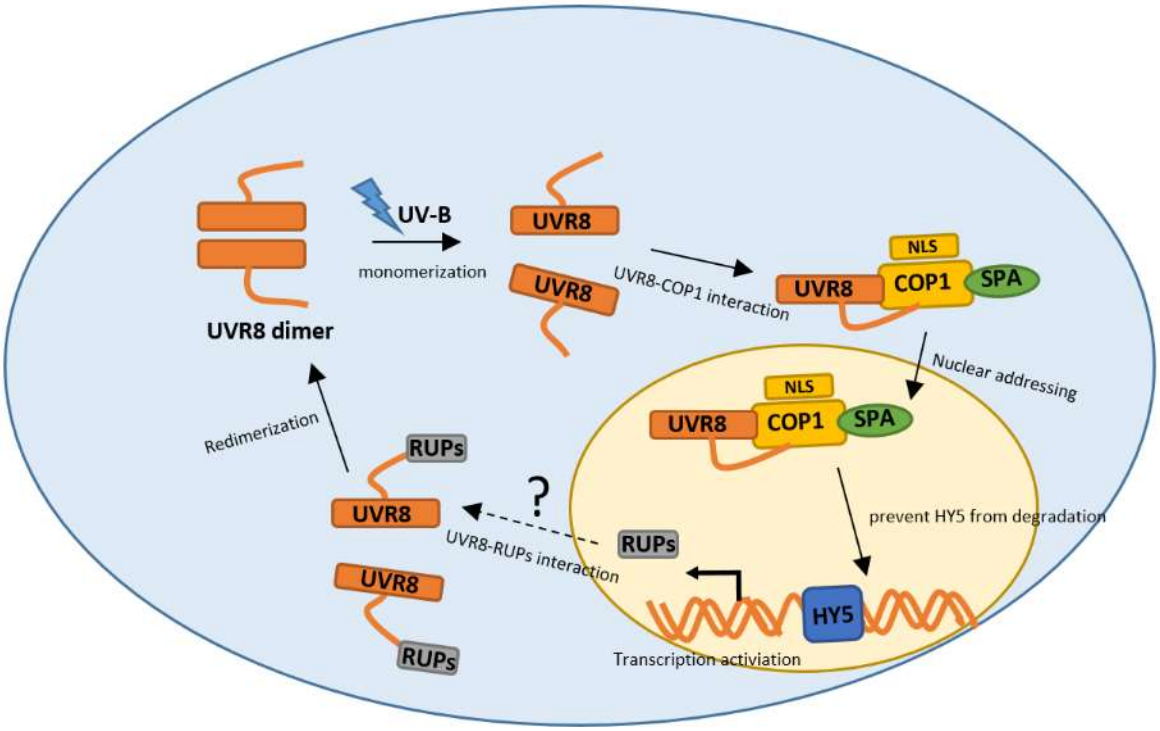


Figure 4. Possible signalling pathways regulated by UV-B light in plants. The unknown processes are with question marks.

3.1 UV resistance locus 8 (UVR8) UV-B photoreceptor

UVR8 protein has been identified to specifically mediate photomorphogenic UV-B responses by acting as a UV-B photoreceptor[20]. UVR8 is involved in orchestrating the expressions of more than 100 genes and is suggested to have appeared early in the evolution of plants to survive exposure to the UV-B fraction of sunlight[16]. UVR8 exists as a homodimeric protein located in the cellular cytosol and dissociates into a signaling-active monomer upon UV-B exposure[21]. Each subunit consists of an N-terminal domain containing a seven-bladed β-propeller repeat domain and C-terminal domain containing 55 amino acids which are missing from the crystal structure. The dimer is held together by electrostatic interactions between charged amino acids

across the dimer interface[22, 23], shown in figure 5. An action spectrum of UVR8 function indicates that UVR8 is most effective at 280nm with the critical actions at 290nm and 300nm[18]. Since 280nm rarely reaches the surface of the earth, the longer wavelengths (290nm and 300nm) are the most physiologically relevant[24]. UVR8 is a unique photoreceptor and utilizes tryptophan amino acids absorbing UV-B instead of attached chromophore for the absorption of specific wavelengths[23]. The C-terminal domain is necessary and efficient for the downstream signaling [25]. The UVR8 redimerization is apparent within approximately 30-min post UV-B exposure and it is complete around 2-h in vivo and more than 24-h in vitro indicating a reversible process[29].

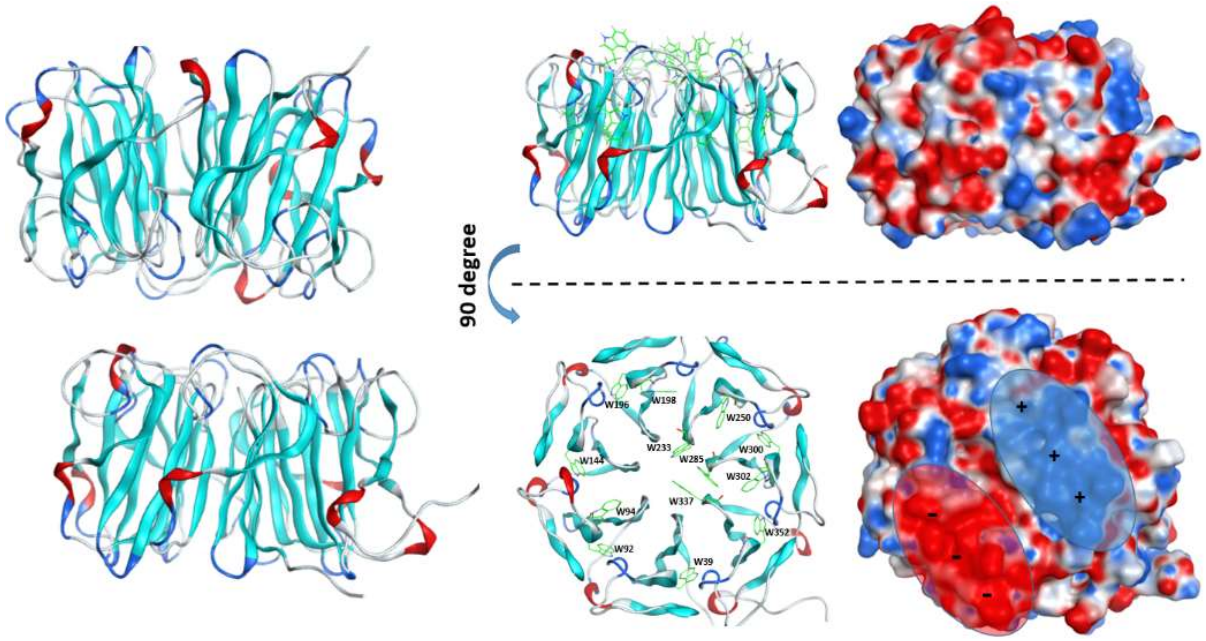


Figure 5. The 3D crystal structure of UVR8 dimer, UVR8 monomer with all tryptophan residues presented in crystal structure highlighted and the electrostatic surfaces of UVR8 monomer with the positively charged surfaces are presented in blue and negatively charged surfaces are presented in red.

3.1.1 Tryptophan residues in UVR8 protein.

A. thaliana UVR8 has 14 tryptophan residues that are highly conserved among plant species. The positions of 13 tryptophans presented in the crystal structure are shown in figure 5, whereas residue W400 is located in the missing C-terminus region[22, 23]. Residues W39, W196, W300 and W92 are identified to take part in the maintenance of the seven-bladed β-propeller repeat domain by hydrogen bonds and hydrophobic interactions and residues W94, W198, W233, W250, W296,

W302 and W337 are positionally located in the dimeric interface[26]. Both dimer and monomer structures are observed in the absence of UV-B after W233A and W337A mutations. W285A and W39A mutants, on the other hand, are constitutive monomers and W233F and W285F mutants are constitutive dimers both in the absence and in the presence of UV-B [22, 23, 26, 27]. It is shown that residue W285 plays a critical role both in dimerization of UVR8 monomers and in UV-B perception, and residue W233 plays an essential role in UV-B perception[22, 23]. The key tryptophan residues W285 and W233 which regulate the monomerization of UVR8 dimer are discovered theoretically in Paper IV.

3.1.2 Salt bridges across the UVR8 dimer interface

The salt bridges interactions exist in the UVR8 homodimer interface. The electrostatic interactions of salt bridges hold together two UVR8 monomers[22, 23]. The salt bridges involving arginines R146, R286, R338 and R354 interacting with specific negatively charged residues aspartate and glutamate residues on the opposing monomer to form an antisymmetric homodimer structure. Through the single mutant analysis experiment using size exclusion chromatography, gel filtration and theoretical umbrella sampling simulation, the charged residues R286 and R338 interacting with corresponding negatively charged residues shown to be essential for maintaining the electrostatic interactions between the UVR8 monomer units in the homodimer complex[22, 23]. The key salt bridge interactions based on residues R286 and R338 are discussed in Paper V.

3.1.3 The proposed mechanisms for UVR8 monomerization

The mechanism by which UV-B induce dissociation of the UVR8 dimer is still unclear. Three different mechanisms have been suggested using theoretical approaches.

1) Photo-induced electron transfer, where an excited electron is sequentially transferred from W233 to W285 and then further on to R338. The electron transfer is accompanied by proton transfer from W233 to nearby residue D129 resulting in the neutralization and disruption of R338 and related salt bridges[28]; This model is studied in Paper VI.
2) Another possibility is that UV-B induces a large dipole moment by the formation of radical ion pair [$W233^{-}$ $W285^{+}$] in each UVR8 unit which may cause the disruptions of critical salt bridges in the dimer interface[29].
3) The UV-B induced excitation of W233 and W285 which may initiate an electron and proton transfer leading to the neutralization of the R286-D107 and/or R338-D44 salt bridges[30].

As to experimental studies, hypotheses of mechanisms were gain:

1) The dissociation of critical salt bridges could result from the diminution of the cation-π interactions between triad Trps and adjacent Args by the photoreception[22].
2) Time-resolved absorption and fluorescence spectroscopy have shown that W285 plays a critical role in primary quenching dynamics, most likely in the process of exciton evolution to a charge-separated state to induce the disruption of salt bridges for initiating dimer dissociation[31].
3) The kinetic analyses through the transient grating (TG) signal and Circular dichroism (CD) spectral measurements indicate a photo-induced conformation change of UVR8 before dissociation[32].
4) Dynamic crystallographic observations suggest that the inter-subunit interactions at the dimer interface are perturbed due to ejecting epicenter water by photo-concerted motions from indole rings and backbones of W285 and W233[33].

3.2 E3 ubiquitin-protein ligase Constitutively Photomorphogentic 1 (COP1)

In the dark, seedling photomorphogenesis is repressed via COP1[34]. COP1 is localized in the nucleus in the dark and move to the cytoplasm in the light[35]. In the dark, the COP1 protein forms a complex in the nucleus with suppressor of phytA-105 (SPA1) proteins[34]. The COP1 protein contains three distinct domains: an N-terminal Zn^{2+} binding RING finger domain, a coil-coil domain and a C-terminal WD40 repeat domain[36, 37] and SPA1 protein contains an N-terminal kinase-like domain, a coiled-coil domain and a C-terminal WD repeated domain. COP1 and SPA1 proteins interact with each other through their coiled-coil domain[38]. Recently, the crystal structures of the COP1 WD40 domain are obtained from two different organisms, *A. thaliana* (PDBID: 5IGO) [39] and human (PDBID: 5C5K) [39], shown in figure 6. Indeed, the COP1 WD40 domain from *A. thaliana* is very much similar to the human version. The WD40 repeated domains forming a seven repeat beta-propeller motif are structurally conserved, even if there is a large difference in amino acid sequences [41]. In the structures, the WD40 domain is in complex with C-terminal peptides derived from Tribbles homologs 1 (TRIB1)[40]. Both in plants and mammals, COP1 is involved in developmental processes. In mammals, COP1 is essential and has been implicated as both being a tumor suppressor and a cancer promoter[40]. In plants, COP1 is known as a repressor of photomorphogenesis and acts as an E3 ubiquitin ligase in darkness but under the

influence of UV-B and UVR8, serves as a promoter of UV-B-specific responses[42]. The WD40 repeated domain of COP1 protein directly interacts with different proteins such as the transcription factors HY5, a bZIP transcription factor in dark, JUN[42] and ETS[43], photoreceptor Cry1[44], salt tolerance-related protein STO and its homologue STH[45] and TRIB[39] proteins. The consensus Val-Pro-charged residue motif which binds to COP1 WD repeated domain has been found in all these proteins[40, 46].

