

# Computational Evidence for the Role of *Arabidopsis thaliana* UVR8 as UV–B Photoreceptor and Identification of Its Chromophore Amino Acids

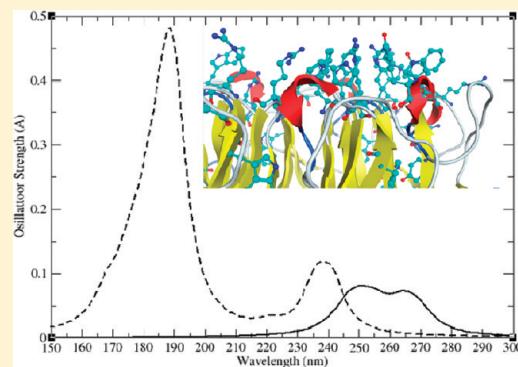
Min Wu,<sup>†</sup> Elin Grahn,<sup>‡</sup> Leif A. Eriksson,<sup>†</sup> and Åke Strid<sup>\*†</sup>

<sup>†</sup>School of Chemistry, National University of Ireland - Galway, Galway, Ireland

<sup>‡</sup>Örebro Life Science Center, School of Science and Technology, Örebro University, 70182 Örebro, Sweden

 Supporting Information

**ABSTRACT:** A homology model of the *Arabidopsis thaliana* UV resistance locus 8 (UVR8) protein is presented herein, showing a seven-bladed  $\beta$ -propeller conformation similar to the globular structure of RCC1. The UVR8 amino acid sequence contains a very high amount of conserved tryptophans, and the homology model shows that seven of these tryptophans cluster at the 'top surface' of the UVR8 protein where they are intermixed with positive residues (mainly arginines) and a couple of tyrosines. Quantum chemical calculations of excitation spectra of both a large cluster model involving all twelve above-mentioned 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. This observation coincides with the published experimentally measured action spectrum for the UVR8-dependent UV–B stimulation of HYS 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.



## ■ INTRODUCTION

Ultraviolet-B radiation (UV–B; 280–315 nm) is a part of the solar spectrum that reaches the surface of the Earth and that potentially has deleterious effects on biological molecules and therefore also on cells and living organisms.<sup>1,2</sup> Especially sessile organisms, such as plants, that are bound to the spot in which they are growing, are highly dependent on their environment and have to adjust their metabolism to reflect both physical challenge and short- and long-term changes. Plants must act on such stimuli and therefore the perception of environmental factors is of great importance. Photosynthetic, blue/UV-A, and infrared radiation are absorbed by light-harvesting chlorophylls and photosystems, cryptochromes and phototropins, and phytochromes, respectively. Since plants also specifically respond to UV–B challenge by alteration in gene expression and metabolism,<sup>3–11</sup> it is tempting to suggest that plants, in addition to the above-mentioned receptors, also contain one or more specific UV–B photoreceptors (PRs), and such PRs have been a prime target for UV–B photobiology research for more than a decade. In addition, to accomplish alteration of gene expression within the cellular nucleus, or to otherwise fine-tune metabolism, the cellular components responsible for the downstream signaling of the perception of UV–B quanta, also have been the targets for identification studies with some successful outcome.

A few signaling components conferring ultraviolet-B-dependent gene expression in *Arabidopsis thaliana*, the most well-studied

plant model organism, have been identified. These include the UV resistance locus 8 (UVR8) protein, the Constitutive photomorphogenesis 1 (COP1) protein, and the bZIP transcription factor Elongated hypocotyl 5 (HYS).<sup>9,11–14</sup> The COP1 protein contains four recognizable structural features: a zinc-binding RING-finger motif at the N-terminus, a potential coiled-coil domain (Helix), a central core domain, and a globular domain with seven WD40 repeats at the C-terminus.<sup>15</sup> Its nuclear abundance is regulated both by white light and by UV–B radiation.<sup>14</sup> In white light, COP1 interacts directly with HYS in the nucleus to regulate its activity negatively through proteolysis.<sup>16,17</sup> Under supplementary UV–B radiation, specific perception of radiation by a UV–B PR results in rapid UVR8/COP1 interaction, breaking up a possibly inactive UVR8 dimer existing in darkness.<sup>18</sup> This UVR8/COP1 interaction is closely linked to the UV–B PR function resulting in a change in the COP1 ubiquitin ligase-directed proteasome substrate specificity away from HYS and in turn leading to UVR8- and HYS-mediated activation of gene expression, including the HYS gene itself.<sup>11,13,14,19</sup> UVR8 and HYS regulate a large subset of all UV–B regulated genes in *Arabidopsis*,<sup>19</sup> although there is also evidence for alternative UV–B-regulatory pathways.<sup>7,19–21</sup> In the case of inducing HYS gene expression, the UVR8/COP1 complex binds to chromatin in the vicinity of the HYS promoter.<sup>22</sup>

Received: January 14, 2011

Published: May 11, 2011



**Figure 1.** a) Multiple alignment of 8 different UVR8 amino acid sequences using the COBALT algorithm at the NCBI Web site (<http://www.ncbi.nlm.nih.gov/tools/cobalt/>). The following abbreviations were used for denoting the different plant species, bryophytes (*Ppa*), and lycopodiophytes (*Smo*): *Ath* (*Arabidopsis thaliana*, AAD43920), *Vvi* (*Vitis vinifera*, XP\_002274569), *Rco* (*Ricinus communis*, XP\_002522929), *Ptr* (*Populus trichocarpa*, XP\_002309939), *Zma* (*Zea mays*, NP\_001141147), *Osa* (*Oryza sativa*, NP\_001052849), *Ppa* (*Physcomitrella patens*, XP\_001778783), *Smo* (*Selaginella moellendorffii*, XP\_002981556), the GenBank accession nos. for each sequence are also given. Sequence identity between the sequences is shown in bold letters high-lighted in yellow. All Trps are shown in bold red letters. All Arg (bold green letters high-lighted in yellow) and Tyr (bold blue letters high-lighted in yellow) that are conserved among all eight species are also indicated. b) Multiple alignment of the *Arabidopsis thaliana* UVR8 amino acid sequence with the human HERC2 (Pdb accession code 3KCI) and RCC1 proteins (Pdb accession code 1A12) using the CLUSTAL W (2.1) algorithm (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>). All Trps in the three sequences are shown in bold red letters, and the UVR8 Trps and those conserved in the other two sequences are also highlighted in yellow. In blue bold letters, all Tyr of the three sequences are shown. Seven of these are conserved in all eight proteins in Figure 1a (highlighted in yellow). All Arg of the three proteins are denoted with bold green letters. Fourteen of these are common to all eight sequences in Figure 1a and high-lighted in yellow.

