Photochemical & Photobiological Sciences (2021) 20:559–570
https://doi.org/10.1007/s43630-021-00039-5

ORIGINAL PAPERS

Unusual shift in the visible absorption spectrum of an active ctenophore photoprotein elucidated by time-dependent density functional theory

Felix N. Tomilin[1,2,3] · Anastasia V. Rogova[2] · Ludmila P. Burakova[2,4] · Olga N. Tchaikovskaya[3] · Pavel V. Avramov[5] · Dmitri G. Fedorov[6] · Eugene S. Vysotski[4]

Received: 19 December 2020 / Accepted: 29 March 2021 / Published online: 8 April 2021

Abstract

Active hydromedusan and ctenophore Ca^{2+}-regulated photoproteins form complexes consisting of apoprotein and strongly non-covalently bound 2-hydroperoxycoelenterazine (an oxygenated intermediate of coelenterazine). Whereas the absorption maximum of hydromedusan photoproteins is at 460–470 nm, ctenophore photoproteins absorb at 437 nm. Finding out a physical reason for this blue shift is the main objective of this work, and, to achieve it, the whole structure of the protein–substrate complex was optimized using a linear scaling quantum–mechanical method. Electronic excitations pertinent to the spectra of the 2-hydroperoxy adduct of coelenterazine were simulated with time-dependent density functional theory. The dihedral angle of 60° of the 6-(*p*-hydroxy)-phenyl group relative to the imidazopyrazinone core of 2-hydroperoxycoelenterazine molecule was found to be the key factor determining the absorption of ctenophore photoproteins at 437 nm. The residues relevant to binding of the substrate and its adopting the particular rotation were also identified.

1 Introduction

Bioluminescence is widespread among marine inhabitants [1, 2]. Many of these organisms generate light using coelenterazine (CTZ) as a substrate in the reactions catalyzed by various bioluminescent proteins [3, 4]. Among those, Ca^{2+}-regulated photoproteins represent a unique class of coelenterazine-utilizing proteins. A photoprotein forms an enzyme–substrate complex consisting of a single-chain globular protein within an internal cavity of which the peroxy-substituted coelenterazine, 2-hydroperoxycoelenterazine (2HP-CTZ, Fig. 1a), is noncovalently bound [3, 5].

In contrast to the common coelenterazine-dependent luciferases [4] in which enzymes catalyze the oxidative decarboxylation reaction of coelenterazine, photoprotein bioluminescence does not require molecular oxygen. Only the binding of calcium ions to a photoprotein is necessary to trigger the decarboxylation reaction of the 2-hydroperoxy adduct of coelenterazine yielding a protein-bound product, coelenteramide (CLM), in an excited state [6, 7]. Relaxation of the excited CLM to the ground state is accompanied by light emission with the spectral maximum in the range of 465–495 nm [8]. However, oxygen is still involved in the photoprotein bioluminescence, because the active

✉ Dmitri G. Fedorov
d.g.fedorov@aist.go.jp

✉ Eugene S. Vysotski
eugene.vysotski@gmail.com

1 Kirensky Institute of Physics SB RAS, Federal Research Center "Krasnoyarsk Science Center SB RAS", Akademgorodok 50/38, Krasnoyarsk 660036, Russia

2 Siberian Federal University, Svobodny 79 pr., Krasnoyarsk 660041, Russia

3 National Research Tomsk State University, Lenin Avenue 36, Tomsk 634050, Russia

4 Photobiology Laboratory, Institute of Biophysics SB RAS, Federal Research Center "Krasnoyarsk Science Center SB RAS", Akademgorodok 50/50, Krasnoyarsk 660036, Russia

5 Kyungpook National University, 80 Daehakro, Bukgu, Daegu 41566, South Korea

6 Research Center for Computational Design of Advanced Functional Materials (CD-FMat), National Institute of Advanced Industrial Science and Technology (AIST), Central 2, Umezono 1-1-1, Tsukuba 305-8568, Japan

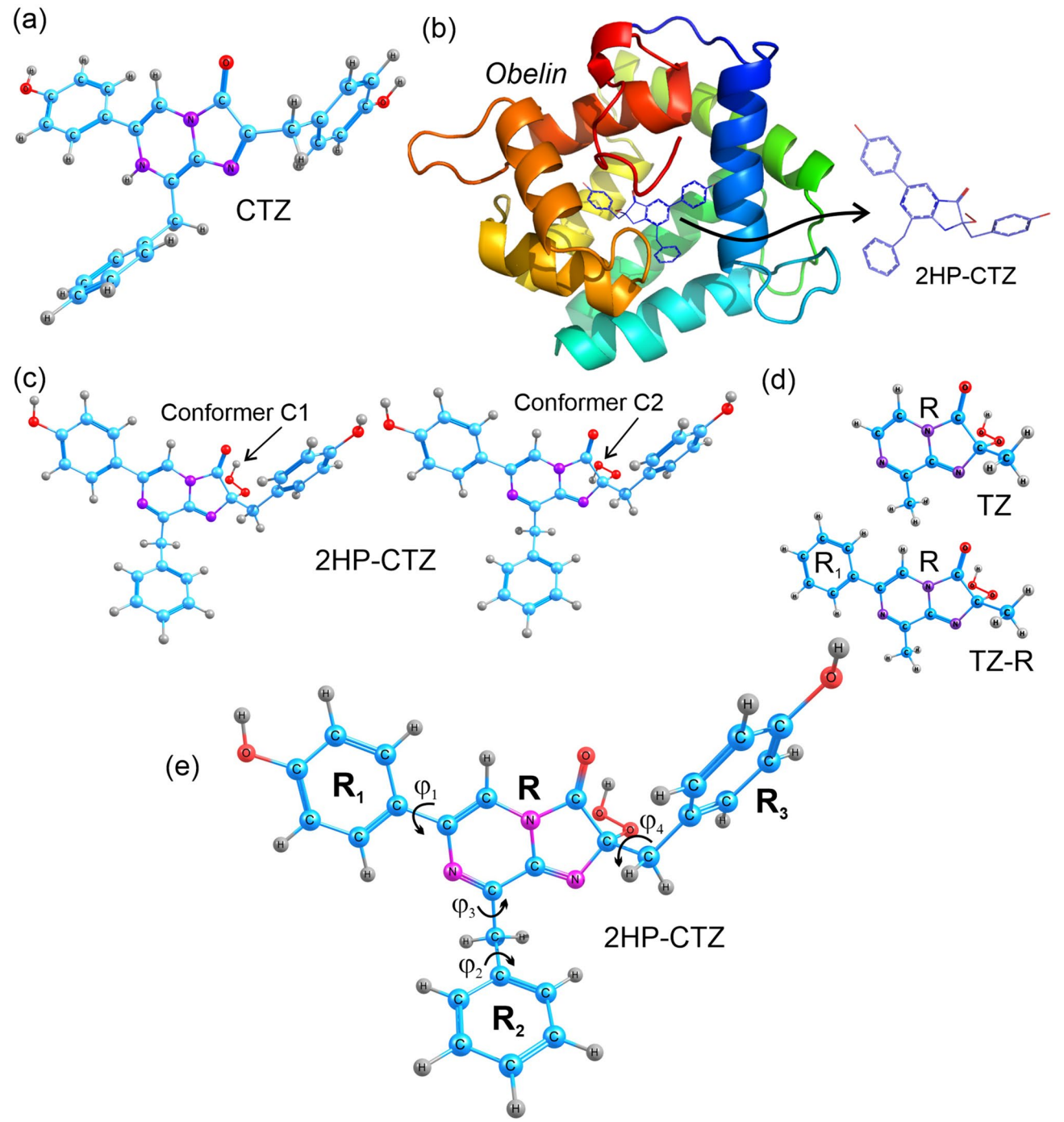

Fig. 1 Molecular structures used to calculate optical absorption spectra: **a** coelenterazine (CTZ); **b** 2-hydroperoxycoelenterazine (2HP-CTZ) in conformation corresponding to that in obelin crystal structure; **c** C1 and C2 conformers of the 2HP-CLZ; **d** TZ and TZ-R artificial molecules used to determine the chromophore group; **e** angles φ_i describing the rotation of the functional groups R_i relative to the chromophore R. The C, O, H, and N atoms are colored in blue, red, grey, and violet, respectively

photoprotein is formed from apoprotein and coelenterazine under Ca^{2+}-free conditions only in the presence of O_2 [9, 10].