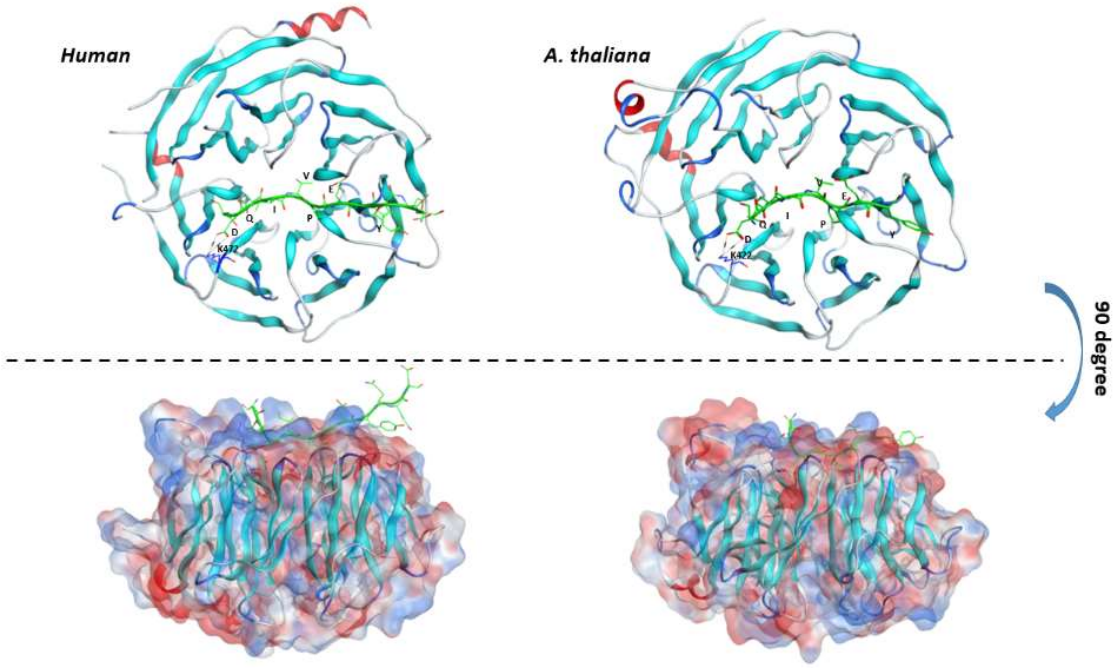


Figure 6. The crystal 3D structures of COP1 WD40 repeated domain in human and *A. thaliana.* C-terminal peptides of TRIB1 binding to COP1 are shown in green. The electrostatic surfaces of COP1 are shown with negatively charged surfaces in red and positively charged surfaces in blue.

3.3 The interactions between COP1 and UVR8 photoreceptor

COP1 is required for UV-B–induced nuclear accumulation of the UVR8 photoreceptor and downstream signaling transduction. The photo-induced UVR8 monomer directly interacts with COP1/SPA1 complex through two domains of UVR8 monomer: a region of 27 amino acids towards the C-terminus of UVR8 (termed C27; AA 397-423) including residues V410 and P411and a region of seven-bladed β-propeller repeated domain[47]. UVR8 has been shown to only bind to COP1 but no direct contact between SPA1 and UVR8[48]. Furthermore, the COP1 protein positively regulates the UV-B-specific response independent of the SPA1 protein[17]. The

interaction between C-terminus peptide of UVR8 and COP1 is UV-B independent but the interaction involving β-propeller domain is UV-B dependent-manner[47]. Whereas, the activity of COP1 regulated by interaction through UVR8 C27 domain is sufficient to mimic UV-B signalling[25]. The present UVR8 crystal structures do not provide any information on the structure and position of 44 AA in C-terminus (termed C44; AA 397-440), indicating that it is flexible. There is also evidence that UV-B exposure causes a conformational change of UVR8 that exposes the C-terminus, leading to its interaction with COP1 protein[22, 23]. Interestingly, in the crystal structure, the C-terminuses of two UVR8 protein in a dimer are pointing towards each other and being in proximity of approximately 15Å, likely suggesting the presence of the interactions between two C-terminus in UVR8 dimer. However, the information on the relationship between dimer dissociation and exposure of C-terminus is still unclear. The interaction of COP1-UVR8^{C27} was studied in Paper VIII and the intrinsically disorder C-terminus of UVR8 was studied in Paper VII.

3.4 Repressor of UV-B photomorphogenesis (RUP) proteins

RUP1 and RUP2 transcripts are rapidly and transiently induced by UV-B in a COP1, UVR8 and HY5 dependent manner[19]. UV-B produced RUP proteins are important repressors of UV-B induced photomorphogenesis and UV-B acclimation[49]. The RUP1 (385 AA) and RUP2 (368 AA) proteins are highly homologous, with approximately 63% identity[19]. Both RUP proteins apparently only consist of WD40 repeated domain, localized to the nucleus and the cytoplasm independent of light condition, and directly interacting with both UVR8 monomer under UV-B exposure and UVR8 dimer post UV-B exposure[19]. However, RUPs are able to access to UVR8 dimer is not because its smaller size compared to COP1. The homology model of RUP1 is shown in figure 7. The C-terminal 27 amino acid of UVR8 (aa 397-423) is determined to be necessary and sufficient for its binding to RUPs through a UVR8 V410-P411 motif[47]. Restoring UVR8 homodimer recovery is observed in present of RUPs[49]. It is not clear that the RUPs-induced UVR8 redimerization is solely due to inactivation of activated residues in UVR8 monomer or an increase of UV-B-induced RUPs limiting the binding between COP1 and UVR8. However, it is proposed that RUPs mediate UVR8 redimerization independently of COP1[49].

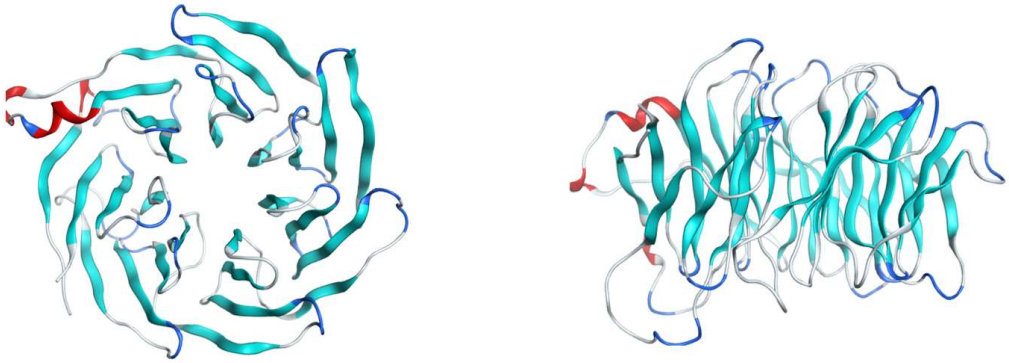


Figure 7. Homology model of RUP1 protein using *A. thaliana* COP1 WD repeated domain as the template (PDB ID: 5IGO).

Chapter 4. Computational methods used in my researches

Theoretical chemistry is a branch of chemistry using mathematical methods combined with fundamental laws of physics to study chemical relevance processes. Computational methods are able to produce approximate results and followed by a refinement to any desired degree of accuracy. The questions in computational chemistry are for instance about molecular geometries, energies of molecules and transition states, chemical reactivities, the interactions of substrates with enzymes and the physical properties of substances. The theoretical methods employed by the computational chemistry are separated into classical mechanics and quantum mechanics. The time-dependent Schrödinger equation is a cornerstone in quantum chemistry, written as:

$$\hat{H}\psi(R,r) = E\psi(R,r) \tag{4.1}$$

where the wave function $\psi(R,r)$ describes a stationary state with energy E, and r and R represent all spatial coordinates of the electrons and nuclei, respectively. The Hamiltonian operator $\hat{H}$ is in this framework given by:

$$\hat{H} = \hat{T}_e + \hat{T}_n + \hat{V}_{ee} + \hat{V}_{nn} + \hat{V}_{en} \tag{4.2}$$

Where $\hat{T}_e$ and $\hat{T}_n$ are the kinetic energy operators of the electrons and nuclei, and $\hat{V}_{ee}$, $\hat{V}_{nn}$ and $\hat{V}_{en}$ are the electron-electron, nucleus-nucleus and electron-nucleus interaction operators respectively. An essential part of solving the Schrödinger equation is the Born-Oppenheimer (BO) approximation. The approximation is based on the assumption that the separation between atomic nuclei and electrons motions in a molecule, resulting in the neglecting of the couplings between the nuclei and electronic motions. The BO approximation is central in theoretical chemistry. The energy calculations and the wave functions of a molecule are simplified by BO approximation. The significant computational efforts in the theoretical chemistry are to solve the electron Schrödinger equations for a set of nuclear coordinates. Once the electronic Schrödinger equation has been solved, the potential energy surface (PES) is known. The motion of the nuclei on the PES can be calculated either using classical mechanics or quantum mechanics [50]. The theoretical methods used in my researches are summarized in this chapter.

4.1 Hartree-Fock Theory

The comprehensive computational methods are required to solve the dynamic questions of a many-electron system. In order to simplify the calculations, the independent-particle models are employed, where the motion of one electron is considered to be independent of the dynamics of all other electrons. The Hartree-Fock (HF) method uses an independent-partial model for describing

electron-electron interactions and is also the starting point for more advanced electronic structure methods. In the HF model, an orbital is used to describe each electron and the total wave function is based on a product of orbitals. The antisymmetric N-electron wave function can be approximated using a single Slater determinant compose of orthonormal spin orbitals. The best set of orbitals is determined by the variational principle, for example, the shape of a given molecular orbitals describes the probability of finding an electron where all the nuclei and average repulsion to all the other electrons are included. The Hartree-Fork (HF) method thus generates solutions to the Schrödinger equation where the real electron-electron interaction is replaced by an average interaction. The HF wave function is not accuracy enough since it typically accounts for ~99% of the total electronic energy and the remaining 1% is usually very crucial for describing chemical phenomena. In a given basis set, a standard way of defining the correlation energy E_{corr} is:

$$E_{corr} = E_0 - E_{HF} \tag{4.3}$$

where E_{HF} is the energy calculated using HF method and E_0 is the lowest possible energy. Physically, the correlation energy corresponds to the motion of electrons being correlated. The electron correlation contains dynamic and static parts. The dynamic part is associated with the electrons occupying the same spatial orbitals. The static part is associated with electrons which occupy different spatial orbitals, leading to avoiding each other. The configuration interaction, coupled cluster and many-body perturbation theory methods are employed to develop the dynamical correlation and the complete active space self-consistent field methods are employed to treat static correlation effects in advanced wave function (*ab initio*) methods. Generally, the standard implementations correlated wave function methods are limited to the large molecule systems (> 100 atoms). An alternative way of treating the electron correlation problem is given by density functional theory (DFT).

4.2 (Time-dependent) Density function theory (TD-DFT/DFT)

The DFT method is the most successful quantum mechanical modelling method used in theoretical chemistry in the treatment of many-body problems in atomic, molecular, solid state and nuclear physics[51]. The goal of DFT methods is to design functionals connection the electron density with the energy. The basis for DFT considers single-electron functions and attempts to calculate the total electronic energy and the overall electronic density distribution, was developed by Hohenberg

and Kohn[52]. The ground state electronic energy is determined entirely by the electron density ρ(r) which is related to many-electron wave function. The electron density ρ(r) is defined as:

$$\rho(r) = N \sum_{s_1} \dots \sum_{s_N} \int dr_2 \dots \int r_N |\psi(r_1, s_1, r_2, s \dots r_N, s_N)|^2 \quad (4.4)$$

$$\int \rho(r) dr = N \quad (4.5)$$

where $r_1, r_2 \dots r_N$ are the spatial variables of electrons; $s_1, s_2 \dots s_N$ are the spin variables of electrons and N is the number of electrons. Compared to HF method, DFT method offers a route for carrying out correlated calculations at a relatively small computational cost.

The modern DFT method introducing orbitals is based on HF theory, as known Kohn-Shan (KS) theory[53]. The KS model is analogous to the HF model since they share the same formulas for the kinetic, nuclear-electron and Coulomb energies. The key advantage of the KS approach is that the complicated the kinetic energy function can be divided into two parts one called exact term and a correction term. The universal functional $E[\rho]$, in the KS scheme, is thus defined as:

$$E[\rho] = T_S[\rho] + J[\rho] + E_{xc}[\rho] \quad (4.6)$$

where $T_S[\rho]$ is the exact kinetic energy of the system containing the non-interacting electrons moving in an effective potential and $J[\rho]$ is the classical Hartree (Coulomb) repulsion term. The third $E_{xc}[\rho]$ term is the exchange-correlation functional.

The self-consistent approach is taken to solve the KS equations. A set of initial orbitals can be obtained by an initial guess of the density and then used into the second iteration until convergence is achieved. The exchange-correlation functional is the key to the success of the density function approach. The classification of exchange-correlation functionals are divided into local density approximation (LDA) based on the uniform electron gas model, generalized gradient approximation (GGA) which adds energy terms depending on the gradient of the electron density, meta-GGA which introduces higher order derivatives into the GGA functional such as the Laplacian of the electron density, and hybrid-GGA methods which introduces a fraction of HF exchange into the GGA functional[54]. There is no systematic way to improve the accuracy of the results in DFT methods, which means that the performance of exchange-correlation functional to new chemical problems has to be carefully assessed through sets of benchmark calculations through the related known chemical properties.