It was previously found<sup>23</sup> that UVR8 accumulates in the nucleus in a UV-B-dependent fashion. It was also shown that addition of a common nuclear localization signal (NLS) N-terminally to UVR8 causes the fusion protein to be constitutively localized in the nucleus. The addition of a nuclear export signal (NES) to UVR8, on the other hand, causes the fusion protein to be fully localized in the cytoplasm. This cytosolic localization is easily overridden by brief UV-B exposure. Nuclear accumulation of UVR8 with an N-terminal 23-amino acid deletion is, however, impaired, which indicates that the N-terminal sequence plays a key role in the nuclear import mechanism. Nonetheless, the N-terminal deletion protein is still capable of binding to chromatin at the *HYS* promoter, similarly to the normal UVR8, indicating that it has the potential to function in transcriptional regulation of the *HYS* gene.<sup>23</sup>

The action spectrum for the UVR8-dependent UV-B stimulation of *HYS* transcription in mature *Arabidopsis thaliana* leaf tissue shows a major peak at 280 nm and a minor peak at 300 nm.<sup>24</sup> 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).<sup>24</sup> Although the main peak of the action spectrum was at 280 nm, substantial level of *HYS* transcript accumulation occurred after absorption of UV-B at the secondary peak (at 300 nm).

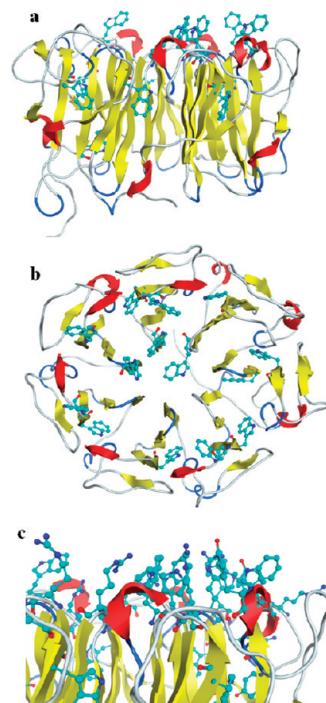
In addition, there are indications for more than one UV-B-dependent signaling pathway in plants<sup>19,20</sup> with slightly different absorption characteristics with respect to wavelength and/or intensity of the radiation.

The first component of the plant UV-B signaling pathways, the PRs, have been the subject for intense studies but have not yet been positively identified, neither by biochemical methods nor by genetic screens.<sup>13,18</sup> An intriguing possibility is that UVR8 itself may be one of these UV-B PRs regulating *HYS* and a large number of other UV-B-dependent genes. In a very recent report, it was concluded that UVR8 exists as a dimer when not activated by UV-B and that absorption by Trp285 is important for monomerization and thereby for subsequent COP1 binding, whereas UVR8<sup>W285A</sup> did not form dimers, and UVR8<sup>W285F</sup> was unable to monomerize.<sup>25</sup> However, no information regarding the possible structure of the UVR8 dimer was given.

In its primary structure, UVR8 has 14 highly conserved tryptophan residues, considerably more than in most proteins of equivalent size, a fact that in theory would permit effective absorption at 280 nm. Interactions between these tryptophans - all or a subset thereof - may also account for the secondary peak of *HYS* gene expression at 300 nm. To address the issue whether UVR8 would be the chief plant UV-B PR, the present computational study was conducted.

## METHODOLOGY

First, the *Arabidopsis thaliana* UVR8 amino acid sequence was aligned, using the COBALT algorithm (<http://www.ncbi.nlm.nih.gov/tools/cobalt/>) with another seven plant, bryophyte or lycopodiophyte UVR8 sequences to examine conserved motifs, with special emphasis on Trp residues (Figure 1a). Since no X-ray crystal structure or NMR structure of UVR8 is available as yet, a homology model was constructed, based on the *Arabidopsis thaliana* UVR8 amino acid sequence (Figure 1), and by performing an automated search of the Brookhaven protein data bank based on the PsiBlast algorithm. The search parameters are



**Figure 2.** Homology model of *Arabidopsis thaliana* UVR8 (residues 26–411): a) side view and b) top view. Tryptophans displayed in ball-and stick in a and b. c) Close-up of Trp/Tyr/Arg clustering of the homology model.  $\alpha$  helical structures are shown in red, and  $\beta$  sheet structures are shown in yellow.

given in Table S1 in the Supporting Information. All homology modeling was performed using the YASARA code and the YASARA2 force field.<sup>26</sup>

The pdb search generated three possible templates, pdb entries 1A12, 1I2M, and 3KCI, with 1A12 and 1I2M being the human Regulator of Chromosome Condensation protein 1 (RCC1). The highest similarity was found for 3KCI, which is the RCC1-like third RLD domain of human HERC2, a crystal structure solved at 1.80 Å resolution.<sup>27</sup> In Figure 1b, the amino acid sequence alignment of these three sequences is shown using the CLUSTAL W algorithm (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>).

After alignment, the homology model was subjected to loop refinements and simulated annealing minimization in aqueous solvent. Out of the total 440 amino acids of *A. thaliana* UVR8, neither the initial 25 amino acids of the N-terminus nor the last 29 residues of the C-terminus were included in the sequence modeled (cf. Figure 1b), which also reasonably fits with the nonconserved regions of the eight UVR8 proteins aligned in Figure 1a. Neither the N- nor the C-terminus contains tryptophan or tyrosine residues, of importance to the current study. The N-terminus has been shown crucial for nuclear import,<sup>24</sup> and the C-terminus plays an important role in UVR8-Chromatin interaction.<sup>22</sup> The resulting 386 amino acid long UVR8 model has 32% sequence identity to the 3KCI template, and the final  $C_{\alpha}$  RMSD value between the homology model and 3KCI is 0.727 Å. In order to relax the modeled UVR8 structure, this was then subjected to a 7.5 ns molecular dynamics (MD) simulation in aqueous solution, again using the YASARA code. Periodic boundary conditions were applied, using a simulation cell of 81 Å × 64 Å × 60 Å, and including a total of 32000 atoms.