The main driving force for studying bioluminescent proteins is the use of the Ca^{2+}-regulated photoproteins as an analytical tool. The major application of photoproteins primarily lies in their capability to emit light upon binding of calcium ions. Now, they are widely applied to measure the intracellular concentration of Ca^{2+} under steady-state

conditions and to study the calcium transients in the cytoplasm of various cells and cell organeles [11, 12].

Among bioluminescent proteins of this type, only photoproteins responsible for the light emission of a variety of marine hydromedusae are well studied [3, 5]. Cloning and sequence analyses have been carried out for the photoproteins, namely, aequorin [13–15], clytin [16–19], mitrocomin [20, 21], and obelin [22–24]. The analysis of amino acid sequences has shown a high degree of identity (~70–60%) of these proteins and the presence of three calcium-binding consensus sequences characteristic of the EF-hand Ca^{2+}-binding proteins [5, 25]. Based on this, Ca^{2+}-regulated photoproteins are assigned to a large family of the EF-hand Ca^{2+}-binding proteins [26] that includes the most numerous and extensively studied members playing an important role in the regulation of multiple processes in various types of cells.

The structures of several Ca^{2+}-regulated photoproteins in different conformational states have been determined for aequorin, obelin, clytin, mitrocomin bound with 2HP-CTZ [21, 27–30], obelin bound with CLM [31, 32], and apo-aequorin and apo-obelin bound with calcium ions [33]. For the photoprotein structure not much affected by the light-emitting reaction [32], it was proposed that in the EF-hand protein family, Ca^{2+}-regulated photoproteins belong rather to the category of Ca^{2+} signal modulators, such as parvalbumin, than to Ca^{2+} sensors, of which calmodulin is the typical and best-known representative [34]. The comparison of spatial structures of the photoprotein before and after bioluminescence reaction as well as comprehensive mutagenesis studies of the residues forming their inner cavity [35] shed light into the role of certain residues of a substrate-binding pocket in the catalytic reaction and emitter formation [8, 36].

The bright bioluminescence of ctenophores also occurs due to Ca^{2+}-regulated photoproteins [3]. Although many of their properties are very similar to those of hydromedusan photoproteins, there are still important differences [37]. The ctenophore photoproteins are extremely sensitive to light; they lose bioluminescence activity upon exposure to light over the entire spectral range [38, 39], including visible region. The bioluminescent capability of these photoproteins can be restored only by their incubation with coelenterazine in the dark at an alkaline pH in the presence of oxygen [39].

Cloning of cDNAs encoding these photoproteins [40–42] has revealed considerable differences in amino acid sequences of ctenophore and hydromedusan photoproteins [41]. At the same time, similar to hydromedusan photoproteins, ctenophore proteins also contain three canonical EF-hand Ca^{2+}-binding sites [41]. The structure of any ctenophore photoprotein bound to oxygenated coelenterazine has not yet been determined, and only the tertiary structures of apo-berovin bound with Ca^{2+} or Mg^{2+} ions [43, 44] and apo-mnemiopsin loaded by Cd^{2+} ions [45] are available now. It is important to note that the spatial structures reveal a common scaffold in ctenophore and hydromedusan photoproteins despite their low amino acid sequence identity [37].

Even though the spatial structures of hydromedusan and ctenophore photoproteins are very similar in shape and geometry of their substrate-binding cavities, their interiors feature large differences. The key amino acid residues, considered to be important for stabilizing the 2-hydroperoxy adduct of coelenterazine and for emitter formation in hydromedusan photoproteins [8, 36], in ctenophore photoproteins are replaced by the residues with very different side chains. This suggests that the amino acids involved in the light emission of ctenophore photoproteins completely differ from those of hydromedusan photoproteins [46].

To identify the key residues that may be involved in bioluminescence reactions, two structural models of ctenophore photoprotein with an embedded 2-hydroperoxy adduct of coelenterazine have been built. Because the structure of these photoproteins apparently undergoes insignificant changes during bioluminescence reactions [37], similar to hydromedusan photoproteins, this approach is plausible. The main difference in these models is the manner of stabilization of the oxygenated coelenterazine within the photoprotein cavity. In the first structural model for berovin, the oxygenated coelenterazine is bound in the form of a 2-peroxy anion adduct which is stabilized via the Coulomb interaction with the positively charged Arg41 guanidinium group that is paired with OH group of Tyr204 [46]. Another structural model has 2-hydroperoxycoelenterazine as a coelenterazine derivative bound within substrate-binding site of ctenophore photoprotein similar to hydromedusan photoproteins [47]. According to this model, the OH group of Tyr204 is close enough to the 2-hydroperoxy group of coelenterazine to form a hydrogen bond, while the guanidinium group of Arg41 forms a hydrogen bond with the OH group of Tyr204 and the carboxylic group of Asp158. Based on these findings, it has been inferred [47] that the Tyr204–Arg41–Asp158 triad carries out the function similar to the Trp–His–Tyr triad in hydromedusan photoproteins, where the tyrosine OH group stabilizes the 2-hydroperoxy group of coelenterazine [8]. However, there is some contradiction with the results obtained for a berovin mutant with a replacement of Tyr204 by Phe that completely retains its bioluminescent activity relative to the wild type berovin [43, 46]. Since this Tyr interacts with the 2-hydroperoxy group, its substitution has to affect the bioluminescence. The substitution of Tyr to Phe in hydromedusan photoproteins, for example, which performs a similar function results in a significant decrease of the bioluminescence activity of mutated photoproteins [48]. Both models look quite plausible, but both do not clarify features of ctenophore photoproteins such as a sensitivity to light of the active photoprotein complex and absorption spectrum in the visible region.

The absorption spectra of ctenophore photoproteins differ from hydromedusan photoproteins spectra. While hydromedusan photoproteins absorb light with a maximum at 460–470 nm [3, 10], the absorption maximum of ctenophore photoproteins is at 437 nm [39] which is very close to that of coelenterazine in methanol (λ_{max} = 435 nm [3]) (Fig. S1). Considering that the photoprotein absorption in the visible region is determined by the bound 2-hydroperoxy adduct of coelenterazine [3, 5], this distinction can be attributed to the difference in the structure of the bound substrate and its amino acid environment within the substrate cavity.

In this study, the effects of geometry, solvents, and substituents on the absorption spectral properties of coelenterazine and its 2-hydroperoxy adduct, as well as the impact of amino acid environment on the absorption spectrum of 2-hydroperoxycoelenterazine embedded into the internal cavity of berovin, are explored with time-dependent density functional theory (TD-DFT).

2 Experimental

2.1 Computational details

The calculation scheme to investigate the optical properties of the active photoprotein substrate is shown in Fig. 1. To select an appropriate method, the electronic excitation for the main peak in the absorption spectrum of a CTZ molecule was calculated using a set of functionals of DFT and two solvation models (Fig. 1a; Table 1). The CTZ was chosen as its absorption spectrum in methanol is known (λ_{max} in the visible region is at 435 nm) [3]. Then, the selected calculation method was applied to investigate the optical properties of a 2HP-CTZ molecule.

The conformation of 2HP-CTZ in the crystal structure of obelin (PDB code 1QV0) was taken as a starting point to optimize the geometry (Fig. 1b). The hydrogen of the peroxide group in 2HP-CTZ is capable of adopting *trans* (C1) and *cis* (C2) conformations (Fig. 1c), which can have different spectral properties, because the peroxide is close to the chromophore (Table 2) and can form a hydrogen bond with it. Therefore, an effect of the peroxide conformations of the peroxide group on the optical properties of the 2HP-CTZ molecule was studied.

The peroxide conformation which gives the best overlap of calculated and experimental absorption spectra of 2HP-CTZ in the visible range (λ_{max} = 460–470 nm [3, 10], Table 2) was selected. For the group of atoms playing the key role in the electronic excitation (chromophore group) to be identified, the TZ and TZ-R model molecules (Fig. 1d, Fig. S2) were designed, and the influence of the rotation of each substituent on the spectral properties of the 2HP-CTZ molecule was studied (Fig. 1e).