Time-dependent density function theory (TD-DFT) is the extension of DFT was developed by E.Runge and E.K.U.Gross to solve the arbitrary time-dependent systems such as excitation energies and general time-dependent phenomena[55]. The time-dependent density $\rho(r, t)$ determines the

external potential by up to an additive function of time. A single-component system is employed in the Hamiltonian in the presence of a time-dependent scalar field. TDDFT is able to describe non-linear phenomena like high-harmonic energy, or multi-photo ionization. The accuracy of TDDFT excitation energies depends on both occupied and virtual KS orbitals and energies.

4.3 Basis sets

Basis sets, also called basis functions, is the other variable in *ab initio* and DFT/TDDFT methods besides the correlation functionals. The basis sets compose of atomic functions to represent the molecular orbitals (MO). An unknown MO can be thought of as a function expended by an infinite number of known functions. However, in the calculations, only a finite number of basis sets can be used so that the larger the basis sets are used, the better the representation is provided. However, on the other hand, the larger the basis sets are used, the more computer time is consumed. Therefore, to choose the basis set used for a particular problem is usually based on comparing the performance against known experimental data. The selection of a suitable computational model is the initial step guided by the experimental data, rather than directly starting the computational procedure.

4.4 Force field methods

The force field (FF) methods, as also named molecular mechanics (MM) methods, is an approximation to the potential energy surface given by the electronic energy for a given nuclear configuration. In the force field methods, the bonding information is provided explicitly, rather than being the result of solving the electronic Schrödinger equation so that the quantum aspects of the nuclear motion are neglected, but the dynamics of the atoms is treated by classical mechanics, such as Newton's second law. The force field energy is written as the sum of the energies containing: bond stretching E_{bonds}, angle bending E_{angles}, torsional energy $E_{dihedrals}$, non-bonded atom-atom interactions such as Van der Waals interaction E_{vdw} and electrostatic interaction E_{el}, and the coupling between the first three terms E_{cross}. There are three different classes in FF: 1) Class I used to treat large systems, such as proteins or deoxyribonucleic acid (DNA), based only on harmonic functions for bond stretching and angle bend, no cross terms are included and the Lennard-Jones potential is used for van der Waals energy; 2) Class II reproducing the small- to medium-size molecules to high degree of accuracy including a number of cross terms and cubic or quartic expansions of bond stretch and angle bend; and 3) Class III for hyperconjugation modelling allowing parameters to depend on neighboring atom types and

including electronic polarization effects. Force fields used for my Ph.D. projects are those designed for treating macromolecules such as AMBER[56]. In Paper III, the parameters of the arginyl radical and hydrogenated arginine side chains are presented using generalized AMBER force field (GAFF) approach. The equation to calculate the generalized AMBER force field energy is given:

$$E(kcal/mol) = \sum_{Bonds} \frac{1}{2} K_r (r - r_0)^2 + \sum_{Angles} \frac{1}{2} K_\theta (\theta - \theta_0)^2 + \sum_{Dihedrals} K_\phi (1 + \cos(n\phi - \phi_s)) + \sum_{LJ} 4\varepsilon \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} + \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \sum_{Coulmob} f \frac{q_i q_j}{\varepsilon_r r_{ij}} \quad (4.7)$$

K_r (kcal/mol/Å^2), K_θ (kcal/mol/rad^2) and K_ϕ (kcal/mol) are the force constants for the bond, bond angles and dihedral angle, respectively; r (Å) and θ (deg) are the bond length and bond angle, respectively; r_0 (Å) and θ_0 (deg) are the bond lengths and angles for the optimized geometry, respectively; n is multiplicity, ϕ (deg) is the dihedral angle and the phase angle (ϕ_s) takes the value of 0 degree. The non-bonded term is represented by Lennard-Jones and Coulomb terms, where, ε (kcal/mol) is the depth of the region surrounding a local minimum of potential energy; lower case subscripts i and j are used to denote particles; σ_{ij} (Å) is the finite distance at which the inter-particle potential is zero; $r_{ij} = r_j - r_i$ where r_i (Å) and r_j (Å) are the position vectors of particles; $f = \frac{1}{4\pi\varepsilon_0} = 332.38154$ kcal Å mol^{-1} e^{-2}; q_i (e) and q_j (e) are the partial atomic charges; ε_r is the dielectric constant including the effect of the medium that is not explicitly represented and equals 1.0 in a typical solvated environment where solvent is represented explicitly. Both scaling factors are identical to the existing AMBER force field for consistency.

4.5 Time-dependent methods for simulation techniques

The majority of chemical reactions are carried out in solution. Biologically relevant process occur in an aqueous system with specific pH and ionic conditions. Molecular properties are sensitive to the environment. Simulations are intimately related with describing solute-solvent interactions. There are two major techniques for generating an ensemble of structures: Monte Carlo and molecular dynamic (MD) methods. MD simulations are employed in my researches. MD simulations create a series of time-correlated points in a trajectory by propagating a starting set of

coordinates and velocities according to Newton's law by a series of finite time steps. In each step of the simulation, a potential function *V(r)* obtained through MM force field is used to evaluate the total force (F) acting on each particle. The initial positions can be obtained from the initial structures and the initial velocities can be generated according to atom types and simulation temperature. The total force is used to calculate the acceleration of each particle. Combined with current position and velocity, the acceleration of each particle is used to calculate the next positions and velocities in the next step of the simulations. Taylor series expansions, as approximations of the positions, velocities and accelerations, are used in several algorithms for the integration of the equation of motion[57, 58]. The numerical errors generated by each step are added up to become significant resulting in non-deterministic trajectories of MD simulations. Typically, the length of the time step is to be 1-2 femtoseconds which is small enough to avoid discretization errors, for example, smaller than the fastest vibrational frequency in the system. Two different statistical ensembles are common setups in MD simulations such as NVT and NPT ensembles, where N, V, T, P are the number of particles, volume V, temperature T and pressure P respectively. In NVT ensemble, the number of particles, system volume and temperature are fixed. Whereas, the number of particles, system pressure and temperature are fixed in NPT ensemble.

In Paper V, VI and VII different molecular dynamic methods are employed, containing normal MD simulation, steered MD simulation, umbrella sampling simulation and bias-exchange meta-dynamic simulation.

4.5.1 Normal MD simulation

Normal MD simulation is limited ability to climb over energy barriers since the uphill motion will generate a force to full the system back towards the minimum tending to only sample the region in phase space close to the starting condition.

4.5.2 Steered MD simulation

Steered MD simulation (SMD), also called force probe simulation, employs a pulling force to cause a change along desired reaction coordinates in a normal MD simulation[59]. Typically, a certain part of the protein is restrained by a harmonic potential and then apply either a constant velocity or a constant force to move the system along the desired reaction pathway.

4.5.3 Umbrella Sampling

Umbrella sampling which is a kind of biased molecular dynamic simulation, is a technique to calculate the free energies along the reaction coordinates[60]. The transition between one minimum

to the other requires a long time scale in the normal MD simulations if the free energy needed to overcome the barrier is high. It is thus impossible to reach the transition by using normal molecule dynamic simulation. Umbrella sampling is an efficient computational method to explore the free energy landscape along one or multi- reaction coordinates. In the umbrella sampling simulations, a biased potential is added to the free-energy landscape, in order to help the system to cross the transition state. A harmonic restraint are used to restrain the sampling at certain reaction coordinates are given:

$$V(s) = \frac{k}{2}(s - s_0)^2 \tag{4.8}$$

Then sample distribution will be:

$$P'(q) \propto P(q)e^{\frac{-k(s_{(q)}-s_0)^2}{2k_bT}} \tag{4.9}$$

where k is force constant; s_0 is the distances of reaction coordinates $s_{(q)}$; k_b is the Boltzmann constant and T is the temperature used in the simulations. The continuous set of simulations along the reaction coordinates are obtained by combining the simulations performed at different distances along the reaction coordinates. Umbrella sampling methods can be used when the reaction coordinate is known. The exact height of the free energy barrier is obtained after the simulations.

4.5.4 Metadynamic simulations

In some cases, the reaction coordinate is unknown, then the adaptive computational methods such as metadynamics simulation[61]. Metadynamics is to enhance the sampling through a bias that considers a set of collective variables (CVs). The bias is obtained by the history-dependent potential constructed as a sum of Gaussian distributions centered along the trajectory of the CVs[61]. The potential is:

$$V_G = w \sum_{t'=rG,2rG,\ldots} \exp\left(-\frac{(s(x)-s(x(t')))^2}{2\sigma_s^2}\right) \tag{4.10}$$

where w is the height of the Gaussian, σ_s is the width of Gaussian for the CVs, the rG is the time which Gaussian distributions are deposited and $s(x)$ represent the CV which provides a coarse-grained description of the system. The bias-exchange metadynamics method allows to correlate the CVs and to get an accurate description of the system and free energy landscape. The bias-exchange metadynamics simulations combine the replica exchanges with metadynamics. Several metadynamics simulations are run in parallel at the same temperature. A time-dependent potential

acting on a different collective variable are exchanged in the different variables are periodically performed based on the replica exchange scheme[62].

4.5.5. Weighted histogram analysis method (WHAM) for free energy calculations

The weighted histogram analysis method (WHAM) is an extension of Ferrenberg and Swendsen's Multiple Histogram Technique, which is used to calculate the free energy for umbrella sampling (free energy perturbation) and the bias-exchange metadynamics methods[63]. The advantages of using WHAM to calculate the free energies are: 1) The sampling errors are provided and thus produce objective estimates of the optimal location and length of additional simulations needed to achieve a desired level of precision, 2) The free energies taking into account all the simulations in order to minimize the statistical errors 3) the better estimates of free energy differences can be evaluated by multiple overlaps of probability distribution[63]. The Hamiltonian $\hat{H}_{\{\lambda\}}$ is given as:

$$\mathrm{H}_{\{\lambda\}} = \sum_{i=0}^{L} \lambda_i V_i\,(r) \tag{4.11}$$

Where *r* represents the coordinates of the molecule, *L* is the functions, $V_i(r)$ are biased potentials. The biased potentials are functions of the molecular coordinates *r*. The multiple biased potentials in which the sampling distributions are shifted along reaction coordinates are used for sampling reaction pathways by separating simulations with different coupling parameters $\{\lambda\}$ for different windows of the reaction pathways. The biased probability distribution functions $P^{i}_{\{\lambda\}}$ are generated in each window. The unbiased distribution function can be simply obtained by reweighting. The full probability distribution functions $P(q)$ can be expressed as the linear combination of $P^{i}_{\{\lambda\}}$:

$$P(q) = \sum C_i\, P^{i}_{\{\lambda\}} \tag{4.12}$$

Where C_i is a set of coefficients that needs to be optimized so as to minimize the total statistical error. C_i is given as:

$$C_i = n_i e^{-\beta V_i} / \langle e^{-\beta V_i} \rangle \tag{4.13}$$

Where n_i is the number of configuration samples in the window and $\sum C_i = 1$. Free energies or potential mean force can be obtained after the restraining potential corrections.

4.6 Computer approaches to build up the three dimensional protein structures.

It is necessary to have a three-dimensional protein and molecule structures for MD simulations and (TD)DFT calculations. However, the number of 3D protein structures are limited in the Protein

Date Base (PDB). Various computational methods can be used to generate the 3D structure of proteins based on physicochemical principles or empirical methods in which predictive schemes are distilled from the analyses of known protein structures.

4.6.1 Homology model

The most reliable way to determine the secondary structure taken up by a polypeptide is to map its amino acid sequence onto that of a homolog of known structure. Based on a template with a similar sequence, homology modelling is able to predict the target protein structure. Homology modelling has proven to be reliable to predict a 3D model of protein when the sequence similarities between the target protein and template protein are more than 25%. The homology modelling can be done through four steps: template selection, target-template alignment, model construction and model assessment. The procedure is shown in figure 8. There are many software tools available for homology modellings such as MODELLER[64], MOE[65], prime[66] and YASARA[67]. In my project, YASARA is the program for homology model generations for COP1 and UVR8 monomer in Paper VIII and IV respectively.

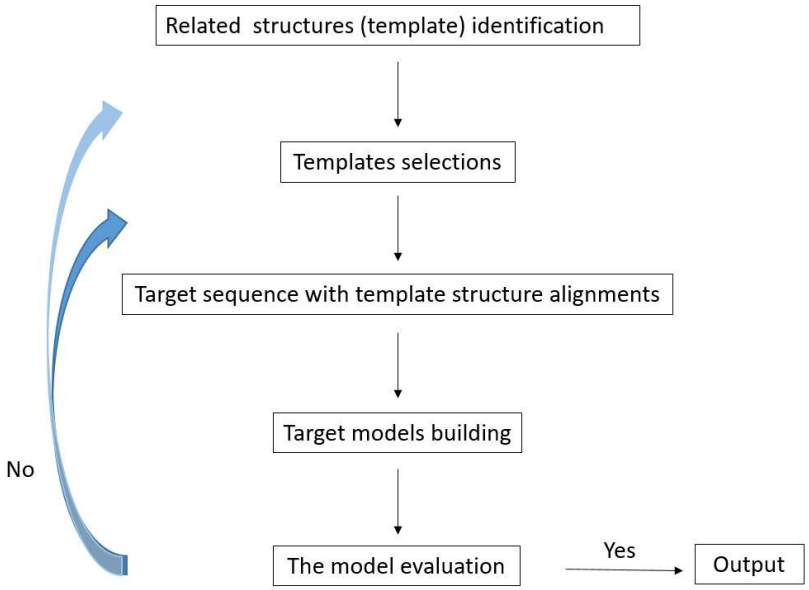


Figure 8. The steps of protein structure homology modelling.