Electrostatics were included using Particle Mesh Ewald summation and a cutoff radius for Coulombic and van der Waals interactions of 10.5 Å. The core domain of the UVR8 model shows no significant changes after the MD simulation, and the main difference in the structure of the homology model before and after MD simulation (rmsd = 1.377 Å) lies in the orientation of the N-terminal loop consisting of 8 amino acids, that curls up under the protein during the simulation (whereas the homology model predicted an extended tail stretching out from the globular bulk of the protein), and a slight shift in orientation of the final 8 amino acids of the C-terminal loop. Neither N- nor C-termini are however assigned an explicit role in the UV spectrum of UVR8.<sup>22,24</sup> In Figure 2, we display the final homology model after MD simulation in side (a) and top (b) view, with all Trp residues displayed in stick model. In Figure 2c we zoom in on the top region, to illustrate the clustering of Trp, Tyr, and Arg residues.

In order to explore the photochemistry of the resulting model, geometry optimizations and excited state calculations were performed on clusters of varying sizes, within the quantum chemical density functional theory (DFT) and time-dependent (TD)-DFT framework. DFT based methodology is the method of choice for optimizations and excited state modeling, when dealing with systems of the current size. Even so, the largest cluster investigated herein presents a considerable challenge for present day supercomputers. To this end, it was necessary to scale down the size of basis sets used to describe the atoms, after initial benchmarking against smaller systems. All optimizations were performed at the hybrid Hartree–Fock - DFT level B3LYP/6-31+G(d,p)<sup>28,29</sup> and absorption spectra were computed using the B3LYP and  $\omega$ B97XD<sup>30</sup> functionals together with the 6-31G and 6-31+G(d,p) basis sets. The accuracy of B3LYP in geometry optimizations of biomolecules is well established and normally provides geometries in very close agreement with experimental data. In terms of excited state calculations, TD-DFT has been employed successfully over the past decade to explore both spectra and photochemical properties on a wide range of systems. The B3LYP functional is known to blue-shift excitations (i.e., predict too high excitation energies) by 0.1–0.2 eV for non charge-transfer (CT) excitations, and the double for absorptions involving CT, relative to experiments.<sup>31,32</sup>  $\omega$ B97XD is an extended version of the long-range corrected  $\omega$ B97 and  $\omega$ B97X<sup>33</sup> functionals with an additional empirical dispersion correction and has been shown in general to perform very well in excited state calculations,<sup>34</sup> including long-range charge transfer excitations.<sup>30</sup> Again, a blue-shift relative to experiment has been noted in many cases. All quantum chemical calculations were performed using the Gaussian09 program.<sup>35</sup>

## RESULTS

**a. Amino Acid Alignments.** Figure 1a shows the amino acid alignment between the *A. thaliana* UVR8 protein and seven UVR8 proteins from other plants, a bryophyte, and a lycopodiophyte. All 14 Trps found in *A. thaliana* UVR8 are conserved also in all the other seven organisms except the first one (amino acid Trp39 in the *A. thaliana* sequence) that has been exchanged for a Phe in the lycopodiophyte *Selaginella moellendorffii*. This indicates a possible role for these Trps in UVR8 function as a UV–B signaling component. In addition, a large number of other amino acid residues are also conserved in UVR8 between species, including for instance a number of Tys (7 of them) and Args (14 conserved ones; Figure 1a). In fact, the total sequence

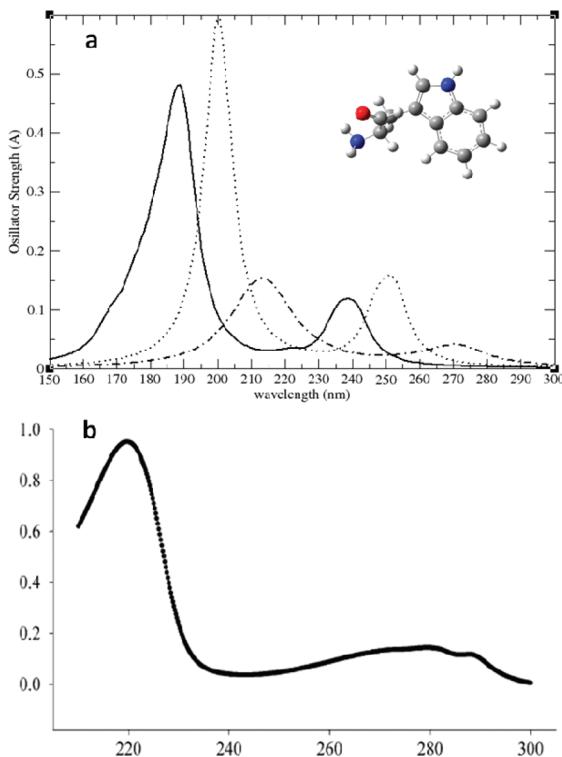
identity between the eight species is 52%, calculated on the basis of the 440 amino acid *A. thaliana* protein. In addition to this high identity score, a large number of amino acids are identical to most but not all eight proteins, and there are also a considerable number of conserved amino acid substitutions, which makes the overall homology between the eight UVR8s even greater.

In Figure 1b is shown the alignment between *A. thaliana* UVR8 and the human HERC2 and RCC1 proteins. Of the 14 Trps in UVR8, 2 are conserved in all three proteins, whereas another 4 UVR8 Trps are conserved in HERC2. In addition, HERC2 and RCC1 contain another 2 Trps each that are not conserved in the other sequences. Of the 14 tryptophans in UVR8, our homology model shows that seven are distributed on the top-surface of the protein, while the rest are buried within the core of the protein. The ones found in both the UVR8 model and the HERC2 and the RCC1 proteins are exclusively the ones found in the interior of the protein indicating that they are probably needed for the formation of the  $\beta$ -propeller fold. Of the seven Tys that are conserved in all eight proteins in Figure 1a, six are also conserved in at least one of HERC2 and RCC1. Four of them are present in all three proteins (UVR8, HERC2, and RCC1). Of the 14 Args conserved in all eight proteins shown in Figure 1a, only two are conserved in HERC2 and none in RCC1.

**b. Homology Model.** Using the amino acid sequence of the *A. thaliana* UVR8 protein and homology modeling tools (blast/psi-blast) and a full search of the Brookhaven protein data bank, a homology model was constructed based on the RCC1-like third RLD domain of human HERC2.