TD-DFT calculations for vertical excitations were performed to simulate the CTZ absorption spectra in methanol using SMD for describing solvent effects [49] with the same dielectric constant ε in both ground and excited states. To find the best DFT functional, the spectral characteristics of CTZ were calculated using CAM-B3LYP, PBE0, M06-2X, and M08-SO functionals (Table 1). Some calculations were repeated with a larger basis set aug-cc-pVDZ including the diffuse functions. B3LYP/SMD/cc-pVDZ gives the best agreement with the experimentally determined absorption maximum (λ_{max} = 435 nm [3]). Taking into account that the results for the larger basis set are essentially the same as

Table 2 Absorption wavelengths λ_{abs} (nm) and oscillator strengths f_{abs} (arb. units) for the $S_0 \rightarrow S_1$ excitations in 2HP-CTZ conformers calculated at B3LYP/SMD/cc-pVDZ level of theory

Conformer[a]	Gas phase λ_{abs} (f_{abs})	Methanol λ_{abs} (f_{abs})	Water λ_{abs} (f_{abs})
C1	473 (0.021)[b]	485 (0.032)	486 (0.032)
C2	480 (0.022)	493 (0.027)	494 (0.028)

[a]See Fig. 1c

[b]Experimental absorption maximum of 2HP-CTZ bound within obelin is at 460–470 nm [3, 10]

Table 1 $S_0 \rightarrow S_1$ excitations in CTZ (Fig. 1b) in methanol calculated with several DFT functionals

	B3LYP		CAM-B3LYP		PBE0		M06-2X		M08-SO	
	PCM	SMD	PCM	SMD	PCM	SMD	PCM	SMD	PCM	SMD
λ_{abs}	430 (447)	435 (439)	405 (407)	407 (403)	442	434	415	409	415	410
E	2.76 (2.77)	2.77 (2.82)	3.06 (3.05)	2.84 (2.84)	2.81	2.85	2.99	3.03	2.99	3.03
f	0.171 (0.191)	0.179 (0.218)	0.262 (0.291)	0.179 (0.323)	0.168	0.179	0.223	0.255	0.238	0.271
Λ	0.751 (0.748)	0.750 (0.752)	0.744 (0.738)	0.745 (0.743)	0.748	0.751	0.742	0.747	0.739	0.742

λ_{abs} (nm) is the absorption wavelength, f (arb. units) is the oscillator strength, E is the absorption energy (eV), and Λ is a parameter describing the physical characteristics of the excitation [59]. Calculations were performed with cc-pVDZ basis set (aug-cc-pVDZ in parentheses). PCM and SMD were selected as solvent models to account solvation effect in methanol. The experimental absorption maximum of CTZ in methanol is 435 nm

for cc-pVDZ, the latter was chosen (Table 1), because it is smaller.

The molecular structure of 2HP-CTZ was optimized using DFT at the B3LYP/cc-pVDZ level implemented in GAMESS [50, 51]. The electronic excitations for absorption spectra were calculated with TD-DFT at the optimized ground singlet state. For 2HP-CTZ, calculations were performed in the gas phase ($\varepsilon = 0$), methanol ($\varepsilon = 33$), and water ($\varepsilon = 78$). The initial structure of 2HP-CTZ was taken from the crystal structure of obelin (PDB code 1QV0, Fig. 1b). The results of calculations are summarized in Tables 1 and 2. As an outcome of these validation tests, B3LYP/cc-pVDZ and conformer C1 were chosen.

By closely inspecting the structures of photoproteins, one can speculate that the spectral properties may be related to several dihedral angles. To study the dependence of properties on these angles, the structures of 2HP-CTZ was reoptimized while fixing the angles of 6-(*p*-hydroxy)-phenyl (R_1), 8-benzyl (R_2), and 2-(*p*-hydroxy)-benzyl (R_3) substituents around the pyrazine group (R) to a range of desired values and optimizing other degrees of freedom at the level of B3LYP/cc-pVDZ (Fig. 1e).

Since the molecular structure of any of the ctenophore photoproteins with a bound 2-hydroperoxy adduct of coelenterazine is not available, the structural model of berovin without calcium ions was generated by the I-TASSER server [52–54] which automatically selects the appropriate templates from the Protein Data Bank. No restrictions were applied for selection of appropriate templates. In total, five structures were selected by the server in the Protein Data Bank [mitrocomin (PDB 4NQG), apo-berovin bound with Ca^{2+} (PDB 4MN0), *i*-aequorin (PDB 1UHI], calcium-dependent protein kinase 1 in complex with bumped kinase inhibitor NM-PP1 (PDB 3I7B), and EF-hand calcium-binding protein calerythrin (PDB 1NYA)) to generate five structural models. The best I-TASSER model (C-score of − 0.98, TM-score of 0.59 ± 0.14, and RMSD of 7.5 ± 4.3 Å) (Fig. 2a) revealed a high similarity with the structures of hydromedusan photoproteins bound with 2-hydroperoxycoelenterazine [obelin from *Obelia longissima* (PDB 1EL4), TM-score, 0.852, RMSD, 1.69 Å; mitrocomin (PDB 4NQG), TM-score, 0.851, RMSD, 1.14 Å; *i*-aequorin (PDB 1UHI), TM-score, 0.849, RMSD, 1.40 Å; obelin from *Obelia geniculata* (PDB 1JF0), TM-score, 0.844, RMSD, 1.27 Å] as well as with apo-aequorin (PDB 1SL8) and apo-berovin (PDB 4MN0) loaded with Ca^{2+} (TM-score, 0.738, RMSD, 1.82 Å and TM-score, 0.719, RMSD, 2.78 Å, respectively) was used to embed a 2HP-CLZ molecule into the internal cavity of berovin by AutoDock4 software [55] (Fig. 2b).

The rotational angles of the side chains of the residues Arg41 and Tyr204 in the substrate-binding cavity were slightly adjusted, because the side chains of these residues may be involved in the stabilization of the hydroperoxy group of the oxygenated coelenterazine as suggested by a study of berovin mutants with substitution of Arg41 and Tyr204 [46] (Fig. 2c). The molecular structure of berovin in water was optimized (energy-minimized with the threshold of 10^{-4} Hartree/Bohr) by the fragment molecular orbital method (FMO) [56, 57] in GAMESS at the level of the two-body expansion for the third-order density-functional tight-binding (FMO2-DFTB3) using 3ob-3-1 parameters [58], combined with the conductor polarizable continuum model of solvation (C-PCM) and Grimme's D3(BJ) dispersion correction. FMO-DFTB/PCM is a fast and accurate parametrized quantum–mechanical method suitable to optimize proteins [60]. The protein (Fig. 2c) was divided into one residue per fragment in FMO calculations.

From the located structure of the complex, the 2HP-CTZ substrate was extracted and its electronic excitation was computed with TD-DFT (B3LYP/cc-pVDZ) (Fig. 2d). The cavity is hydrophobic, thus, $\varepsilon = 0$ (as in the gas phase) can be used to compare to experiment; among the two conformers in Table 2, C1 matched the experiment better; therefore, it was chosen for further simulations.

To check the effect of the embedding, an FMO1-TD-B3LYP/6-31*/PCM calculation of a 2HP-CTZ conformer was performed. In FMO1-TDDFT, a polarizable embedding from the protein is added based on the electron density of fragments. The excitation energy with and without the embedding is 433 and 437 nm, respectively. Both match well the absorption maximum ($\lambda_{max} = 435$ nm) of the active berovin [41]. Because the embedding effect is small, the rest of the TDDFT calculations were performed without an embedding.

3 Results and discussion

3.1 Rotation of different substituents of 2HP-CTZ

To reveal the orientation effect of the R_1 and R groups (Fig. 3a, b) on the absorption spectra, the R_1 group was rotated by the angle φ_1 around the R_1–R single bond with a step of 5° relative to the equilibrium geometry (Fig. 3a). The dihedral angle φ_1 was fixed by freezing the coordinates of carbon C1, C2, C3, and C4 atoms, while the rest ones were optimized (Fig. 3a, b). The same approach was used to study the effect of rotations of the R_2 and R_3 groups relative to the chromophore R. As R_2 is bound to R by two single bonds, two rotations were carried out separately (Fig. 3c–f). Only one rotation was relevant for R_3 (Fig. 3g, h). All rotations were treated independently by varying a single angle at a time.