4.6.2 Ab initio protein structure modelling

In some cases, the template does not exist or exist but cannot be identified, homology modelling cannot be used such as the case in Paper VIII. In those situations, models have to be constructed from scratch. Ab initio structure prediction is the method to build the models based on physical

principles rather than previously solved 3D structures. However, the success of ab initio modelling is limited to small proteins (<100 residues)[68]. In my research work, RosettaCommons is the program used to build the 3D structure of UVR8 C-terminus[69]. The procedure is shown in figure 9. The RosettaAbinitioRelax application in RosettaCommons program is based on the assumption that the distribution of structures adopted by the closely related sequences in known protein structures is able to approximate the distribution of conformations sampled in a local segment of a polypeptide chain [70]. In order to build the *ab initio* model of the fragment, all possible three- and nine-residue segments of the fragment were first extracted from the protein structure database by a sequence profile-profile comparison method[70]. The fragment libraries were generated from the RosettaServer: http://robetta.bakerlab.org/. The Monte Carlo procedure is employed to search the conformation defined by these fragments with an energy function that favors compact structures with paired β-strands and buried hydrophobic residues. For each query sequence, a large number of independent simulations were carried out starting from different random number seeds. The samples with "low-resolution" are first generated, followed by an all-atom refinement to obtain "high-resolution" models. The output structures have low free energy nonlocal interactions and consistency with the local conformational biases inherent in the sequence. Finally, the samples were divided into several clusters. In each cluster, Cα root means square derivation (RMSD) values among the structures are higher than a certain value or the structures with energy higher than certain value were filtered out. The centre structures of the largest clusters were selected as the highest confidence models.

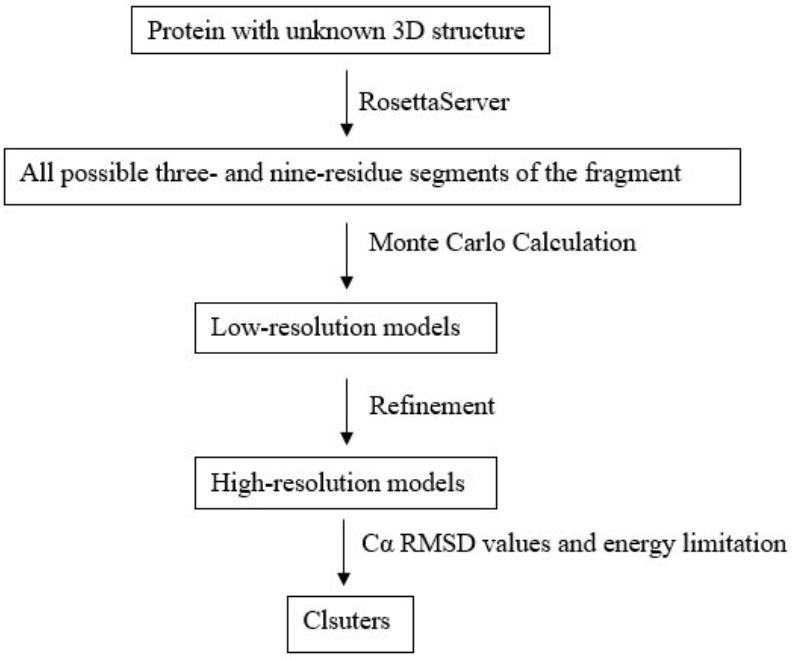


Figure 9. Illustration of the *Ab initio* protein structure modelling

4.6.3 Docking

Molecular docking is a widely used method to predict the binding poses of a molecule with its specific targets such as DNA and protein. Docking can be divided into four main steps: 1) active binding sites are identified from the targets; 2) a conformation searching of the docked molecule is performed using docking algorithms; 3) scoring functions are applied to assess the strengths of the binding poses throughout all predicted conformations of the ligand in step 2 and 4) the refinement process is applied to the binding poses with highest scores in step 3. In step 3), the protein is treated as fixed. Meanwhile, the ligand molecule is flexible. The ligand is able to move in order to take into account the ligand conformations' degrees of freedom. The binding sites on the target are not able to undergo any structure change. For a blind run, the RMSD values in the results have no physical meaning. Instead, it is a way to spread out the points in order to identify outliers. The vast majority of ligand conformations will generate unfavorable interactions between the partners. Therefore, the natural interactions should stand out from most of the binding poses with a low relative score. In step 4, the binding poses with lowest scores are taken from step 3,

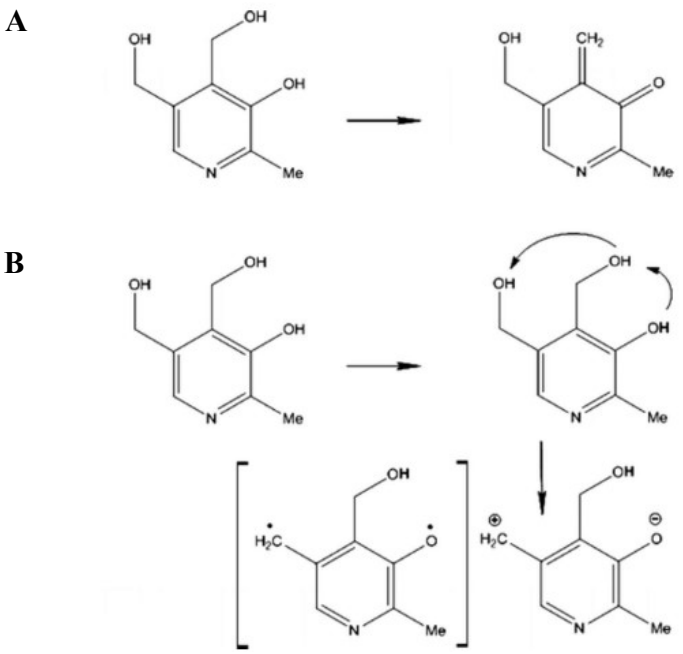
followed by a binding pose searching with both acceptor and ligand treated as flexible. The final docked poses are evaluated using a scoring function which estimates the binding free energy of a docked molecule to the receptor. It is a crucial step to distinguish the correct pose from those incorrect ones. Generally, the scoring functions make various assumptions and simplifications in the evaluation of binding free energy, in order to speed up the docking. Scoring functions can be classified into forcefield-based, empirical and knowledge-based scoring function. The forcefield-based scoring functions take into account, for instance, the nonbonded interactions such as van der Waals interactions and electrostatic interactions and solvent effects by implicit solvation methods such as Poisson-Boltzmann surface area (PBSA) or Generalized Born surface area (GBSA) [71]. The empirical scoring function based on counting the changes in solvent accessible surface area (ΔSASA) in the complex compared to the individual ligand and protein. The statistical results of intermolecular close contacts in large 3D structure databases such as PDB are employed by the knowledge-based scoring function. The forcefield-based scoring functions were employed in paper VIII.

Chapter 5. Summary of papers

My Ph.D. projects focus on the light-induced photoreaction in biomolecules and macromolecules such as proteins. I summary eight research papers here. The Papers I and II describe theoretical (TD)DFT studies on Vitamin B_6 and H_2O_2 molecules. The Paper III is about new AMBER ff99 parameters for the arginyl radical and hydrogenated arginyl side chains. The next five papers focus on the UV-B induced signaling pathways including UVR8 dimer monomerization mechanisms, UVR8-COP1 interaction, and the intrinsically disordered properties of UVR8 C-terminal.

5.1 The possible reactions of pyridoxine (vitamin B6) induced by UV-radiation

Pyridoxine, also known as Vitamin B_6, is an important nutrient in the human body. It is one of eight water-soluble B vitamins, with plants being the major source of intake for humans. Vitamin B6 is able to quench singlet oxygen and superoxide [72], being thus a potential antioxidant. In fact, plants apply vitamin B6 for self-protection against reactive oxygen species (ROS) [73]. The aim of this work is to explore the possible reactions of pyridoxine induced by UV-radiation. It is concluded that the aromatic ring-opening and photodegradation reactions regarding the loss of ring-bound functional groups (hydroxymethyl groups, a methyl group and a hydroxyl group) are not likely to occur upon exposure to UV-radiation because of the extremely high energy barriers in the excited transition states. However, it is well known that Pyridoxine generates quinone methides (QMs) by photolysis in aqueous solution [74]. QMs widely occur as reactive intermediates in the chemistry of phenols and related compounds [75]. They are believed to be fundamental intermediates in the biosynthesis of lignin as well as critical intermediates in the mechanism of action of many anticancer drugs. In thermal chemistry, *ortho*- and *para*-QMs are the most common isomers. The *meta*-QM is an example of a non-Kekulé molecule which is only accessible via photochemical routes. Due to the presence of $-CH_2OH$ groups in ortho and meta positions with respect to the aromatic hydroxyl group in pyridoxine (Scheme 2), competition will exist in this case. In this work, we found that the formation of *ortho*-QMs has a lower energy barrier than the non-Kekulé *m*-QM formation. This accounts for the selective formation of the *o*-QM compound after UV-irradiation of pyridoxine.



Scheme 2. Generation of QMs from Pyridoxine via Exposure to UV Radiation; (A) *o*-QM, (B) *m*-QM.

5.2 Hydrogen peroxide contributes to the ultraviolet-B (280-315 nm) induced oxidative stress of plant leaves through multiple pathways

Production of a variety of reactive oxygen species (ROS) was demonstrated in photosynthetic organisms in response to the exposure to high UV-B doses [76-78]. In controlled environment UV experiments, plants are typically first grown indoors in the absence of UV and then only mature leaves are exposed to supplementary UV radiation. Such an experiment, using UV-B tubes giving 170–180% of ambient doses, resulted in increased H_2O_2 concentrations in tobacco leaves [79]. Photobiologically, H_2O_2 is of interest because it also absorbs at UV wavelengths contained in sunlight. In this paper we report the possible roles of H_2O_2 in plants in response to UV-B irradiation. We present evidence for UV-B-induced enhancement of H_2O_2 production, and UV-B-induced conversion of both 'natural' and 'extra' H_2O_2 to ·OH radicals.

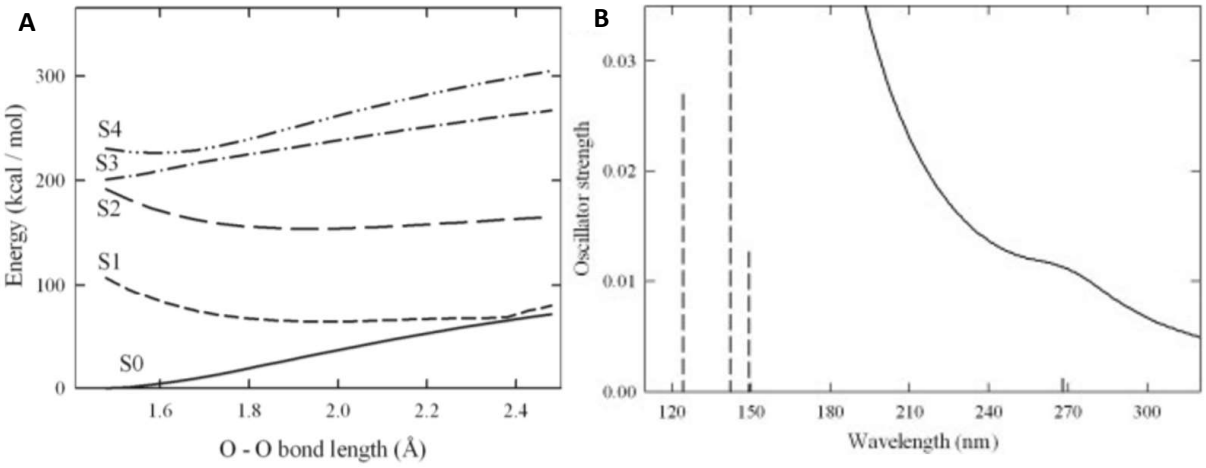


Figure 10. A) Energy curves for stretching the O-O bond in H_2O_2 in the ground state and the four lowest single excited states, B) UV absorption spectra of H_2O_2 and OH radical obtained at the B3LYP/6-31+G(d,p) level.