The UVR8 homology model has a seven-bladed  $\beta$ -propeller structure (Figure 2). The innermost region forms a barrel, which is covered on the exterior by seven 'propeller blades' aligned along the surface (best seen in Figure 2b), forming a highly compact globular protein. In this model, a clustering exists largely on the exterior of the protein at the top of the  $\beta$ -propeller structure of the "excess tryptophans". In between the Trp residues, we also find a couple of Tyr residues, four Arg residues and one Lys residue. Other possible aromatic side chains (Phe and His), which may assist in UV absorption, are scattered throughout the remainder of the protein. From the structure, it is intuitively clear that any UV perception activity of the protein would arise from the surface clustering of the many tryptophans. Figure 2c shows a close-up of the cluster with the above-mentioned Trp, Tyr, and Arg/Lys residues displayed. The cluster forms a tight hydrogen bonded and cation -  $\pi$ -system network (Arg-Trp and Arg-Tyr), with distances between interacting atoms in the range of 3.3–4.3 Å. The cluster structure is very stable also after 8.5 ns MD simulation in water.

Several UVR8 mutants have been investigated experimentally<sup>12,13,18</sup> (especially in the Supporting Information of ref 18). In the current work, three UVR8 exon 5 mutants were explored, G199R (mutant *uvr8-10*), G202R (*uvr8-9*), and G205E (*uvr8-11*). These mutants are known to be deficient in COP1 interaction,<sup>18</sup> but their UV–B-absorbing properties have not been reported. Homology models containing these mutants were generated, and it was found that all three mutants are located in the region between one of the propeller blades and the b-sheet 'core' of the protein. Most likely, the replacement of the above glycines for bulky and charged residues will force the propeller blade to extend away from the body of the protein. This may in turn impair dimer formation<sup>25</sup> or other protein–protein interactions. None of the three mutations are located in the vicinity of the crucial Trp/Arg/Tyr cluster.



**Figure 3.** a) Theoretical spectrum for free Trp in solution using different methods: B3LYP/6-31+G(d,p) dot-dashed,  $\omega$ B97XD/6-31+G(d,p) dotted, and  $\omega$ B97XD/6-31G solid line and b) experimental spectrum.<sup>26</sup>

**c. Absorption Spectra.** In order to benchmark the excited state calculations, the absorption spectrum of Trp alone was studied, using the different methodologies as outlined above. In Figure 3 we show the structure of tryptophan (3a insert), its computed spectra using different methods (Figure 3a) and the experimental spectrum<sup>36</sup> (Figure 3b), respectively. The tryptophan molecule was first geometry-optimized at the B3LYP/6-31+G(d,p) level, followed by TD-DFT calculations of the absorption spectrum.

The first peak in the experimental spectrum is located at 280 nm and the second, stronger, peak is found at 220 nm (Figure 3b). The first peak in the theoretically predicted spectrum computed at the optimization level B3LYP/6-31+G(d,p), Figure 2a, dot-dashed line, is predicted at 270 nm, and the larger peak at 215 nm, in good agreement with experimental data. The well-established blue-shift for B3LYP computed excitations compared to experiments (0.1–0.2 eV for non charge-transfer transitions) is seen also for this system. Using the  $\omega$ B97XD functional with the same basis set (Figure 3a, dotted line) gives an additional blue-shift, by ~15 nm. Using the smaller 6-31G basis set and the  $\omega$ B97XD functional renders peaks with a further 12–15 nm blue-shift (Figure 3a, solid line). All three methods reproduce the shape of the experimental spectrum and the relative heights of the two peaks. The oscillator strengths (the probability for a transition to occur) using the B3LYP method are, however, very low, and we thus chose to continue the study using the  $\omega$ B97XD functional, bearing the additional blue-shift in mind.

As a second system, we used the 12 amino acid cluster at the top of the barrel, as obtained from the homology model of UVR8 (referred to as the full cluster). Residue labeling and numbering is

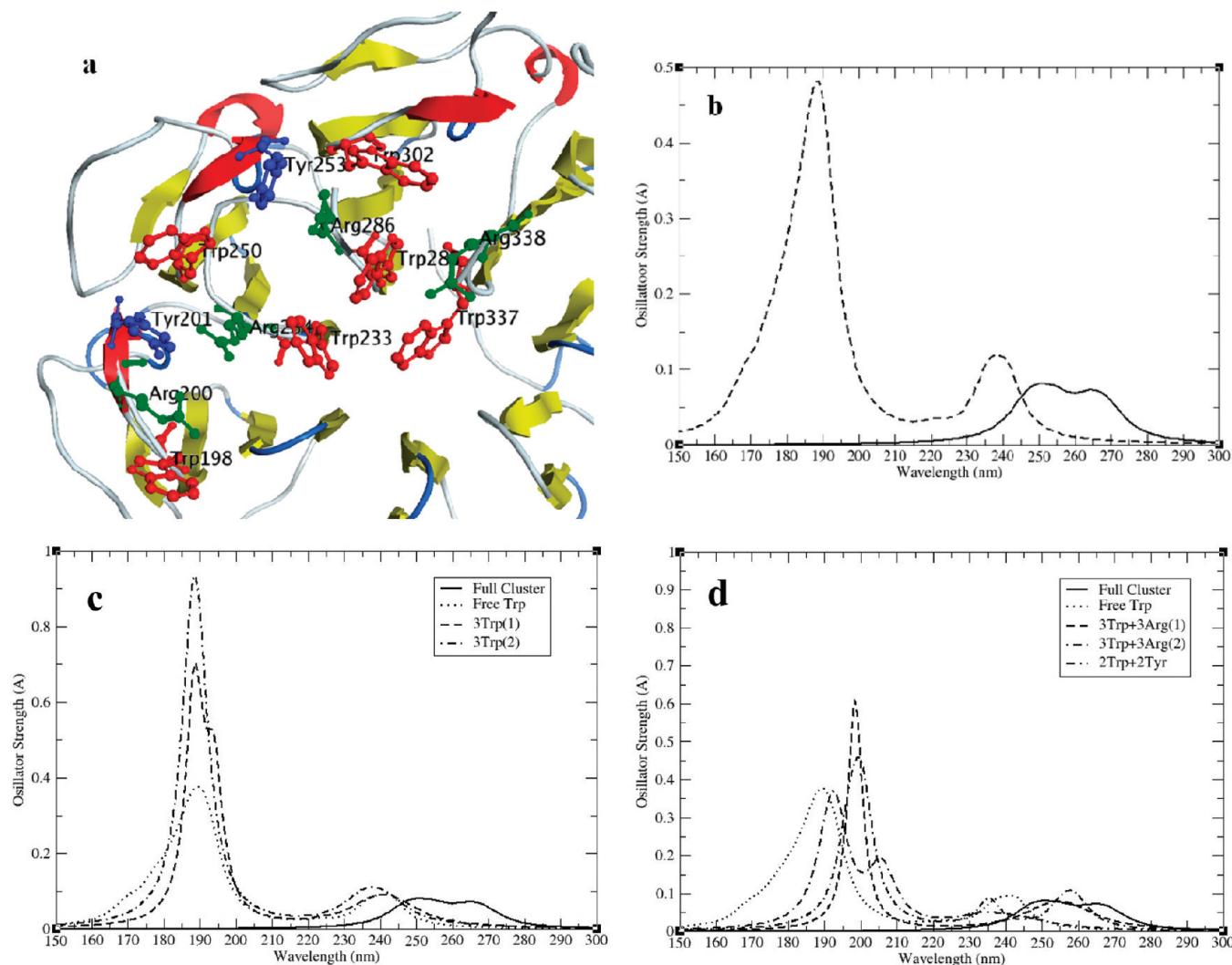
given in Figure 4a. The six Trp, four Arg, and two Tyr residues highlighted at the top of the protein were cut out and optimized at the B3LYP/6-31G level keeping the backbone alpha-carbons fixed in order to ensure an equivalent overall spatial distribution. The spectrum of the cluster was computed using the  $\omega$ B97XD/6-31G method (Figure 4b). Due to the large size of the system (298 atoms; 1660 basis functions), calculations using larger basis sets were not attempted.