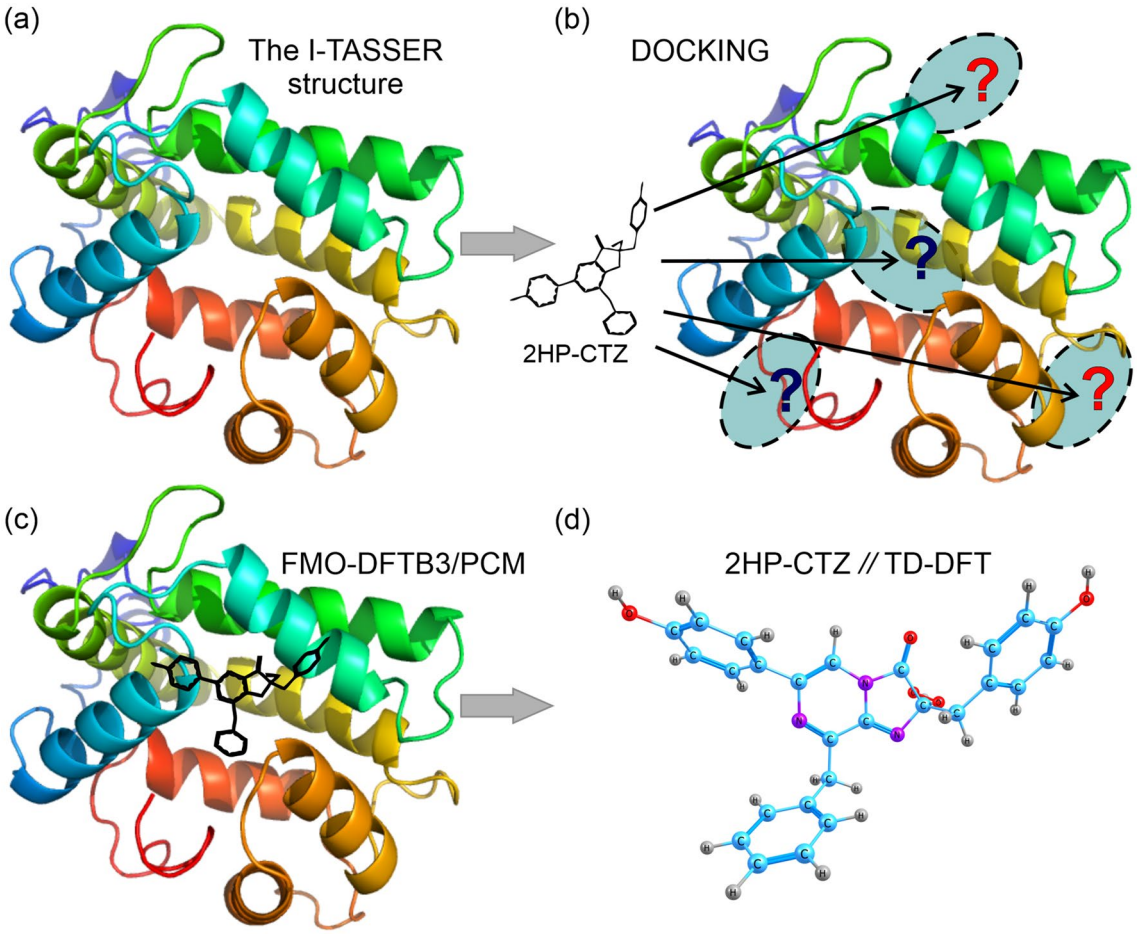

Fig. 2 Calculation workflow. **a** Prediction of berovin structure using I-TASSER. **b** Docking 2HP-CTZ to a berovin cavity by AutoDock4. **c** Geometry optimization (FMO-DFTB3/PCM) of the protein-substrate complex. **d** Calculation of the electronic excitation of 2HP-CTZ using TD-DFT

3.2 Effect of rotation of different substituents on the electronic structure and the absorption spectrum of 2HP-CTZ

As is shown in Fig. 4a–d, the LUMO is consistently localized on the *R* group with energy of − 2.34 to − 2.35 eV. The HOMO (Fig. 4e–h) is localized on the *R*, R_1, and R_3 groups with a distinctive dependence upon the rotation of the R_1 group. At the energy minimum (Figs. 3a and 4a) for the angles $\varphi_1 = 7°$ and up to 35°, the HOMO is localized on the *R* and R_1 groups (Fig. 4a, b). For the angle from 40 to 80°, the HOMO is localized on the *R*, R_1, and R_3 groups (Fig. 4c). For the angle $\varphi_1 = 85°–90°$, the HOMO is localized on the *R* and R_3 groups (Fig. 4d). As the substituent R_1 rotates, the orbital distribution expands from the central area to the left and right arms. The energy of the HOMO for $\varphi_1 = 7°–90°$ varies from − 5.49 to − 5.85 eV. The energy gap (ΔE HOMO − LUMO) changes from 3.15 to 3.51 eV. This gap (394–353 nm) correlates with the excitation energy; therefore, understanding of the changes of HOMO and LUMO and the gap between them can be helpful to elucidate the absorption wavelength, and to propose a way to control it via a modification of the adduct and/or the protein.

From the values of the excitation energy and the delocalization pattern of the HOMO in Fig. 4 (the LUMO delocalization is very similar for all angles), it can be inferred that the substituent R_1 does not affect the excitation energy considerably, as the extent of the HOMO delocalization does not affect the excitation energy much. On the other hand, as φ_1 increases and the HOMO delocalization over R_1 decreases and increases over R_3, it is can be seen that the delocalization over R_3 substantially changes the excitation energy.

At the B3LYP/cc-pVDZ level of theory, the rotation of the R_1 group leads to a spectral blue shift from 474 to 416 nm (Fig. 5). For a small change of φ_1 (5°), there is a small increase of the absorption wavelength, followed by its decrease up to $\varphi_1 = 75°$. Other angles are found to