According to B3LYP/6-31+G(d,p) level calculations (figure 10B), the computed first absorption wavelength of H_2O_2 is located at 268 nm, and at 149, 142 and 124 nm for the second, third, and fourth absorption wavelengths. A scanning approach was employed to investigate the singlet excited state energies along the different pathways, in order to explore the possibility of OH radical generation by UV-induced H_2O_2 dissociation. The O–O bond was scanned from 1.48 Å to 2.38 Å, with a 0.05 Å step size. For each new point, the structure was re-optimized in the ground state and the vertical singlet excitation energies were calculated. The four lowest singlet excited states were included in the calculations along the reaction coordinate. Singlet state energy curves along the reaction coordinate obtained at the TD-B3LYP/6-31+G(d,p) level showed that the energies for H_2O_2 dissociation to OH are highly endothermic in the ground state and in the third and fourth singlet excited states throughout the scan of the O–O bonds shown in figure 10A. However, in the first and second excited states, the dissociation of H_2O_2 to OH radical is spontaneous and a calculated absorption wavelength of approximately 268 nm (figure 10B) corresponds to an expected experimental absorption maximum at approximately 280 nm due to the above-mentioned 10–15 nm blue-shift of 250–300 nm absorption wavelengths calculated by the current TD-DFT methodology.

5.3 Development of non-standard arginine residue parameters for use with the AMBER force fields

Amino acid radicals are often involved as intermediates in biological processes but are difficult to be experimentally captured. Computational modelling can be employed to study the features of the species involved. The neutral arginyl radical has previously been detected experimentally using ECD and ETD spectroscopy. Protonation of the radical can occur on the guanidinium carbon, depending on the peptide structure and protein environment. Accurate force fields are essential for reproducing the conformational and dynamic behaviour of these intermediates. New AMBER ff99 parameters for the arginyl radical and hydrogenated arginine side chains were presented in this work, based on *ab initio* calculations at the HF/6-31G* level. The optimized structures of standard arginine, arginyl radical and hydrogenated arginine were included in figure 11. The new AMBER parameters provide a very good agreement with the results obtained from the quantum calculations both in the structure of minimum energy, and torsional scans for arginyl radical and hydrogenated arginine side chains. Using the new parameters, the potential energy surfaces of dihedral angles showed average ΔE-RMSDs of less than 3 kcal/mol relative to the *ab initio* calculations. It is furthermore concluded that the AMBER parameters for arginyl radical and hydrogenated arginine cannot be estimated using the existing parameters from standard arginine residue in the AMBER force field with only the partial charges recalculated.

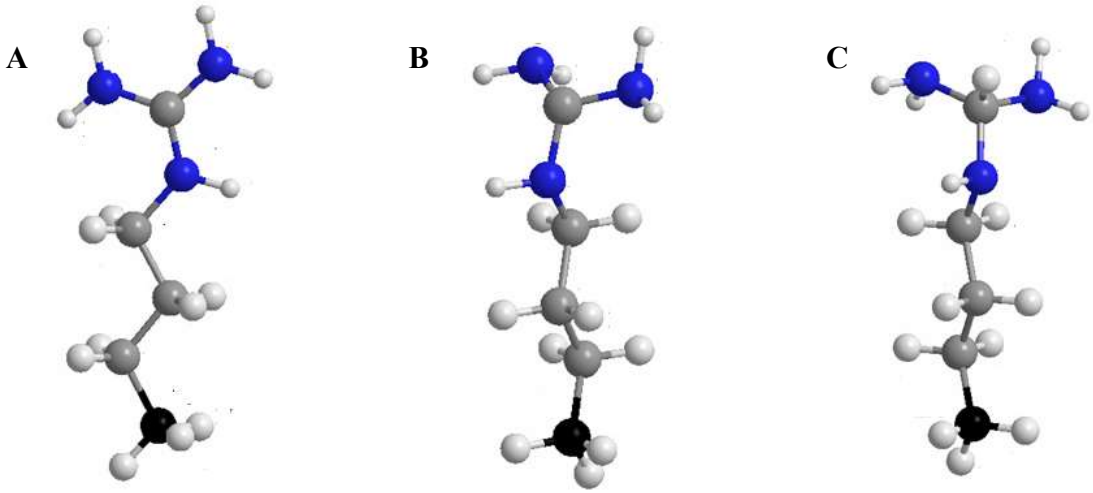


Figure 11. A) Optimized Standard arginine side chain, B) arginyl radical side chain and C) hydrogenated arginine side chain, obtained at the HF/6-31G* level.

5.4 Computational evidence for the role of *Arabidopsis thaliana* UVR8 as UV-B photoreceptor and identification of its chromophoric Amino Acids

The UVR8 protein is identified to be UV-B response photoreceptor that regulates more than 100 genes [16]. The action spectrum for the UVR8-dependent UV-B stimulation of HY5 transcription in mature *Arabidopsis thaliana* leaf tissue shows a major peak at 280 nm and a minor peak at 300 nm [18]. The reciprocal relationship of dose–response curves recorded at 300 nm, over a range of fluence rates, indicates that the UVR8-dependent pathway is controlled by photochemical reaction(s). Although the main peak of the action spectrum was at 280 nm, the substantial level of HY5 transcript accumulation occurred after absorption of UV-B at the secondary peak (at 300 nm). In this work, a homology model of the *Arabidopsis thaliana* UV resistance locus 8 (UVR8) protein was built, showing a seven-bladed β-propeller conformation similar to the globular structure of RCC1. The UVR8 amino acid sequence contains a very high amount of conserved tryptophan residues, and the homology model shows that seven of these tryptophan's cluster are located at the 'top surface' of the UVR8 protein where they are intermixed with positive residues (mainly arginines) and a couple of tyrosines. Quantum chemical (TD)DFT calculations of excitation spectra of both a large cluster model involving twelve residues and smaller fragments thereof reveal that absorption maxima appearing in the 280–300 nm range for the full cluster result from interactions between the central tryptophans and surrounding arginines seeing figure 12. This observation is in agreement with the published experimentally measured action spectrum for the UVR8-dependent UV-B stimulation of HY5 transcription in mature *A. thaliana* leaf tissue. In total, these findings suggest that UVR8 has in fact in itself the ability to be an ultraviolet-B photoreceptor in plants and the critical tryptophans W285 and W233 act as the chromophores in UVR8 photoreceptor.

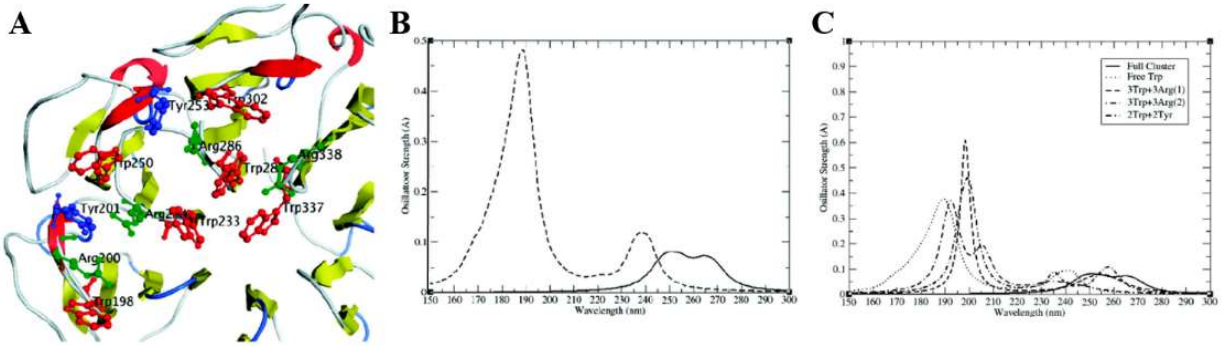


Figure 12. A) Residues used in the large cluster calculations of UV absorption spectra: 6 Trps (red), 4 Args (green), and 2 Tyrs (blue). B) Spectra of the optimized cluster (solid) and free Trp in solution

(dotted). C) Spectra of the full cluster, free Trp in solution, the 3Trp+3Arg (1) cluster containing (W302, W285, W233, R234, R286 and R338), the 3Trp+3Arg (2) cluster containing residues W337, W285, W233, R234, R286 and R338, and the peripheral 2Trp+2Tyr cluster containing residues Tyr201, Trp250, Tyr253, and Trp302. All spectra were computed at the ωB97XD/6-31G level of theory.

5.5 Interactions and stabilities of the UV Resistance Locus 8 (UVR8) protein dimer and its key mutants

The 3D structure of the wild-type UVR8 core domain published in 2012, shows the dimer structure with a seven-bladed propeller WD40 repeat domain in each monomer [23, 80]. However, the 11 most N-terminal amino acids and the 55 C-terminal amino acids are missing from the crystal structure. *A. thaliana* UVR8 has 14 tryptophan residues that are highly conserved among plant species [23, 27, 81]. The positions of 13 of the tryptophan residues are present in the crystal structure [23, 80], whereas residue W400 is located in the missing C-terminus region. Several salt bridges involving arginines R146, R286, R338 and R354 were found in the crystal structure [23, 80], but the crystal data are not unambiguous. In one study, the $UVR8^{R286A}$ mutant protein was shown to retain the dimeric structure and undergo UV-B-induced monomerization, whereas the dimeric structure was disrupted entirely by the double mutation R286A/R338A, when assayed by size exclusion chromatography [23]. However, in another paper, the $UVR8^{R286A}$mutant protein was demonstrated to be monomeric already in the absence of UV-B, when investigated by gel filtration.[80]. In order to investigate the strengths of the monomer-monomer interactions in the wild type UVR8 dimer and its mutants R286A, R338A and R286A/R338A. The free energy (ΔG) of dissociation of the wild type UVR8 dimer and its mutants R286A, R338A and R286A/R338A was calculated in this work using molecular dynamics (MD), steered molecular dynamics (SMD), seeing figure 13A-D, and umbrella sampling simulations employing the WHAM method [82-85]. Energy values are then obtained from PMF curves as a function of the reaction coordinate ξ [86-88], as shown in figure 13B.

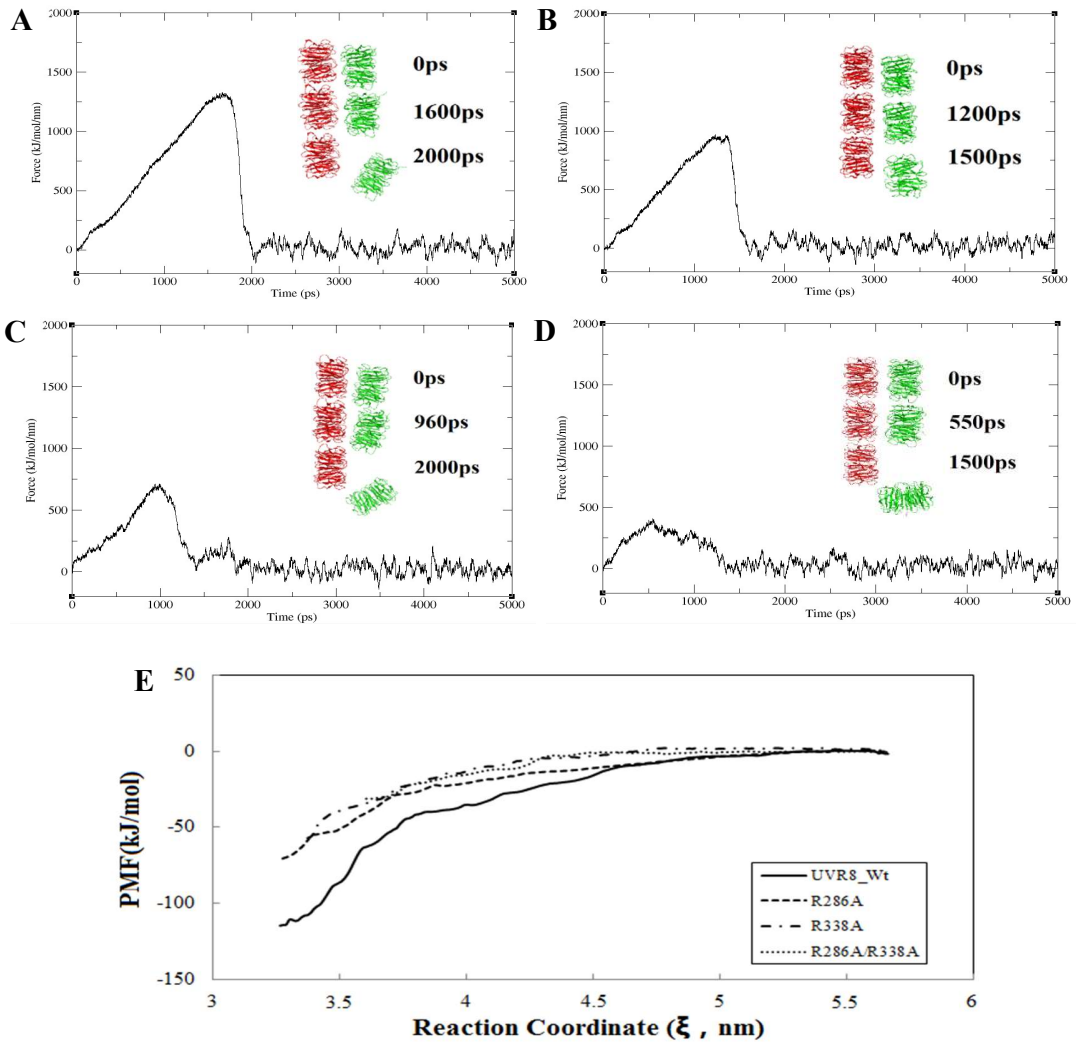


Figure 13. Dissociation pathways of the UVR8 wild type dimer and its mutants, with the corresponding plots of force vs time: A) wild type, B) R286A, C) R338A, D) R286A/R338A and Potential of mean force (PMF) curves for dissociation of wild type (solid), R286A mutant (dashed), R338A mutant (dot-dashed), and R286A/R338A double mutant (dotted) UVR8 dimers.