Comparing the spectrum of the full cluster to that of the single Trp residue computed at the same level, we note a large red-shift of the first excitation upon inclusion of the full model (Figure 4b). In addition, the first single peak broadens and splits into two distinct peaks with approximately 18 nm separation. It should be noted that due to the very large number of orbitals involved in the full cluster, the number of excitations computed herein (200), albeit being very exhaustive from a computational point of view, was not sufficient to generate excitations all the way up to the large peak at around 200 nm; for the cluster model, the excitations computed are all found between 230 and 300 nm.

The first of the longer wavelength peaks in the spectrum of the large cluster, at 263–269 nm, primarily involves Trp302, Trp285, and Trp233 in the central part of the cluster (cf. Figure 4a). The corresponding orbital distributions are displayed in Figure S1 in the Supporting Information. The peak includes two main excitations. The electron transfer at 268 nm mainly involves excitations from orbital 540 to orbitals 558 and 559 (Table 1). Orbital 540 is for the most part distributed on Trp285 and Trp302 and is of  $\pi$  nature, and orbitals 558 and 559 are mostly localized on Trp285 and is of  $\pi^*$  nature. The excitation is hence a mix of vertical  $\pi \rightarrow \pi^*$  (within Trp285) and charge transfer (from Trp302( $\pi$ ) to Trp285( $\pi^*$ )). We also note that Arg234 is a minor acceptor of electron density in this excitation. The electronic excitation at 264 nm is mostly from orbital 536 to orbital 557. These are both located on Trp233 and is a pure  $\pi \rightarrow \pi^*$  excitation.

The second peak at longer wavelength, with a maximum at around 250 nm, includes the first vertical excitations (UV-absorptions) for the other Trps as listed in Table 2. The first excitations for Trp302 and Trp337 occur at approximately 240 nm (same as for a single Trp residue at this level; solid line of Figure 3b), but for Trp250 and Trp198 the interactions with surrounding residues in the full cluster lead to a slightly red-shifted excitations compared to the single Trp residue. Taking also into account the 35–40 nm blue-shift at the current level of theory, for a single Trp residue relative to experiments, the two peaks seen for the full cluster are predicted to appear at approximately 275 and 300 nm, in very good agreement with that seen in the action spectrum of UVR8 dependent UV-B stimulation of HY5 transcription in *A. thaliana* leaves.<sup>24</sup>

To explore if the full cluster is required to generate the first new peak seen in the full cluster (at  $\lambda = 269$  nm), a set of smaller clusters was also studied, based on the spatial distribution of the amino acids as in the full cluster: i) A small 3-amino acid cluster was constructed using Trp233, Trp285, and Trp302, positioned as in the full cluster model (see Figure 4a), labeled 3Trp(1). These were the tryptophans shown in the orbital analyses to be the ones responsible for the absorptions at the longest wavelengths, and we thus explore if they are able to generate the extra peak independently of interaction with the other moieties. ii) A second cluster of three Trp moieties, this time including the three centrally located Trp233, Trp285, and Trp337 (labeled 3Trp(2)). The difference over the above model is the internal



**Figure 4.** 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, and the two three-tryptophan clusters. d) Spectra of the full cluster, free Trp in solution, the two 3Trp+3Arg clusters, and the peripheral 2Trp+2Tyr cluster. All spectra were computed at the  $\omega$ B97XD/6-31G level of theory.

**Table 1. UV-Absorptions Included in the First Absorption Peak of the Full Cluster, at 263–269 nm**

wavelength	oscillator strength	wavelength	oscillator strength
268.7 nm	$f = 0.066$	263.7 nm	$f = 0.064$
Orbitals Involved in the Excitation			
537 → 558	-0.108	524 → 565	0.128
537 → 559	-0.146	534 → 557	-0.172
540 → 557	-0.128	536 → 557	0.624
540 → 558	-0.273	536 → 559	-0.138
540 → 559	-0.369		
540 → 561	-0.152		
544 → 558	0.204		
544 → 559	0.276		
544 → 561	0.115		

orientation of the tryptophans. iii) and iv) Two intermediate clusters including both sets of three Trps in the two models

above, along with the three arginines surrounding these, Arg234, Arg286, and Arg338. These two Arg-containing clusters were labeled 3Trp(1)+3Arg and 3Trp(2)+3Arg, respectively. The aim with these clusters was to explore whether the closest surroundings affect the absorption of the tryptophans. v) A peripheral cluster consisting of Tyr201, Trp250, Tyr253, and Trp302, all lying along the 'exterior' of the full cluster (cf. Figure 4a); labeled 2Trp+2Tyr. In this case the aim was to investigate whether the tyrosines influence the absorption of the tryptophans via  $\pi$ - $\pi$  interactions. In all cases above, the amino acids were excised and geometry optimized at the B3LYP/6-31+G(d,p) level keeping the  $\alpha$ -carbons fixed in the same way as for the full cluster. The spectra were subsequently computed at the  $\omega$ B97XD/6-31G level of theory (Figure 4c,d).