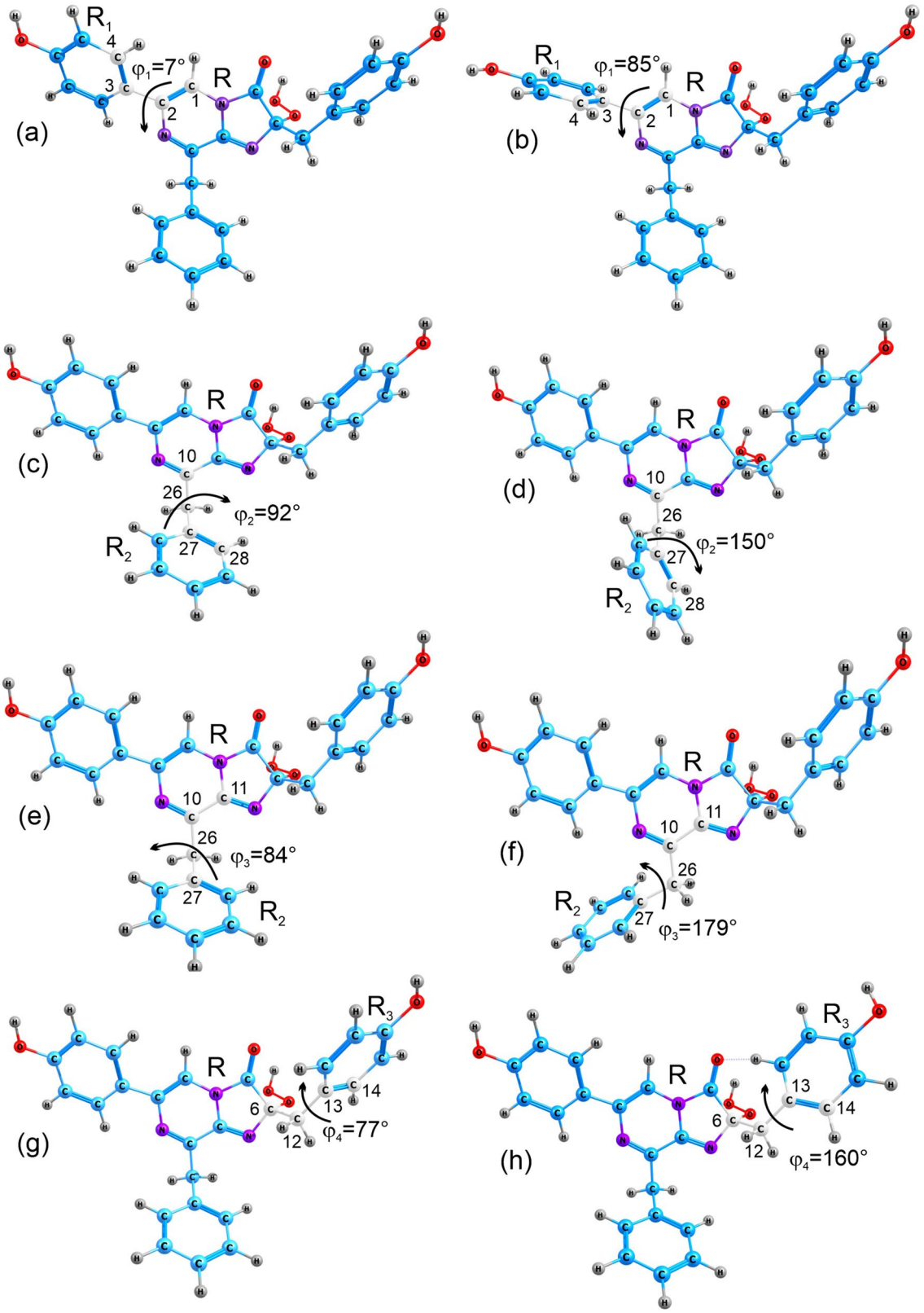

Fig. 3 Rotations of the R_1, R_2, and R_3 substituents relative to R around a single bond defined by the φ_i angle. In 2HP-CTZ, the frozen carbon atoms are indicated by numbers R_1 (1, 2, 3, 4), R_2 (10, 26, 27, 28), R_2 (11, 10, 26, 27), and R_3 (6, 12, 13, 14). **a**, **c**, **e**, **g** The ground state with $\varphi_1 = 7°$, $\varphi_2 = 92°$, $\varphi_3 = 84°$, and $\varphi_4 = 77°$. **b**, **d**, **f**, **h** The system with $\varphi_1 = 85°$, $\varphi_2 = 150°$, $\varphi_3 = 179°$, and $\varphi_4 = 160°$

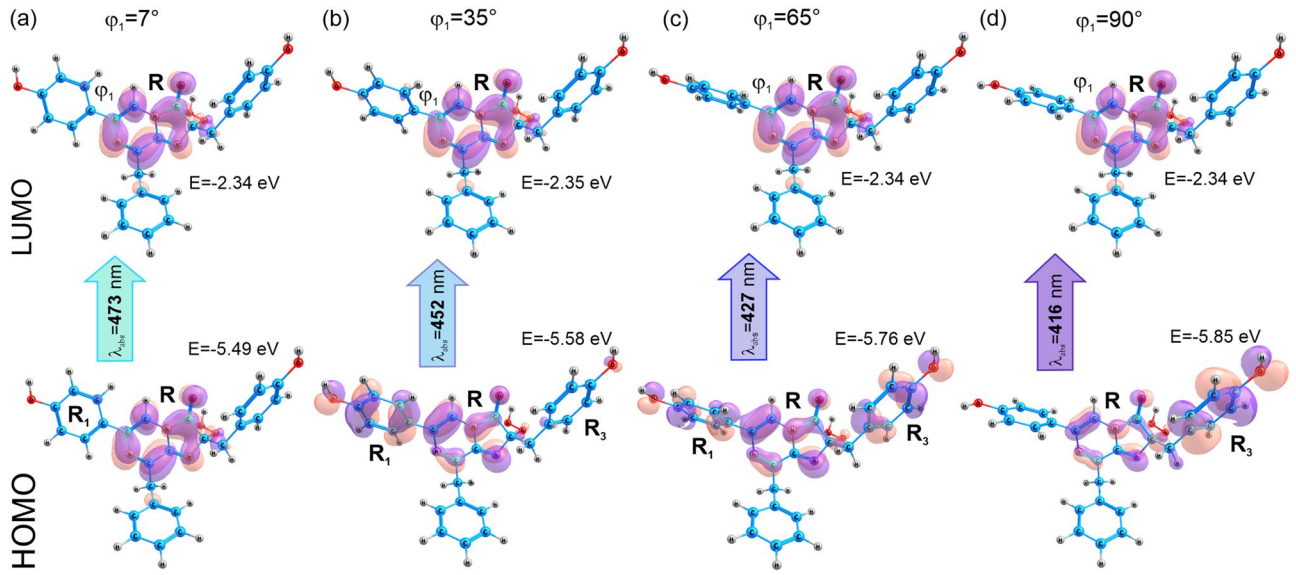

Fig. 4 HOMO and LUMO localization for 2HP-CTZ depending on the rotation of the substituent R_1. LUMOs are above the corresponding HOMOs. φ_1 is the dihedral angle (Fig. 3a, b). **a** The ground state $\varphi_1 = 7°$, **b** $\varphi_1 = 35°$, **c** $\varphi_1 = 65°$, and **d** $\varphi_1 = 90°$. *E* is the energy of a molecular orbital in eV

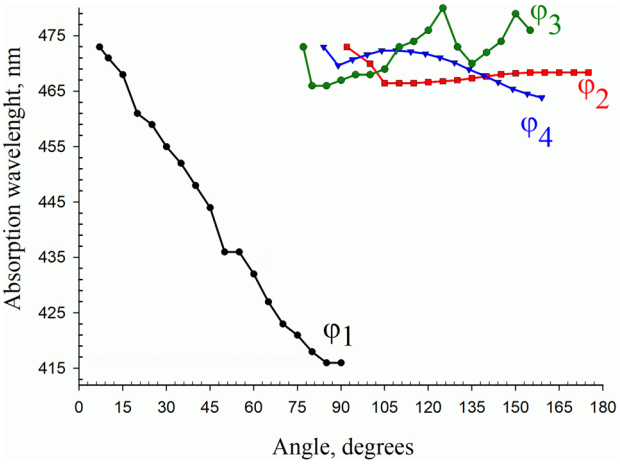

Fig. 5 Dependence of the computed absorption wavelength $\lambda(\varphi_i)$ upon the rotation dihedral angles φ_i in 2HP-CTZ at the level of B3LYP/cc-pVDZ, shown in Fig. 3. The *x*-axis denotes the absolute values of the dihedral angle of rotation relative to the ground state for each substituent

exert a much weaker influence on the wavelength, especially φ_2 (Fig. 5). These results show that the rotation of the R_1 group has the largest effect on the maximum of the absorption spectrum. The other groups (R_2 and R_3) affect the spectral properties to a small extent.

It is clear why R_2 does not influence the spectrum, since both LUMO and HOMO are not localized on it (Fig. 4). Although these orbitals do extend to R_3, this group is separated from the central core group *R* by a $–CH_2–$ group (see Fig. 1). This methylene stopper may be the reason for a weaker conjugation and a small effect of R_3 relative to R_1; likewise, R_2 is also separated from *R* by a $–CH_2–$ group.

The effect of these angles upon the total energy of the ground state is shown in Fig. 6. Since the zero point corresponds to the minimum total energy, all energies rise as φ_i increases. Among the four angles, φ_3 affects the energy most strongly (up to 19.1 kcal/mol), and φ_4 most weakly (up to 3.6 kcal/mol). For φ_1 and φ_3, there are multiple local minima.

3.3 Effect of the protein environment on the 2HP-CTZ absorption spectrum

The binding of 2HP-CTZ in the conformation corresponding to the obelin crystal structure (PDB code 1QV0) causes a shift of some amino acids in the binding pocket to give enough space to accommodate 2HP-CTZ. In total, 35 residues are found at a distance of 5 Å from 2HP-CTZ atoms. Among them, the side chains of Arg41, Phe45, Leu47, Glu57, Val58, Trp61, Arg64, Met65, Val69, Met79, Phe86, Phe87, Ile126, Leu129, Ser130, Tyr133, Tyr134, Met153, Met154, Phe157, Phe167, Leu184, Phe188, Trp192, and Tyr204 are directed inward the substrate-binding cavity.