5.6 Photochemical reaction mechanism of UV-B-induced monomerization of UVR8 dimers as the first signaling event in UV-B-regulated gene expression in plants

The *A. thaliana* UVR8 protein has been identified to specifically mediate photomorphogenic UV-B responses by acting as a UV-B photoreceptor. The dimeric structure of the UVR8 protein dissociates into signaling-active monomers upon UV-B exposure, and the monomers rapidly

interact with downstream signaling components to regulate gene expression. UVR8 monomers revert to dimers in the absence of UV-B radiation, thereby reversing transcription activation [22, 23]. UVR8 amino acid residues W233 and W285 have been identified to play critical roles in the UVR8 dimer for the response to UV-B irradiation. In the present work, the photoreaction mechanism for UVR8 monomerization is explored with quantum chemical cluster calculations and evaluated by molecular dynamics simulations using the wild-type UVR8 dimer and novel force field parameters developed for intermediate radicals formed in the photochemical process. Three different models are investigated as shown in figure 14, suggesting that the preferred mechanism for UVR8 monomerization involves electron transfer from residue W233 to W285 and onward to R338 initiated by UV-B irradiation, coupled to simultaneous proton transfer from W233 to D129 leading to the formation of protonated D129, a deprotonated W233 radical, and a neutral R338 radical. Due to the formation of the neutral R338 radical, salt bridges involving this residue are disrupted together with the concomitant interruption of several other key salt bridges R286-D96, R286-D107, R338-D44, R354-E43, and R354-E53. The resulting large decrease in the protein–protein interaction energy, arising from this sequence of events, leads to the monomerization of the UVR8 dimer. The mechanism presented is in accord with all experimental data available to date.

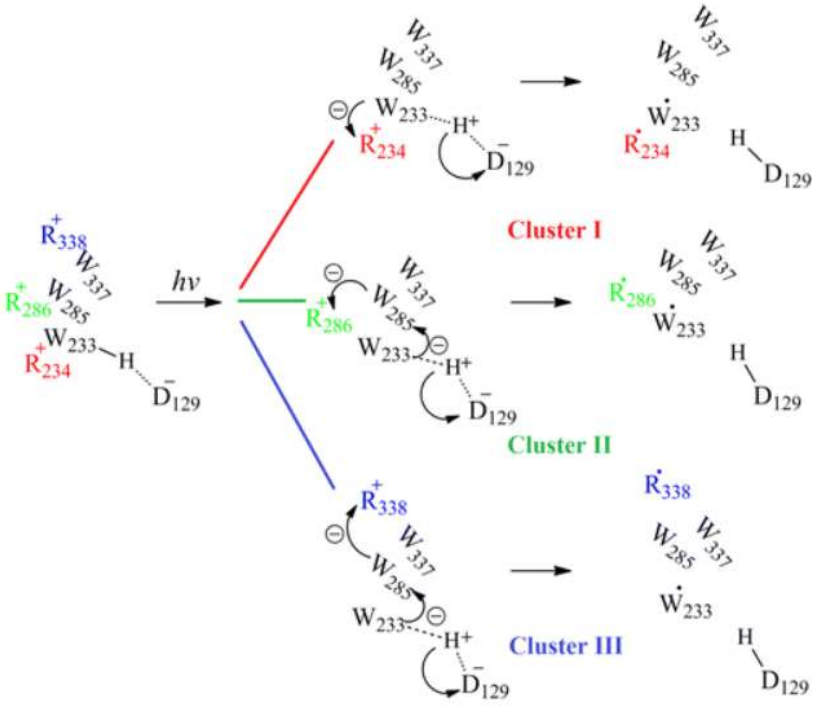


Figure 14. Three possible photoinduced reaction mechanism pathways triggering UVR8 dimer monomerization.

5.7 The intrinsically disordered domain of UVR8. Its inverted free energy landscape and predicted interactions with cancer promoter protein COP1.

UVR8 is a UV-B photoreceptor responsible for initiating UV-B signalling in plants. UVR8 exists as a homodimer in its inactivated form. However, upon the absorption of UV-radiation, the protein monomerize into its photoactivated state [22, 23]. In its monomeric form, UVR8 is available to bind the E3 ubiquitin ligase COP1 [25], hence triggering the subsequent UV-B dependent photomorphogenic development in plants. Recent *in vivo* experiments have shown that the UVR8 C-terminus region (AA 397-423) alone is sufficient to regulate the activity of COP1. In this work presented here, bias-exchange metadynamics simulations have been performed to evaluate the free energy landscape of $UVR8^{397-423}$ and $UVR8^{397-440}$. Simulations reveal an inverted free energy landscape where a disordered structure is found in the global energy minima (figure 15A). Flanking the global energy minima, more structured states are found at higher energies as shown in figure 15B. We confirmed the theoretical predictions with experiments, showing that the UVR8 C-terminal is an unstructured domain, and indeed, at higher temperatures the domain gains secondary structure and thus, a more ordered structure is obtained. Furthermore, stabilization of the low

energy disordered state has been pinpointed to a Proline residue, P411. This Proline is also a key residue in the binding to COP1 where a polyproline II appears to be the most favourable conformation in the complex formed.

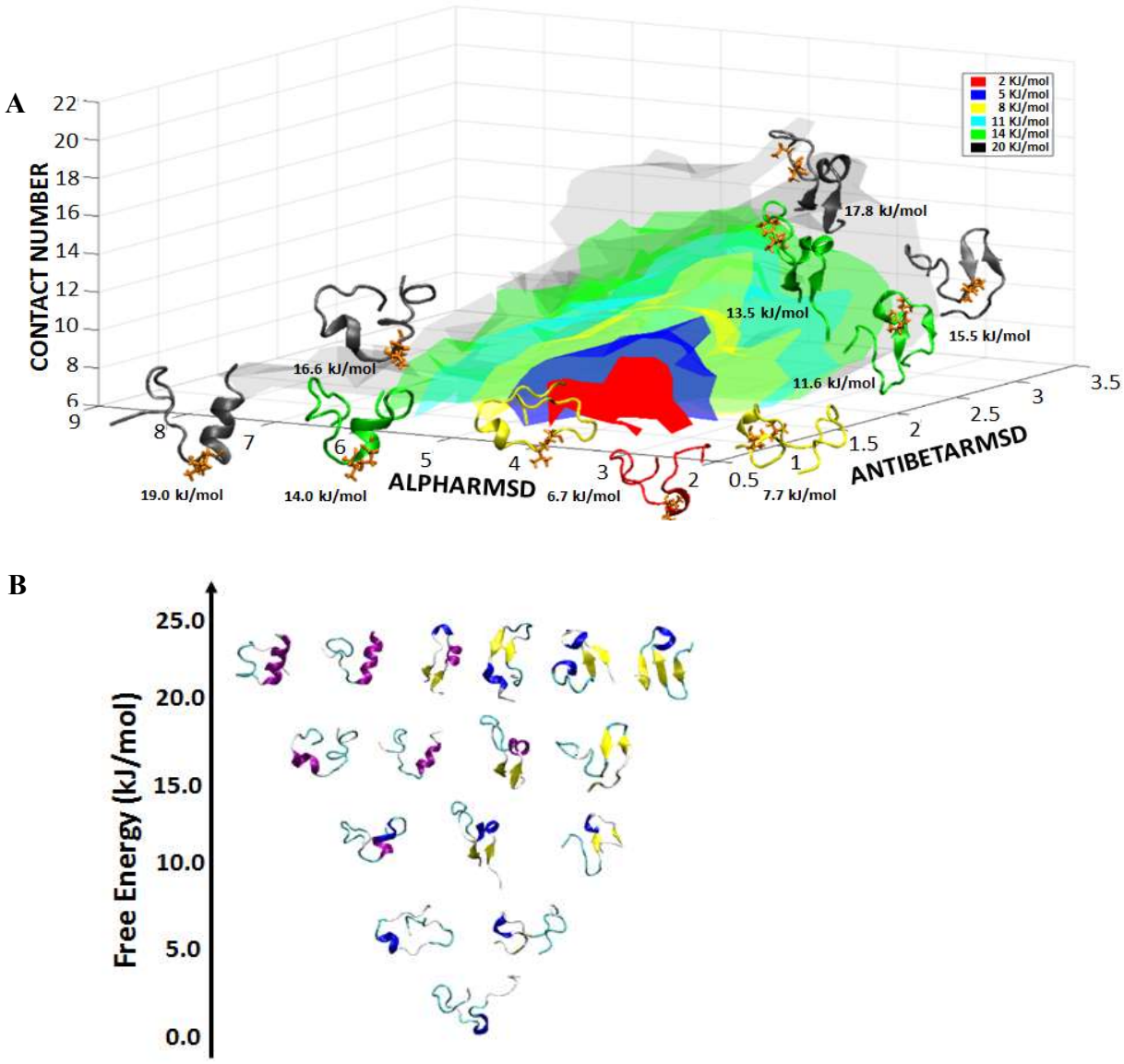


Figure 15. A) Three-dimensional representation of the free energy landscape of UVR8$^{397-423}$ peptide as a function of three collective variables: contact number (the number of hydrophobic contacts), AlphaRMSD representing α-helix content and AntiBetaRMSD for anti-parallel β-sheet content. Isosurfaces are shown at 2 (red), 5 (blue), 8 (yellow), 11 (cyan), 14 (green) and 20kJ/mol (gray); the white region is not reached as they have higher energies. Representative structures

sampled during the simulation are also shown with their relative free energies. B) Characterization of the high free energy structures of the UVR8$^{397-423}$ peptide.

5.8 Theoretical prediction of the protein–protein interaction between Arabidopsis thaliana COP1 and UVR8

In plants, ultraviolet-B radiation (280–315 nm) regulates gene expression and plant morphology through the UVR8 photoreceptor. The first signaling event after quantal absorbance is the interaction of the UVR8 C-terminus with the E3 ubiquitin ligase COP1 [25]. The nature of the interaction between these two proteins is hitherto unknown. A homology model of the *A. thaliana* COP1 protein which has seven-bladed propeller WD40 repeated domain and de novo folds of the C-terminal 27 amino acid (amino acids 397–423) peptide of Arabidopsis UVR8 (UVR8$^{397-423}$), were reported in this work, shown in figures 16A and B. Using a theoretical computational docking protocol, the interaction between COP1 and UVR8 was predicted as shown in figure 16C and D. A core motif was identified in UVR8$^{397-423}$ comprising adjacent hydrophobic residues V410 and P411 together with a charged residue D412, homologous to the corresponding motifs in other COP1-binding proteins, such as HY5, Cry1, and salt tolerance proteins STO/STH. The protein–protein interaction between the COP1 WD40 repeated domain and UVR8$^{397-423}$ reveals binding within a region of COP1 overlapping with the binding site for HY5 and the other COP1-interacting proteins. This study provides a framework for understanding the docking between UVR8 and COP1, which in turn gives clues for experimental testing of UVR8/COP1 interaction.