The first and second peak in the spectrum of the small 3-tryptophan clusters occur at essentially identical positions as for the free Trp amino acid, although the cluster environment helps to provide considerably higher probability of the transitions to occur (Figure 4c). The red-shifted additional peak seen in the full cluster is no longer present. The different UV absorption

**Table 2.** Components of the Second Peak of the Full Cluster, at around 250 nm

Trp250	Trp198	Trp302	Trp337
254.5 nm <i>f</i> = 0.081	249.1 nm <i>f</i> = 0.060	244.9 nm <i>f</i> = 0.035	243.4 nm <i>f</i> = 0.034

wavelengths for the three residues involved in the small cluster investigated at the  $\omega$ B97XD/6-31G(d,p) level were again based on  $\pi \rightarrow \pi^*$  excitations within the individual Trps. The large second peak, at 190 nm (200 nm at the  $\omega$ B97XD/6-31+G level), involves orbital contributions from all three Trp residues.

For the intermediate 3Trp+3Arg clusters the spectrum is red-shifted  $\sim$ 10 nm relative to the small cluster, but the excitations are all local  $\pi \rightarrow \pi^*$  excitations within the respective tryptophans (Figure 4d). Interestingly, the peak at 250–260 nm overlaps perfectly with the second of the two peaks in the full cluster, which indicates that the environment - in this case the three arginines - impacts on the position of the peaks. The local absorption of Trp285 occurs at 258 nm in 3Trp(1)+3Arg, and 255 nm in 3Trp(2)+3Arg, and for Trp233 at 255 nm in 3Trp(1)+3Arg, and 259 nm in 3Trp(2)+3Arg. Mutagenesis experiments have confirmed that Trp285 is important for the functionality of UVR8 and that Trp233 mutants do not dimerize.<sup>25</sup> For the 'peripheral' cluster 2Trp+2Tyr, a weak Trp absorption at 235–240 nm is seen followed by absorptions by the tyrosines at 205/206 nm and the Trps at 191/192 nm (Figure 4d). From the above analyses, we can thus conclude that the arginines surrounding the Trp cluster are crucial for modulating the absorptions and that Trp285 and Trp233 give rise to the absorption at the longest wavelength.

## ■ DISCUSSION AND CONCLUSIONS

UVR8 is known to have a crucial role in UV-B-dependent signaling in plant cells. Amino acid alignment of eight UVR8 proteins from plants or plant-related species was conducted, showing that 14 unique residues of the UV-absorbing amino acid Trp were in principle conserved in all eight species, in addition to a number of other amino acids including a large number of Tyrs and Args, leading to an overall amino acid identity score of 52%. Alignment of the *Arabidopsis thaliana* UVR8 with the two similar but UV-B-independent proteins HERC2 and RCC1 from *Homo sapiens* showed that 6 of these Trps are conserved in HERC2 but only two in RCC1; these are however tryptophans found in the interior of the proteins. Out of the seven Tyrs conserved in the plant species, as many as six were conserved in either HERC2 or RCC1, four of them in both proteins. In contrast, out of the 14 conserved Args in the plant species, only two were conserved in HERC2 and none in RCC1.

Based on the amino acid sequence of *A. thaliana* UVR8 protein, a homology model thereof was created. The homology model displays a 7-fold  $\beta$ -propeller arrangement, as seen for the RCC1 protein and similar to the WD40 repeat structure predicted for COP1. The structure shows an excess accumulation of aromatic residues (mainly the above-mentioned tryptophans) clustered on the 'top' surface of the propeller, intertwined by arginines.

The excitation spectra of different sized amino acid clusters were explored, based on these excess tryptophans; one containing six Trp, four Arg, and two Tyr residues; and several containing either three of the Trp residues found to be responsible for the excitations at the longest wavelengths of the full cluster with or without the neighboring arginines, or forming the 'peripheral' region of alternating Trp and Tyr residues. Initially,

the absorption spectrum of free tryptophan in solution was explored using different methods and basis sets, in order to benchmark against the experimental spectrum thereof. It is concluded that for the method later used on the full cluster, a blue-shift of 35–40 nm relative to experiments should be taken into account.

The absorption spectrum of the full cluster displayed a markedly different shape at longer wavelengths, than that of a single free Trp amino acid. Instead of a single peak at 240 nm ( $\omega$ B97XD/6-31G level), the spectrum is red-shifted and displays two peaks, at 263–269 nm and 250–255 nm. The first of these involves transitions between residues Trp302 and Trp285 (and Arg 234) and a highly localized  $\pi \rightarrow \pi^*$  excitation within Trp233. The peak centered at  $\lambda = 250$  nm involves the first vertical local excitation of tryptophans 250, 198, 302, and 337. Performing calculations on a smaller cluster containing only Trp233, 285, and 302 or only Trp233, 285, and 337 gives essentially identical spectrum to that of free Trp in solution. This implies that the protein environment - i.e. the clustered structure of the residues as mentioned above - strongly influences the absorption spectra of the tryptophans to longer wavelengths. This is manifested in the intermediate clusters, consisting of the above three Trps and the surrounding three Args 234, 286, and 338, resulting in a red-shift of the entire 3Trp spectrum to longer wavelengths, by approximately 10 nm. In particular Trp285 and Trp233 are essential for the absorption at the longest wavelengths (found at 255–260 nm in the above 3Trp-3Arg clusters). Given the 35–40 nm difference between the calculated spectrum at the  $\omega$ B97XD/6-31G level and the experimental spectrum of free Trp in solution (Figure 1b,c), it is likely that the theoretically calculated peaks at 250 and 263–269 nm for the full cluster accounts for the experimental peaks at 280 and 300 nm seen for UVR8. Of particular importance is the role of the 300 nm peak in the action spectrum of HY5 gene expression<sup>24</sup> and other UV-B-regulated genes.<sup>20</sup> UV absorption by these amino acid residues would thus be the initiators of the transcription of these genes, i.e. they constitute the chromophore controlling gene expression. In addition, this implies that the UVR8 protein itself is the more important of at least two likely photoreceptors, regulating the bulk of the plant genes induced by UV-B radiation.