In this complex, the hydrogen of the hydroperoxy group of 2HP-CTZ is oriented towards the side chains of Arg41 and Tyr204 which are involved in the stabilization of the hydroperoxide group through a hydrogen-bond network [51]. The 6-(*p*-hydroxy)-phenyl substituent (R_1) is located

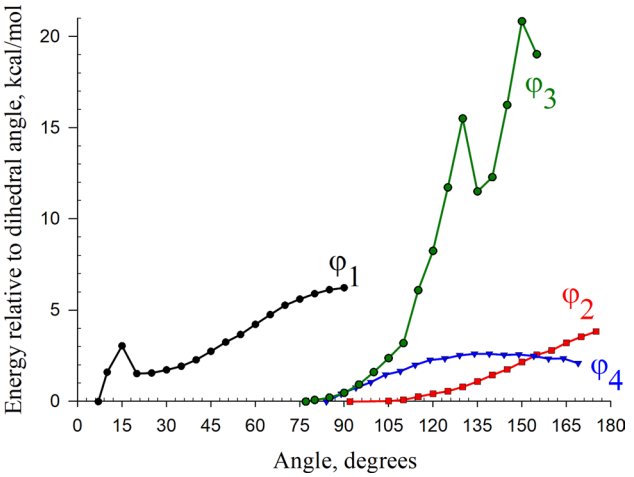

Fig. 6 Dependence of the computed ground state energy $E(\varphi_i)$ upon the dihedral angles φ_i in 2HP-CTZ at B3LYP/cc-pVDZ level (Fig. 3). Rotation by each angle on the absolute values of φ_i relative to the minimum value is plotted on the *x*-axis. Zero energy corresponds to the ground state (Fig. 3a)

within the hydrophobic pocket formed by Met65, Tyr133, Leu137, Met153, Met154, and Phe157. The interaction with these residues determines the rotation of the R_1 group relative to the equilibrium geometry of the isolated molecule. Optimization of the berovin structure with FMO2-DFTB3/PCM leads to a decrease of φ_1 to 60° compared to the geometry of the molecule in the photoprotein obelin ($\varphi_1 = 7°$) because of a readjustment of both 2HP-CTZ molecule and the surrounding residues. In optimized berovin structure, the 6-(*p*-hydroxy)-phenyl substituent geometry is ensured by Tyr133, Tyr134, Met153, Phe167, and Leu184 (Fig. 7).

3.4 A physical picture of how substrate binds to berovin

The amino acids in the active center have different functions: some are important for binding the substrate, others for its retention, some affect the conformation of the substrate, others affect its spectral properties, etc. Some useful information about the physical aspects defining the role of residues (although not directly the above functions) can be obtained by performing pair interaction energy decomposition analysis (PIEDA) for the ground state [61].

PIEDA calculations were done at the level of resolution-of-the-identity second-order Møller–Plesset perturbation theory (RI-MP2) with the basis set cc-pVDZ for the solute, and the solvent was treated with the polarizable continuum model (PCM). The structure for the analysis was taken to be the equilibrium geometry of berovin obtained at the FMO-DFTB3 level, truncated to include only the residues in the binding pocket (35 residues) and the substrate. As a result of PIEDA calculations, one obtains the interaction energy of each fragment (residue) with the substrate, and a decomposition of this total value into several components [electrostatic, quantum mechanical (charge transfer and exchange repulsion) and dispersion (van der Waals)]. The total value for each residue–substrate pair and the dispersion component (describing non-polar hydrophobic interactions) may be the most important characteristics. The solvent screening was included in total interactions. Substantial interactions are plotted in Fig. 8. Negative and positive values mean attraction and repulsion, respectively.

Among the residues surrounding 2HP-CTZ for $\varphi_1 = 77°$, only Trp103 reveals a repulsion. The Arg41 (− 18.5 kcal/mol) mainly interacts with the hydroperoxide group of 2HP-CTZ (electrostatics and dispersion are the main

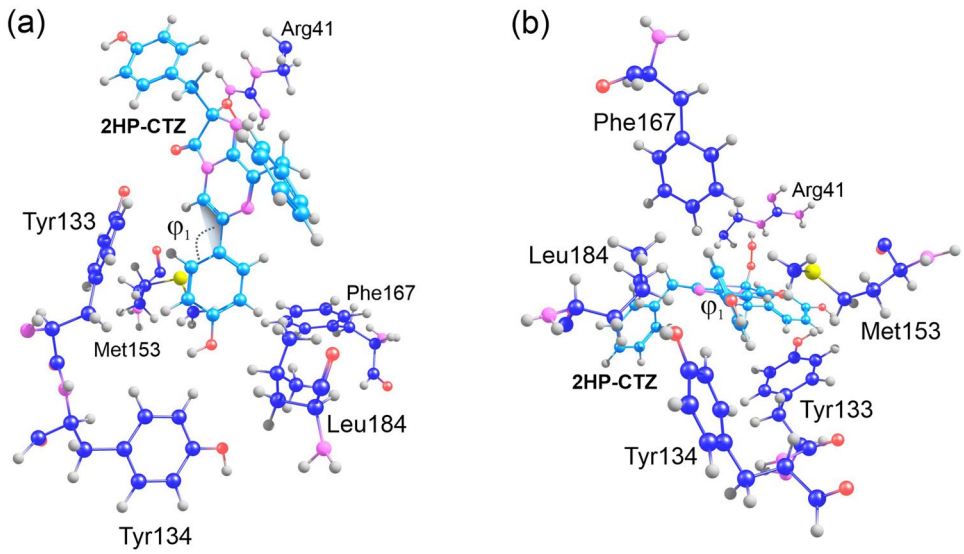

Fig. 7 Substrate-binding pocket of berovin with 2HP-CTZ from two different viewpoints. φ_1 is the dihedral angle (Fig. 3). C atoms of 2HP-CTZ are colored in light blue, C atoms of the protein are dark blue, O atoms are red, S atoms are yellow, N atoms are purple, H atoms are grey

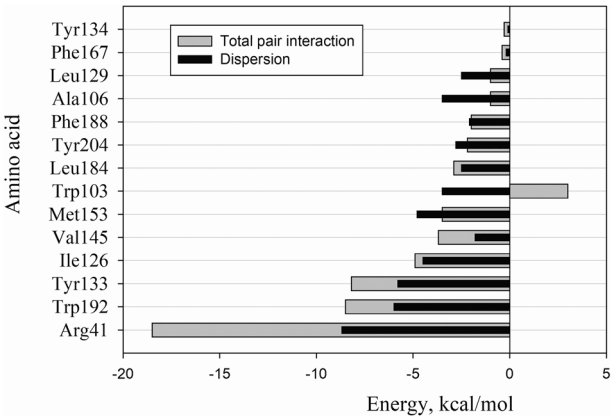

Fig. 8 Interactions between the residues in berovin and the substrate computed with PIEDA at RI-MP2/cc-pVDZ level in solution. Amino acids are shown in order of decreasing of interaction energy

components); Ile126 (− 4.9 kcal/mol) binds the R_2 substituent (dispersion); Tyr133 (− 8.2 kcal/mol), Tyr134 (− 1.0 kcal/mol) (dispersion is the main component), Met153 (− 3.5 kcal/mol), and Phe167 (− 0.4 kcal/mol) bind the R_1 group (dispersion is the main component); Trp192 (− 8.5 kcal/mol) and Trp103 (+ 3.0 kcal/mol) bind R_3 (dispersion is the main component); Tyr204 (− 2.2 kcal/mol) stabilizes the R_1 substituent (dispersion is the main component); Leu184 (− 2.9 kcal/mol) hinders the rotations of the R_1 and R_2 substituents (dispersion is the main component).

The reason why Arg41 has a large electrostatic interaction may be attributed to the hydrogen bond it forms with the substrate and close contact between the dipole on the hydroperoxide group of the substrate and the + 1 charge on the residue. Furthermore, because the substrate has several aromatic rings, it exhibits strong dispersion (hydrophobic) interaction especially with the residues that have aromatic rings as well, such as Phe, Tyr, and Trp. The importance of many of the above-mentioned residues for berovin bioluminescence was shown earlier by site-directed mutagenesis [46].