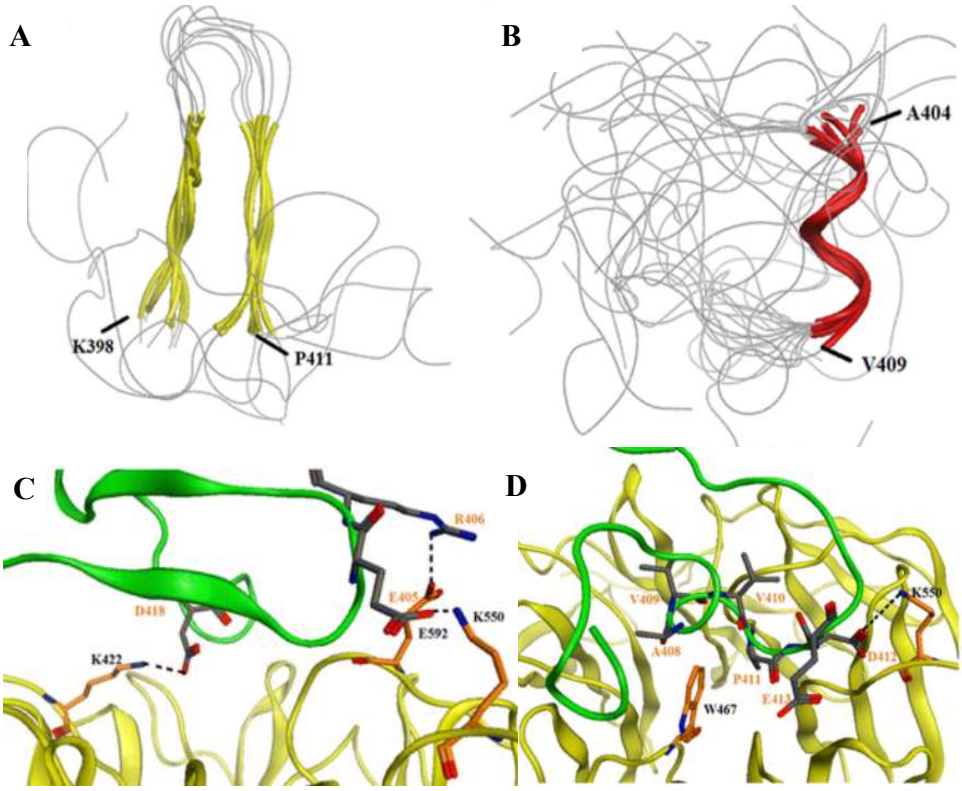


Figure 16. A) Six different UVR8$^{397-423}$ structures dominated by β-sheets (yellow) superposed on residues K398–V401 and residues A408–P411; B) Nineteen different UVR8$^{397-423}$structures dominated by an α-helix on residues A404–V409 (central red segment) superposed; C) The complex between wild-type COP1 and the modeled β-sheet-containing peptide UVR8$^{397-423}$. COP1 backbone colored yellow and UVR8 green; D) The modeled structure of the interaction complex between COP1 and the α-helix-dominated structure model of the UVR8$^{397-423}$ fragment. The carbon atoms in UVR8$^{397-423}$ are colored gray and in COP1 orange. Oxygen atoms are red, and nitrogen atoms are dark blue.

Future outlook

Some questions remaining for UVR8 action

Some important questions remain to be addressed in future are discussed below.

First, the 14 tryptophan residues in UVR8 are highly conserved in number and position in sequences from diverse species. It is evident that residues W233 and W285 as chromophores are of crucial importance in the mechanism of photoreception. However, many light-harvesting hypothesizes have been proposed. Hence, detailed photobiological studies need to be performed in order to determine the roles of the tryptophans in UVR8 action in vivo.

Second, how does UVR8 interact with COP1? UVR8 interacts with COP1 WD40 repeated domain via its C27 region in the sequences of 397-423 and residues V410 and P411 are essential for the interaction. The position of the C27 region is unknown because it is missing in the crystal structures. It has been suggested that the C-terminus could interact with the N-terminus of UVR8 to form a β-sheet structure that impairs interaction with COP1[48]. However, it is still unknown how structure changes are associated with photoreception and monomerization.

Third, how and where does UVR8 redimierization take place? UVR8 active monomers are rapidly redimerized in the present of RUP proteins. It is not clear that the RUPs-induced UVR8 redimerization is solely due to the inactivation of activated residues in UVR8 monomer or the accumulation increase of UV-B-induced RUPs limiting the binding between COP1 and UVR8. Furthermore, whether the redimerization process takes place in the nucleus or the cytosol is still not clear.

Fourth, how does UVR8 activated monomer or dimer move back to the cytosol? Nuclear localizations of UVR8 activated monomers take place in the present of COP1 indicating that the NLSs of COP1 proteins are able to address UVR8 activated monomers to nuclear from cytosol because of the physical interactions between COP1 and UVR8. On the other hand, COP1 proteins possess the nuclear export sequence (NES) domains indicating that COP1 may also regulate UVR8 proteins nuclear export. However, the mechanisms of the UVR8 proteins' nuclear import and export via COP1 proteins are still unclear.

Photoreceptor employed for optogenetic systems

Optogenetics is a method used to control the electrical activity of excitable cells by a combination of genetic and optic approaches. In this case, light can be employed to regulate the signals in living cells and organism. Nowadays, a number of photoreceptor proteins such as Phytochromes[89-91], Cryptochromes[92, 93], LOV domains[94-97] and UVR8[98-100] proteins that change their conformations in response to different wavelengths of light and regulate a large number of signaling pathways in plants have been engineered to control protein localization, pathway activation, gene expressions and so on, listed in Table 1.

Table 1. The photoreceptors engineered to control different physiological processes in either yeast or mammalian cells are presented through various papers.

light-controlled physiological processes	Phytochromes	Cryptochromes	LOV	UVR8
Nuclear addressing	-	-	[95]	[98]
Transcription activiation	[91]	[93]	[96]	-
Conformation change	[90]	[92]	[97]	-
Membrane addressing	[89]	[93]	[94]	-
Homo/Heterodimerization	[89]	[93]	[94]	[99, 100]

The advantage of using light-sensitive proteins are that the light-controlled proteins are much easier to deliver to the targets because they can be expressed instead of injection[101]. How the UVR8 protein works in the other living cells such as the mammalian cell can be studied in the future.

Acknowledgement

First of all, I would like to thank my supervisor, **Prof. Leif A. Eriksson,** without whom this thesis would not have been possible. Thank you, Leif, for all your invaluable help, support and guidance during the past eight years and for introducing me into the field of computational chemistry.

I would also like to thank my co-supervisor **Prof. Åke Strid** for all his help and assistance, and for introducing me to plant field.

I would also like to thank my advisor **Prof. Graham Fleming** who give me an opportunity to study spectroscopies in his group for two semesters. Thank you, Graham, for your support and organization during my stay.

I would also like to thank **Dr. Jiejie Chen** for her helpful discussions and suggestions during the time in Fleming's group.

Thanks to **Helena Bergkvist** for taking care of all administrative issues and **Prof. Johan Boman** for administrating my teaching.

All present and former colleagues in the Eriksson's group in Galway, Örebro, and Gothenburg deserve great thanks for making my working days more pleasant. My special thanks go to **Dr. Boxue Tian, Dr. Emma Eriksson, Dr. Yaoquan Tu, Dr.Samuel Genheden and Dr. Anna Reymer** for discussions.

I want to thank **the Wallenberg Foundation** and **Lundgrenska foundations** for the financial support of my travelling during my Ph.D. and I want to thank the donation from **Paul och Marie Berghaus donationsfond** and **Stipendiefonden Viktor Rydbergs** for the financial support of my exchange Ph.D. study in University of California Berkeley, U.S.

I want to thank **the national metacentre-Swedish National infrastructure for computing** including **C3SE** at Chalmers, Göteborg and **PDC** at Roy institute of Technology (KTH), Stockholm and **the Irish Center for High-End Computing (ICHEC)** in Ireland to provide me with the computer time in my Ph.D. studies.

Finally, I want to thank my family for their love and continuous support and encouragement. Special thanks to my husband **Bisheng Chen**, my daughters (**Richelle Chen** and **Ria Chen**), my parents (**Jianliang Wu** and **Cuiying Zhou**) and parents in law (**Xuehao Chen** and **Lili Quan**).

References

1. Hansen, J., et al., *Earth's energy imbalance: Confirmation and implications.* Science, 2005. **308**(5727): p. 1431-1435.

2. Snyder, C.S., et al., *Review of greenhouse gas emissions from crop production systems and fertilizer management effects.* Agriculture Ecosystems & Environment, 2009. **133**(3-4): p. 247-266.

3. Moan, J., *Visible light and UV radiation. Radiation at home, outdoors and in the workplace.* 2001, Oslo: Scandinavian Science. 69-85.

4. Holick, M.F., *Sunlight, UV-radiation, vitamin D and skin cancer: How much sunlight do we need?* Sunlight, Vitamin D and Skin Cancer, 2008. **624**: p. 1-15.

5. Hideg, E., et al., *UV-B exposure, ROS, and stress: inseparable companions or loosely linked associates?* Trends in Plant Science, 2013. **18**(2): p. 107-115.

6. Pham-Huy, L.A.et al., Free radicals, antioxidants in disease and health. International Journal of Biomedical Science, 2008. 4(2): p. 89-96.

7. Agati, G. and M. Tattini, *Multiple functional roles of flavonoids in photoprotection.* New Phytologist, 2010. **186**(4): p. 786-793.

8. Schreiner, M., et al., *UV-B-Induced Secondary Plant Metabolites - Potential Benefits for Plant and Human Health.* Critical Reviews in Plant Sciences, 2012. **31**(3): p. 229-240.

9. Nawkar, G.M., et al., *UV-Induced Cell Death in Plants.* International Journal of Molecular Sci., 2013. **14**(1): p. 1608-1628.

10. Nimse, S.B. and D. Pal, F*ree radicals, natural antioxidants, and their reaction mechanisms.* Rsc Advances, 2015. **5**(35): p. 27986-28006.

11. Czegeny, G., et al., *Hydrogen peroxide contributes to the ultraviolet-B (280-315 nm) induced oxidative stress of plant leaves through multiple pathways.* FEBs Letters, 2014. **588**(14): p. 2255-2261.

12. Nita, M. and A. Grzybowski, *The Role of the Reactive Oxygen Species and Oxidative Stress in the Pathomechanism of the Age-Related Ocular Diseases and Other Pathologies of the Anterior and Posterior Eye Segments in Adults.* Oxidative Medicine and Cellular Longevity, 2016. **2016**(2016):p. 1-23.

13. Tripathy, B.C. and R. Oelmuller, *Reactive oxygen species generation and signaling in plants.* Plant Signaling & Behavior, 2012. **7**(12): p. 1621-1633.

14. Li, J., et al*., Phytochrome signaling mechanisms.* The Arabidopsis Book, 2011. **9**: p. e0148.

15. Huang, X., et al., *Photoactivated UVR8-COP1 module determines photomorphogenic UV-B signaling output in Arabidopsis.* PLoS Genet, 2014. **10**(3): p. e1004218.

16. Brown, B.A., et al., *A UV-B-specific signaling component orchestrates plant UV protection.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(50): p. 18225-18230.

17. Oravecz, A., et al., *CONSTITUTIVELY PHOTOMORPHOGENIC1 is required for the UV-B response in Arabidopsis.* Plant Cell, 2006. **18**(8): p. 1975-1990.

18. Brown, B.A., et al., *UV-B Action Spectrum for UVR8-Mediated HY5 Transcript Accumulation in Arabidopsis.* Photochemistry and Photobiology, 2009. **85**(5): p. 1147-1155.

19. Gruber, H., et al., *Negative feedback regulation of UV-B-induced photomorphogenesis and stress acclimation in Arabidopsis.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(46): p. 20132-20137.

20. Tilbrook, K., et al., *The UVR8 UV-B Photoreceptor: Perception, Signaling and Response.* The Arabidopsis Book, 2013. **11**: p. e0164.

21. Jenkins, G.I*., The UV-B Photoreceptor UVR8: From Structure to Physiology.* Plant Cell, 2014. **26**(1): p. 21-37.

22. Wu, D., et al., *Structural basis of ultraviolet-B perception by UVR8.* Nature, 2012. **484**(7393): p. 214-219.

23. Christie, J.M., et al., *Plant UVR8 Photoreceptor Senses UV-B by Tryptophan-Mediated Disruption of Cross-Dimer Salt Bridges.* Science, 2012. **335**(6075): p. 1492-1496.

24. Gupta, A., et al., *Ultraviolet Radiation in Wound Care: Sterilization and Stimulation.* Advances in Wound Care (New Rochelle), 2013. **2**(8): p. 422-437.

25. Cloix, C., et al., *C-terminal region of the UV-B photoreceptor UVR8 initiates signaling through interaction with the COP1 protein.* Proceedings of the National Academy of Sciences of the United States of America, 2012. **109**(40): p. 16366-16370.

26. O'Hara, A. and G.I. Jenkins, *In vivo function of tryptophans in the Arabidopsis UV-B photoreceptor UVR8.* Plant Cell, 2012. **24**(9): p. 3755-3766.

27. Rizzini, L., et al., *Perception of UV-B by the Arabidopsis UVR8 protein.* Science, 2011. **332**(6025): p. 103-106.

28. Wu, M., et al., *Photochemical Reaction Mechanism of UV-B-Induced Monomerization of UVR8 Dimers as the First Signaling Event in UV-B-Regulated Gene Expression in Plants.* Journal of Physical Chemistry B, 2014. **118**(4): p. 951-965.

29. Voityuk, A.A.,el al., *On the mechanism of photoinduced dimer dissociation in the plant UVR8 photoreceptor.* Proceedings of the National Academy of Sciences of the United States of America, 2014. **111**(14): p. 5219-5224.

30. Li, X., et al., *Theoretical Study on the UVR8 Photoreceptor: Sensing Ultraviolet-B by Tryptophan and Dissociation of Homodimer.* Journal of Chemical Theory and Computation, 2014. **10**(8): p. 3319-3330.

31. Mathes, T., et al., *Proton-Coupled Electron Transfer Constitutes the Photoactivation Mechanism of the Plant Photoreceptor UVR8.* Journal of the American Chemical Society, 2015. **137**(25): p. 8113-8120.

32. Miyamori, T., et al., *Reaction dynamics of the UV-B photosensor UVR8.* Photochemical & Photobiological Sciences, 2015. **14**(5): p. 995-1004.