Strengthening this notion is the fact that neither of these Trps are among the 6 (out of 14) Trps that are shared by both UVR8 and/or the non-UV-B-active human proteins HERC2 and RCC1, indicating that evolution of UV-B absorption and signaling is dependent on the ensemble of the particular Trps that are not found in the human proteins. In contrast, as many as six of the seven conserved Tyrs in the plant UVR8s are also found in HERC2 and RCC1, indicating that the Tyrs play a minimal role in the absorption of UVR8 in the UV-B part of the spectrum, but rather are important for the formation of the core of the protein fold. Finally, only 2 of the 14 Args conserved in the plant UVR8 proteins are also found in HERC2 (none in RCC1), again inferring a larger role for these Args in UV-B activation than for the Tyrs, and which we also show for the 3Trp+3Arg(1) and 3Trp+3Arg(1) clusters (Figure 3d).

In their recent report, Rizzini et al.<sup>25</sup> also discuss the importance of the UVR8 Trps for function of the protein Trps285, 337, and 233. Mutations in these amino acids e.g. abolish

dimerization of UVR8 and Trp285 seems to play a crucial role in interaction with COP1. Therefore, their data corroborate the importance of these Trps as shown by us in this study, where we especially have emphasized the absorption properties of the chromophore. To these findings we have been able to add information about the importance of the Arg in the particular absorption bands at 280 and 300 that are characteristic for the UV-B receptor. In addition, the importance of especially Trp285<sup>25</sup> strengthens our discussion below of a coordinating "special" Trp that funnels the signal about absorption of UV-B by UVR8 onward by activating the signaling pathway, including the event of complexation between UVR8 and COP1.

Thus, the UV-B PR converts a physical signal, in this case the absorption of UV quanta, to chemical changes that in turn, through the actions of the PR and other signaling components, result in biological signals. In UVR8, the Trps can each be absorbing the UV quanta. However, the chemical signal or signals and the resulting events (UVR8 import into the nucleus,<sup>23</sup> UVR8 monomerization and COP1 interaction,<sup>18,25</sup> and induction of gene expression<sup>11</sup>) can either be transduced from each of the Trps or, more likely, coordinated through one of the Trps (or Arg) that therefore would serve as a collection point of the information of quantal absorbance. In the latter case, the other Trps would function as UV-B antennae. Further analysis of each of these Trps (and Arg) would resolve which of these two simplified cases would be most closely resembling the situation *in vivo*. Indeed, if all six Trps would have the same importance for signal transduction, deletion of each of them would diminish the signal amplitude by approximately 17%. On the other hand, if one of the Trps is the collection point of the information, e.g. Trp285,<sup>25</sup> removal of this central amino acid residue would completely terminate the signal. Of course, other Trp-dependent signaling permutations within UVR8 can also be envisaged. Therefore, careful mutant analysis, on a larger scale than in ref 25, based on the data presented in this paper would be needed to resolve the roles of the individual tryptophans (and arginines) in the cluster (structural/π-stacking interaction, UV-absorption, signal modulation, etc.).

## ■ ASSOCIATED CONTENT

**Supporting Information.** Homology model settings are found in Table S1, and molecular orbitals of full cluster are found in Figure S1. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

### Corresponding Author

\*Phone: +46-19 303603. Fax: +46-19 303566. E-mail: ake.strid@oru.se.

## ■ ACKNOWLEDGMENT

The National University of Ireland, Galway is gratefully acknowledged for financial support (L.A.E.), as is the Faculty of Business, Science and Technology at Örebro University (Å.S. and E.G.).