4 Conclusions

Applying the density functional theory, the key reasons for differences in visible absorption spectra of ctenophore and hydromedusan photoproteins have been revealed. It has been shown that the difference in the dihedral angles of the 6-(*p*-hydroxy)-phenyl group relative to the imidazopyrazinone core in 2-hydroperoxycoelenterazine bound to the ctenophore and hydromedusan photoproteins (50°–60° and 7°, respectively) accounts for the blue shift in the visible absorption spectrum of ctenophores versus hydromedusan photoproteins.

The reason why this angle affects the excitation is linked to the conjugation due to a direct contact of the 6-member ring of the substituent and the core. It was also demonstrated that the rotation of the 6-(*p*-hydroxy)-phenyl substituent by 60° within the inner cavity of the ctenophore photoprotein is stabilized by the side chains of Arg41, Trp192, Tyr133, and Ile126.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s43630-021-00039-5.

Acknowledgements The ab initio quantum chemical calculations were funded by RFBR and NSFC as the research project No. 19-54-53004 and RFBR research project No. 20-04-00085. The development of structural atomistic model of berovin without calcium ions generated by the I-TASSER server was funded by project 0721-2020-0033 of the Russian Ministry of Science and Education.

Declarations

Conflict of interest There are no conflicts to declare.

References

1. Haddock, S. H. D., Moline, M. A., & Case, J. F. (2010). Bioluminescence in the sea. *Annual Review of Marine Science, 2*, 443–493.
2. Widder, E. A. (2010). Bioluminescence in the ocean: origins of biological, chemical, and ecological diversity. *Science, 328*, 704–708.
3. Shimomura, O. (2006). *Bioluminescence: chemical principles and methods*. . World Scientific Publishing Co.
4. Markova, S. V., & Vysotski, E. S. (2015). Coelenterazine-dependent luciferases. *Biochemistry (Mosc), 80*, 714–732.
5. Vysotski, E. S., Markova, S. V., & Frank, L. A. (2006). Calcium-regulated photoproteins of marine coelenterates. *Molecular Biology, 40*, 404–417.
6. Shimomura, O., & Johnson, F. H. (1972). Structure of the light-emitting moiety of aequorin. *Biochemistry, 11*, 1602–1608.
7. Cormier, M. J., Hori, K., Karkhanis, Y. D., Anderson, J. M., Wampler, J. E., Morin, J. G., & Hastings, J. W. (1973). Evidence for similar biochemical requirements for bioluminescence among the coelenterates. *Journal of Cellular Physiology, 81*, 291–297.
8. Vysotski, E. S., & Lee, J. (2004). Ca^{2+}-regulated photoproteins: structural insight into the bioluminescence mechanism. *Accounts of Chemical Research, 37*, 405–415.
9. Shimomura, O., & Johnson, F. H. (1975). Regeneration of the photoprotein aequorin. *Nature, 256*, 236–238.
10. Eremeeva, E. V., Natashin, P. V., Song, L., Zhou, Y., van Berkel, W. J., Liu, Z. J., & Vysotski, E. S. (2013). Oxygen activation of apoobelin-coelenterazine complex. *ChemBioChem, 14*, 739–745.
11. Bonora, M., Giorgi, C., Bononi, A., Marchi, S., Patergnani, S., Rimessi, A., Rizzuto, R., & Pinton, P. (2013). Subcellular calcium measurements in mammalian cells using jellyfish photoprotein aequorin-based probes. *Nature Protocols, 8*, 2105–2118.
12. Alonso, M. T., Rodríguez-Prados, M., Navas-Navarro, P., Rojo-Ruiz, J., & García-Sancho, J. (2017). Using aequorin probes to measure Ca^{2+} in intracellular organelles. *Cell Calcium, 64*, 3–11.
13. Prasher, D., McCann, R. O., & Cormier, M. J. (1985). Cloning and expression of the cDNA coding for aequorin, a bioluminescent

calcium-binding protein. *Biochemical and Biophysical Research Communications, 126*, 1259–1268.
14. Inouye, S., Noguchi, M., Sakaki, Y., Takagi, Y., Miyata, T., Iwanaga, S., Miyata, T., & Tsuji, F. I. (1985). Cloning and sequence analysis of cDNA for the luminescent protein aequorin. *Proceedings of the National Academy of Sciences of the United States of America, 82*, 3154–3158.
15. Prasher, D. C., McCann, R. O., Longiaru, M., & Cormier, M. J. (1987). Sequence comparisons of complementary DNAs encoding aequorin isotypes. *Biochemistry, 26*, 1326–1332.
16. Inouye, S., & Tsuji, F. I. (1993). Cloning and sequence analysis of cDNA for the Ca^{2+}-activated photoprotein, clytin. *FEBS Letters, 315*, 343–346.
17. Inouye, S. (2008). Cloning, expression, purification and characterization of an isotype of clytin, a calcium-binding photoprotein from the luminous hydromedusa *Clytia gregarium*. *Journal of Biochemistry, 143*, 711–717.
18. Markova, S. V., Burakova, L. P., Frank, L. A., Golz, S., Korostileva, K. A., & Vysotski, E. S. (2010). Green-fluorescent protein from the bioluminescent jellyfish *Clytia gregaria*: cDNA cloning, expression, and characterization of novel recombinant protein. *Photochemical and Photobiological Sciences, 9*, 757–765.
19. Fourrage, C., Swann, K., Garcia, J. R. G., Campbell, A. K., & Houliston, E. (2014). An endogenous green fluorescent protein photoprotein pair in *Clytia hemisphaerica* eggs shows co-targeting to mitochondria and efficient bioluminescence energy transfer. *Open Biology, 4*, 130206.
20. Fagan, T. F., Ohmiya, Y., Blinks, J. R., Inouye, S., & Tsuji, F. I. (1993). Cloning, expression and sequence analysis of cDNA for the Ca^{2+}-binding photoprotein, mitrocomin. *FEBS Letters, 333*, 301–305.
21. Burakova, L. P., Natashin, P. V., Markova, S. V., Eremeeva, E. V., Malikova, N. P., Cheng, C., Liu, Z. J., & Vysotski, E. S. (2016). Mitrocomin from the jellyfish *Mitrocoma cellularia* with deleted C-terminal tyrosine reveals a higher bioluminescence activity compared to wild type photoprotein. *Journal of Photochemistry and Photobiology B: Biology, 162*, 286–297.
22. Illarionov, B. A., Markova, S. V., Bondar, V. S., Vysotski, E. S., & Gitelson, J. I. (1992). Cloning and expression of cDNA for the Ca^{2+}-activated photoprotein obelin from the hydroid polyp *Obelia longissima*. *Doklady Akademii Nauk, 326*, 911–913.
23. Illarionov, B. A., Bondar, V. S., Illarionova, V. A., & Vysotski, E. S. (1995). Sequence of the cDNA encoding the Ca^{2+}-activated photoprotein obelin from the hydroid polyp *Obelia longissima*. *Gene, 153*, 273–274.
24. Markova, S. V., Vysotski, E. S., Blinks, J. R., Burakova, L. P., Wang, B. C., & Lee, J. (2002). Obelin from the bioluminescent marine hydroid *Obelia geniculata*: cloning, expression, and comparison of some properties with those of other Ca^{2+}-regulated photoproteins. *Biochemistry, 41*, 2227–2236.
25. Ohmiya, Y., & Hirano, T. (1996). Shining the light: the mechanism of the bioluminescence reaction of calcium-binding photoproteins. *Chemistry & Biology, 3*, 337–347.
26. Kawasaki, H., Nakayama, S., & Kretsinger, R. H. (1998). Classification and evolution of EF-hand proteins. *BioMetals, 11*, 277–295.
27. Head, J. F., Inouye, S., Teranishi, K., & Shimomura, O. (2000). The crystal structure of the photoprotein aequorin at 2.3 Å resolution. *Nature, 18*, 372–376.
28. Liu, Z. J., Vysotski, E. S., Chen, C. J., Rose, J. P., Lee, J., & Wang, B. C. (2000). Structure of the Ca^{2+}-regulated photoprotein obelin at 1.7 Å resolution determined directly from its sulfur substructure. *Protein Science, 9*, 2085–2093.
29. Liu, Z. J., Vysotski, E. S., Deng, L., Lee, J., Rose, J. P., & Wang, B. C. (2003). Atomic resolution structure of obelin: soaking with calcium enhances electron density of the second oxygen atom substituted at the C2-position of coelenterazine. *Biochemical and Biophysical Research Communications, 311*, 433–439.
30. Titushin, M. S., Feng, Y., Stepanyuk, G. A., Li, Y., Markova, S. V., Golz, S., Wang, B. C., Lee, J., Wang, J., Vysotski, E. S., & Liu, Z. J. (2010). NMR-derived topology of a GFP-photoprotein energy transfer complex. *Journal of Biological Chemistry, 285*, 40891–40900.
31. Deng, L., Markova, S. V., Vysotski, E. S., Liu, Z. J., Lee, J., Rose, J., & Wang, B. C. (2004). Crystal structure of a Ca^{2+}-discharged photoprotein: implications for mechanisms of the calcium trigger and bioluminescence. *Journal of Biological Chemistry, 279*, 33647–33652.
32. Liu, Z. J., Stepanyuk, G. A., Vysotski, E. S., Lee, J., Markova, S. V., Malikova, N. P., & Wang, B. C. (2006). Crystal structure of obelin after Ca^{2+}-triggered bioluminescence suggests neutral coelenteramide as the primary excited state. *Proceedings of the National Academy of Sciences of the United States of America, 103*, 2570–2575.
33. Deng, L., Vysotski, E. S., Markova, S. V., Liu, Z. J., Lee, J., Rose, J., & Wang, B. C. (2005). All three Ca^{2+}-binding loops of photoproteins bind calcium ions: the crystal structures of calcium-loaded apo-aequorin and apo-obelin. *Protein Science, 14*, 663–675.
34. Nelson, M. R., & Chazin, W. J. (1998). Structures of EF-hand Ca^{2+}-binding proteins: diversity in the organization, packing and response to Ca^{2+} binding. *BioMetals, 11*, 297–318.
35. Eremeeva, E. V., & Vysotski, E. S. (2019). Exploring bioluminescence function of the Ca^{2+}-regulated photoproteins with site-directed mutagenesis. *Photochemistry and Photobiology, 95*, 8–23.
36. Vysotski, E. S., & Lee, J. (2007). Bioluminescent mechanism of Ca^{2+}-regulated photoproteins from three-dimensional structures. In V. R. Viviani & Y. Ohmiya (Eds.), *Luciferases and fluorescent proteins: principles and advances in biotechnology and bioimaging*. (pp. 19–41). Transworld Research Network.
37. Burakova, L. P., & Vysotski, E. S. (2019). Recombinant Ca^{2+}-regulated photoproteins of ctenophores: current knowledge and application prospects. *Applied Microbiology and Biotechnology, 103*, 5929–5946.
38. Ward, W. W., & Seliger, H. H. (1976). Action spectrum and quantum yield for the photoinactivation of mnemiopsin, a bioluminescent photoprotein from the ctenophores *Mnemiopsis sp*. *Photochemistry and Photobiology, 23*, 351–363.
39. Anctil, M., & Shimomura, O. (1984). Mechanism of photoinactivation and reactivation in the bioluminescence system of the ctenophore *Mnemiopsis*. *Biochemistry, 221*, 269–272.
40. Aghamaali, M. R., Jafarian, V., Sariri, R., Molakarimi, M., Rasti, B., Taghdir, M., Sajedi, R. H., & Hosseinkhani, S. (2011). Cloning, sequencing, expression and structural investigation of mnemiopsin from *Mnemiopsis leidyi*: an attempt toward understanding Ca^{2+}-regulated photoproteins. *Protein Journal, 30*, 566–574.
41. Markova, S. V., Burakova, L. P., Golz, S., Malikova, N. P., Frank, L. A., & Vysotski, E. S. (2012). The light-sensitive photoprotein berovin from the bioluminescent ctenophore *Beroe abyssicola*: a novel type of Ca^{2+}-regulated photoprotein. *FEBS Journal, 279*, 856–870.
42. Powers, M. L., McDermott, A. G., Shaner, N., & Haddock, S. H. (2013). Expression and characterization of the calcium-activated photoprotein from the ctenophore *Bathocyroe fosteri*: insights into light-sensitive photoproteins. *Biochemical and Biophysical Research Communications, 431*, 360–366.
43. Stepanyuk, G. A., Liu, Z. J., Burakova, L. P., Lee, J., Rose, J., Vysotski, E. S., & Wang, B. C. (2013). Spatial structure of the novel light-sensitive photoprotein berovin from the ctenophore *Beroe abyssicola* in the Ca^{2+}-loaded apoprotein conformation state. *Biochimica et Biophysica Acta, 1834*, 2139–2146.