33. Zeng, X.L., et al., *Dynamic crystallography reveals early signalling events in ultraviolet photoreceptor UVR8.* Nature Plants, 2015. **1**(1). 14006-14020

34. Hofmann, N.R., *A Mechanism for Inhibition of COP1 in Photomorphogenesis: Direct Interactions of Phytochromes with SPA Proteins.* Plant Cell, 2015. **27**(1): p. 8-8.

35. Yan, H.L., et al., *Nuclear Localization and Interaction with COP1 Are Required for STO/BBX24 Function during Photomorphogenesis.* Plant Physiology, 2011. **156**(4): p. 1772-1782.

36. Mcnellis, T.W., et al., *Overexpression of Arabidopsis Cop1 Results in Partial Suppression of Light-Mediated Development - Evidence for a Light-Inactivable Repressor of Photomorphogenesis*. Plant Cell, 1994. **6**(10): p. 1391-1400.

37. Deng, X.W., et al., *Cop1, an Arabidopsis Regulatory Gene, Encodes a Protein with Both a Zinc-Binding Motif and a G-Beta Homologous Domain.* Cell, 1992. **71**(5): p. 791-801.

38. Chen, S., et al., *The functional divergence between SPA1 and SPA2 in Arabidopsis photomorphogenesis maps primarily to the respective N-terminal kinase-like domain.* BMC Plant Biology, 2016. **16**:p. 165-177.

39. Murphy, J.M., et al., *Molecular Mechanism of CCAAT-Enhancer Binding Protein Recruitment by the TRIB1 Pseudokinase.* Structure, 2015. **23**(11): p. 2111-2121.

40. Uljon, S., et al., *Structural Basis for Substrate Selectivity of the E3 Ligase COP1.* Structure, 2016. **24**(5): p. 687-696.

41. Paoli, M., *Protein folds propelled by diversity.* Progress in Biophysics & Molecular Biology, 2001. **76**(1-2): p. 103-130.

42. Migliorini, D., et al., *Cop1 constitutively regulates c-Jun protein stability and functions as a tumor suppressor in mice.* Journal of Clinical Investigation, 2011. **121**(4): p. 1329-1343.

43. Vitari, A.C., et al., *COP1 is a tumour suppressor that causes degradation of ETS transcription factors.* Nature, 2011. **474**(7351): p. 403-406.

44. Yang, H.Q., et al., *The signaling mechanism of Arabidopsis CRY1 involves direct interaction with COP1.* Plant Cell, 2001. 13(12): p. 2573-2587.

45. Holm, M., et al., *Identification of a structural motif that confers specific interaction with the WD40 repeat domain of Arabidopsis COP1.* Embo Journal, 2001. **20**(1-2): p. 118-127.

46. Holm, M., et al., *Two interacting bZIP proteins are direct targets of COP1-mediated control of light-dependent gene expression in Arabidopsis.* Genes & Development, 2002. **16**(10): p. 1247-1259.

47. Yin, R.H., et al., *Two Distinct Domains of the UVR8 Photoreceptor Interact with COP1 to Initiate UV-B Signaling in Arabidopsis.* Plant Cell, 2015. **27**(1): p. 202-213.

48. Yang, X.J., et al., *How Does Photoreceptor UVR8 Perceive a UV-B Signal?* Photochemistry and Photobiology, 2015. **91**(5): p. 993-1003.

49. Findlay, K.M. and G.I. Jenkins, *Regulation of UVR8 photoreceptor dimer/monomer photo-equilibrium in Arabidopsis plants grown under photoperiodic conditions.* Plant Cell Environ, 2016. **39**(8): p. 1706-1714.

50. Greco, C., et al., *Combining experimental and theoretical methods to learn about the reactivity of gas-processing metalloenzymes.* Energy & Environmental Science, 2014. **7**(11): p. 3543-3573.

51. van Mourik, T., et al., *Density functional theory across chemistry, physics and biology.* Philosophical transactions. Series A, Mathematical, physical, and engineering sciences, 2014. **372**(2011): p. 20120488-20120493.

52. Hohenberg, P. and W. Kohn, *Inhomogeneous Electron Gas.* Physical Review B, 1964. **136**(3b): p. B864-872.

53. Kohn, W. and L.J. Sham, *Self-Consistent Equations Including Exchange and Correlation Effects.* Physical Review, 1965. **140**(4a): p. 1133-1138.

54. Perdew, J.P. and K. Schmidt, *Jacob's ladder of density functional approximations for the exchange-correlation energy.* Density Functional Theory and Its Application to Materials, 2001. **577**: p. 1-20.

55. Runge, E. and E.K.U. Gross, *Density-Functional Theory for Time-Dependent Systems.* Physical Review Letters, 1984. **52**(12): p. 997-1000.

56. Weiner, S.J., et al*., A New Force-Field for Molecular Mechanical Simulation of Nucleic-Acids and Proteins.* Journal of the American Chemical Society, 1984. **106**(3): p. 765-784.

57. Hockney, R.W., *Computational Similarity (Reprinted from Concurrency Practice and Experience 7:p.. 147-166, 1995)*. Supercomputer, 1995. **11**(4): p. 102-123.

58. Verlet, L., *Computer Experiments on Classical Fluids .I. Thermodynamical Properties of Lennard-Jones Molecules.* Physical Review, 1967. **159**(1): p. 98-111.

59. Li, M.S. and B.K. Mai, *Steered Molecular Dynamics-A Promising Tool for Drug Design.* Current Bioinformatics, 2012. **7**(4): p. 342-351.

60. Torrie, G.M. and J.P. Valleau, *Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling.* Journal of Computational Physics, 1977. **23**(2): p. 187-199.

61. Barducci, A., et al., *Metadynamics.* Wiley Interdisciplinary Reviews-Computational Molecular Science, 2011. **1**(5): p. 826-843.

62. Granata, D., et al., *The inverted free energy landscape of an intrinsically disordered peptide by simulations and experiments.* Scientific Reports, 2015. **5**: p. 15449-15451.

63. Kumar, S., et al., *The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules .1. The Method.* Journal of Computational Chemistry, 1992. 13(8): p. 1011-1021.

64. Sali, A. and T. Blundell, *Comparative Protein Modeling by Satisfaction of Spatial Restraints.* Protein Structure by Distance Analysis, **1994**: p. 64-86.

65. Inc., C.C.G., Molecular Operating Environment (MOE), 2013.08. 2017: 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada.

66. Jacobson, M.P., et al., *A hierarchical approach to all-atom protein loop prediction.* Proteins-Structure Function and Bioinformatics, 2004. **55**(2): p. 351-367.

67. Krieger, E. and G. Vriend, *YASARA View-molecular graphics for all devices-from smartphones to workstations.* Bioinformatics, 2014. **30**(20): p. 2981-2982.

68. Raman, S., et al., *Structure prediction for CASP8 with all-atom refinement using Rosetta.* Proteins-Structure Function and Bioinformatics, 2009. **77**: p. 89-99.

69. Bonneau, R., et al., *Rosetta in CASP4: Progress in ab initio protein structure prediction.* Proteins-Structure Function and Genetics, **2001**: p. 119-126.

70. Bradley, P., et al., *Toward high-resolution de novo structure prediction for small proteins.* Science, 2005. 309(5742): p. 1868-1871.

71. Kitchen, D.B., et al., *Docking and scoring in virtual screening for drug discovery: methods and applications.* Nature Reviews Drug Discovery, 2004. **3**(11): p. 935-949.

72. Karuppanapandian, T., et al., *Reactive oxygen species in plants: their generation, signal transduction, and scavenging mechanisms.* Australian Journal of Crop Science, 2011. **5**(6): p. 709-725.

73. Havaux, M., et al., *Vitamin B6 deficient plants display increased sensitivity to high light and photo-oxidative stress.* BMC Plant Biology, 2009. **9**:130-155.

74. Brousmiche, D. and P. Wan, *Photogeneration of an o-quinone methide from pyridoxine (vitamin B-6) in aqueous solution.* Chemical Communications, **1998**(4): p. 491-492.

75. Wan, P., et al., *Quinone methide intermediates in organic photochemistry.* Pure and Applied Chemistry, 2001. **73**(3): p. 529-534.

76. Barta, C., et al., *Differences in the ROS-generating efficacy of various ultraviolet wavelengths in detached spinach leaves.* Functional Plant Biology, 2004. **31**(1): p. 23-28.

77. Hideg, E. and I. Vass, *UV-B induced free radical production in plant leaves and isolated thylakoid membranes.* Plant Science, 1996. **115**(2): p. 251-260.

78. Majer, P. and E. Hideg, *Developmental stage is an important factor that determines the antioxidant responses of young and old grapevine leaves under UV irradiation in a green-house.* Plant Physiology and Biochemistry, 2012. **50**: p. 15-23.

79. Coles, B.F. and F.F. Kadlubar, *Detoxification of electrophilic compounds by glutathione S-transferase catalysis: Determinants of individual response to chemical carcinogens and chemotherapeutic drugs? (Reprinted from Thiol Metabolism and Redox Regulation of Cellular Functions)*. Biofactors, 2003. **17**(1-4): p. 115-130.

80. Wu, D., et al., *Structural basis of ultraviolet-B perception by UVR8.* Nature, 2012. **484**: p. 214-219.

81. Wu, M., et al., *Computational evidence for the role of Arabidopsis thaliana UVR8 as UV-B photoreceptor and identificantion of its chromophore amino acids.* Journal of Chemical Information and Modeling, 2011. **51**.1287-1295

82. Jarzynski, C., *Nonequilibrium equality for free energy differences.* Physical Review Letters, 1997. **78**(14): p. 2890-2693.

83. Kumar, S., et al., *The weighted histogram analysus method for free-energy calculations on biomolecules. I. The method.* Journal of Computational Chemistry, 1992. **13**(8): p. 1011-1021.

84. Hub, J.S., et al., *g_wham-A free weighted histogram analysis implementation including rohust error and autocorrelation estimates.* Journal of Chemical Theory and Computation, 2010. **6**(12): p. 3713-3720.

85. Hub, J.S. and B.L. de Groot, *Does CO2 permeate through Aquaporin-1?* Biophysical Journal, 2006. **91**: p. 842-848.

86. Torrie, G.N. and J.P. Valleau, *Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling.* Journal of Computational Physics, 1977. **23**(2): p. 187-199.

87. Torrie, G.N. and J.P. Valleau, *Monte Carlo free energy estimates using non-Boltzmann sampling:Application to the sub-critical Lennard-Jones fluid.* Chemical Physics Letters, 1974. **28**(4): p. 578-581.

88. Patey, G.N. and J.P. Valleau, *The free energy of spheres with dipoles: Monte Carlo with multistage sampling.* Chemical Physics Letters, 1973. **21**(2): p. 297-300.

89. Levskaya, A., et al., *Spatiotemporal control of cell signalling using a light-switchable protein interaction.* Nature, 2009. **461**(7266): p. 997-1001.

90. Folcher, M., et al., *Mind-controlled transgene expression by a wireless-powered optogenetic designer cell implant.* Nature Communications, 2014. **5**: p. 5392-5394.

91. Shimizu-Sato, S., et al., *A light-switchable gene promoter system.* Nature Biotechnology, 2002. **20**(10): p. 1041-1044.

92. Idevall-Hagren, O., et al., *Optogenetic control of phosphoinositide metabolism.* Proceedings of the National Academy of Sciences of the United States of America, 2012. **109**(35): p. E2316-2323.

93. Kennedy, M.J., et al*., Rapid blue-light-mediated induction of protein interactions in living cells.* Nature Methods, 2010. **7**(12): p. 973-975.

94. Pathak, G.P., et al., *Benchmarking of optical dimerizer systems.* ACS Synthetic Biology, 2014. **3**(11): p. 832-838.

95. Niopek, D., et al., *Engineering light-inducible nuclear localization signals for precise spatiotemporal control of protein dynamics in living cells.* Nature Communications, 2014. **5**: p. 4404-4408.

96. Polstein, L.R. and C.A. Gersbach, *Light-inducible spatiotemporal control of gene activation by customizable zinc finger transcription factors.* Journal of the American Chemical Society, 2012. **134**(40): p. 16480-16483.

97. Wu, Y.I., et al., *A genetically encoded photoactivatable Rac controls the motility of living cells.* Nature, 2009. **461**(7260): p. 104-108.

98. Crefcoeur, R.P., et al., *Ultraviolet-B-mediated induction of protein-protein interactions in mammalian cells.* Nature Communications, 2013. **4**: p. 1779-1781.

99. Chen, D., et al., *A light-triggered protein secretion system.* The Journal of Cell Biology, 2013. **201**(4): p. 631-640.

100. Muller, K. and W. Weber, *Optogenetic tools for mammalian systems*. Molecular BioSystems, 2013. **9**(4): p. 596-608.

101. Tischer, D. and O.D. Weiner, *Illuminating cell signalling with optogenetic tools*. Nature Reviews Molecular Cell Biology, 2014. **15**(8): p. 551-558.