## ■ REFERENCES

- Bornman, J. F. Target sites of UV-B radiation in photosynthesis of higher plants. *J. Photochem. Photobiol. B: Biol.* **1989**, *4*, 145–158.
- Strid, Å.; Chow, W. S.; Anderson, J. M. UV-B damage and protection at the molecular level in plants. *Photosynth. Res.* **1994**, *39*, 475–489.
- Brosché, M.; Fant, C.; Bergkvist, S. W.; Strid, Å.; Svensk, A.; Olsson, O.; Strid, Å. Molecular markers for UV-B stress in plants: alteration of the expression of four classes of genes in *Pisum sativum* and the formation of high molecular mass RNA adducts. *Biochim. Biophys. Acta* **1999**, *1447*, 185–198.
- A-H-Mackerness, S. Plant responses to UV-B (UV-B: 280–320 nm) stress: What are the key regulators?. *Plant Growth Regul.* **2000**, *32*, 27–39.
- Sävenstrand, H.; Brosché, M.; Strid, Å. Regulation of gene expression by low levels of ultraviolet-B radiation in *Pisum sativum*: Isolation of novel genes by suppression subtractive hybridisation. *Plant Cell Physiol.* **2002**, *43*, 402–410.
- Brosché, M.; Schuler, M. A.; Kalbina, I.; Connor, L.; Strid, Å. Gene regulation by low level UV-B radiation: identification by DNA array analysis. *Photochem. Photobiol. Sci.* **2002**, *1*, 656–664.
- Brosché, M.; Strid, Å. Molecular events following perception of ultraviolet-B radiation by plants: UV-B induced signal transduction pathways and changes in gene expression. *Physiol. Plant* **2003**, *117*, 1–10.
- Sävenstrand, H.; Brosché, M.; Strid, Å. *Arabidopsis* brassinosteroid mutants are defective in UV-B-regulated defence gene expression. *Plant Physiol. Biochem.* **2004**, *42*, 687–694.
- Ulm, R.; Baumann, A.; Oravecz, A.; Máté, Z.; Ádám, É.; Oakeley, E. J.; Schäfer, E.; Nagy, F. Genome-wide analysis of gene expression reveals function of the bZIP transcription factor HY5 in the UV-B response of *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 1397–1402.
- Kalbina, I.; Strid, Å. Supplementary ultraviolet-B irradiation reveals differences in stress responses between *Arabidopsis thaliana* ecotypes. *Plant Cell Environ.* **2006**, *29*, 754–763.
- Jenkins, G. I. Signal transduction in responses to UV-B Radiation. *Annu. Rev. Plant Biol.* **2009**, *60*, 407–431.
- Kliebenstein, D. J.; Lim, J. E.; Landry, L. G.; Last, R. L. *Arabidopsis* UVR8 regulates UV-B signal transduction and tolerance and contains sequence similarity to human *Regulator of Chromatin Condensation 1*. *Plant Physiol.* **2002**, *130*, 234–243.
- Brown, B. A.; Cloix, C.; Jiang, G. H.; Kaiserli, E.; Herzyk, P.; Kliebenstein, D. J.; Jenkins, G. I. A UV-B-specific signaling component orchestrates plant UV protection. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 18225–18230.
- Oravecz, A.; Baumann, A.; Máté, Z.; Brzezinska, A.; Molinier, J.; Oakeley, E. J.; Ádám, É.; Schäfer, E.; Nagy, F.; Ulm, R. Constitutively Photomorphogenic 1 is required for the UV-B response in *Arabidopsis*. *Plant Cell* **2006**, *18*, 1975–1990.
- Stacey, M. G.; Kopp, O. R.; Kim, T.-H.; von Arnim, A. G. Modular domain structure of *Arabidopsis thaliana* COP1. reconstitution of activity by fragment complementation and mutational analysis of a nuclear localization signal in planta. *Plant Physiol.* **2000**, *124*, 979–990.
- Chen, M.; Chory, J.; Fankhauser, C. Light signal transduction in higher plants. *Annu. Rev. Genet.* **2004**, *38*, 87–117.
- Yi, C.; Deng, X. W. COP1 - from plant photomorphogenesis to mammalian tumorigenesis. *Trends Cell Biol.* **2005**, *15*, 618–625.
- Favory, J. J.; Stec, A.; Gruber, H.; Rizzini, L.; Oravecz, A.; Funk, M.; Albert, A.; Cloix, C.; Jenkins, G. I.; Oakeley, E. J.; Seidlitz, H. K.; Nagy, F.; Ulm, R. Interaction of COP1 and UVR8 regulates UV-B-induced photomorphogenesis and stress acclimation in *Arabidopsis*. *EMBO J.* **2009**, *28*, 591–601.
- Brown, B. A.; Jenkins, G. I. UV-B signaling pathways with different fluence-rate response profiles are distinguished in mature *Arabidopsis* leaf tissue by requirement for UVR8, HY5, and HYH. *Plant Physiol.* **2008**, *146*, 576–588.
- Kalbina, I.; Li, S.; Kalbin, G.; Björn, L. O.; Strid, Å. Two separate UV-B radiation wavelength regions control expression of different molecular markers in *Arabidopsis thaliana*. *Funct. Plant Biol.* **2008**, *35*, 222–227.
- Wargent, J. J.; Geges, V. C.; Jenkins, G. I.; Doonan, J. H.; Paul, N. D. UVR8 in *Arabidopsis thaliana* regulates multiple aspects of cellular

differentiation during leaf development in response to ultraviolet B radiation. *New Phytol.* **2009**, *183*, 315–326.

(22) Cloix, C.; Jenkins, G. I. Interaction of the *Arabidopsis* UV-B-specific signaling component UVR8 with chromatin. *Mol. Plant* **2008**, *1*, 118–128.

(23) Kaiserli, E.; Jenkins, G. I. UV-B promotes rapid nuclear translocation of the UV-B-specific signaling component UVR8 and activates its function in the nucleus. *Plant Cell* **2007**, *19*, 2662–2673.

(24) Brown, B. A.; Headland, L. R.; Jenkins, G. I. UV-B action spectrum for UVR8-mediated HYS transcript accumulation in *Arabidopsis*. *Photochem. Photobiol.* **2009**, *85*, 1147–1155.

(25) Rizzini, L.; Favery, J.-J.; Cloix, C.; Faggionato, D.; O'Hara, A.; Kaiserli, E.; Baumeister, R.; Schäfer, E.; Nagy, F.; Jenkins, G. I.; Ulm, R. Perception of UV-B by the *Arabidopsis* UVR8 Protein. *Science* **2011**, *332*, 103–106.

(26) Krieger, E.; Koraimann, G.; Vriend, G. Increasing the precision of comparative models with YASARA NOVA - a self-parameterizing force field. *Proteins* **2002**, *47*, 393–402.

(27) Walker, J. R.; Qiu, L.; Vesterberg, A.; Weigelt, J.; Bountra, C.; Arrowsmith, C. H.; Edwards, A. M.; Bochkarev, A.; Dhe-Paganon, S. Structure of the third RLD domain of HERC2. *RCSB Protein Data Bank*. 2009, doi:10.2210/pdb3kci/pdb. <http://www.pdb.org/pdb/explore.do?structureId=3KCI> (accessed Nov 3, 2009).

(28) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(29) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *J. Phys. Chem.* **1994**, *98*, 11623–11627.

(30) Chai, J. D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.

(31) Perpète, E. A.; Wathelet, V.; Preat, J.; Lambert, C.; Jacquemin, D. Toward a Theoretical Quantitative Estimation of the  $\lambda_{\max}$  of Anthraquinones-Based Dyes. *J. Chem. Theory Comput.* **2006**, *2*, 434–440.

(32) Jacquemin, D.; Wathelet, V.; Perpète, E. A.; Adamo, C. Extensive TD-DFT Benchmark: Singlet-Excited States of Organic Molecules. *J. Chem. Theory Comput.* **2009**, *5*, 2420–2435.

(33) Chai, J. D.; Head-Gordon, M. Systematic optimization of long-range corrected hybrid density functionals. *J. Chem. Phys.* **2008**, *128*, 084106. <http://jcp.aip.org/resource/1/jcpsa6/v128/i8> (accessed Feb 27, 2008).

(34) Tian, B.; Eriksson, E. S. E.; Eriksson, L. A. Can Range-Separated and Hybrid DFT Functionals Predict Low-Lying Excitations? A Tookad Case Study. *J. Chem. Theory Comput.* **2010**, *6*, 2086–2094.

(35) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A. Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09. Revision A.02; Gaussian, Inc.: Wallingford, CT, 2009.

(36) Yu, A. Z.; Yu, A. B.; Dadayan, A. K.; Myasoedov, N. F. Isotopic effects in the electronic spectra of tryptophan. *Amino Acids* **2006**, *31*, 403–407.