44. Burakova, L. P., Natashin, P. V., Malikova, N. P., Niu, F., Pu, M., Vysotski, E. S., & Liu, Z. J. (2016). All Ca^{2+}-binding loops of light-sensitive ctenophore photoprotein berovin bind magnesium ions: the spatial structure of Mg^{2+}-loaded apo-berovin. *Journal of Photochemistry and Photobiology B: Biology, 154*, 57–66.
45. Molakarimi, M., Gorman, M. A., Mohseni, A., Pashandi, Z., Taghdir, M., Naderi-Manesh, H., Sajedi, R. H., & Parker, M. W. (2019). Reaction mechanism of the bioluminescent protein mnemiopsin1 revealed by X-ray crystallography and QM/MM simulations. *Journal of Biological Chemistry, 294*, 20–27.
46. Burakova, L. P., Stepanyuk, G. A., Eremeeva, E. V., & Vysotski, E. S. (2016). Role of certain amino acid residues of the coelenterazine binding cavity in bioluminescence of light-sensitive Ca^{2+}-regulated photoprotein berovin. *Photochemical and Photobiological Sciences, 15*, 691–704.
47. Molakarimi, M., Mohseni, A., Taghdir, M., Pashandi, Z., Gorman, M. A., Parker, M. W., Naderi-Manesh, H., & Sajedi, R. H. (2017). QM/MM simulations provide insight into the mechanism of bioluminescence triggering in ctenophore photoproteins. *PLoS ONE, 12*, e0182317.
48. Eremeeva, E. V., Markova, S. V., Frank, L. A., Visser, A. J., van Berkel, W. J., & Vysotski, E. S. (2013). Bioluminescent and spectroscopic properties of His-Trp-Tyr triad mutants of obelin and aequorin. *Photochemical and Photobiological Sciences, 12*, 1016–1024.
49. Marenich, A. V., Cramer, C. J., & Truhlar, D. G. (2009). Universal solution model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *The Journal of Physical Chemistry B, 113*, 6378–6396.
50. Schmidt, M. W., Baldridge, K. K., Boatz, J. A., et al. (1993). General atomic and molecular electronic structure system. *Journal of Computational Chemistry, 14*, 1347–1363.
51. Barca, G. M. J., Bertoni, C., Carrington, L., et al. (2020). Recent developments in the general atomic and molecular electronic structure system. *The Journal of Chemical Physics, 152*, 154102.
52. Zhang, Y. (2008). I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics, 9*, 40.
53. Roy, A., Kucukural, A., & Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols, 5*, 725–738.
54. Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., & Zhang, Y. (2015). The I-TASSER suite: protein structure and function prediction. *Nature Methods, 12*, 7–8.
55. Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2009). AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *Journal of Computational Chemistry, 30*, 2785–2791.
56. Fedorov, D. G., & Kitaura, K. (2004). The importance of three-body terms in the fragment molecular orbital method. *The Journal of Chemical Physics, 120*, 6832–6840.
57. Fedorov, D. G. (2017). The fragment molecular orbital method: theoretical development, implementation in GAMESS, and applications. *WIREs Computational Molecular Science, 7*, e1322.
58. Gaus, M., Goez, A., & Elstner, M. (2013). Parametrization and benchmark of DFTB3 for organic molecules. *Journal of Chemical Theory and Computation, 9*, 338–354.
59. Peach, M. J. G., Benfield, P., Helgaker, T., & Tozer, D. J. (2008). Excitation energies in density functional theory: an evaluation and a diagnostic test. *The Journal of Chemical Physics, 128*, 044118.
60. Nishimoto, Y., & Fedorov, D. G. (2016). The fragment molecular orbital method combined with density-functional tight-binding and the polarizable continuum model. *Physical Chemistry Chemical Physics, 18*, 22047–22061.
61. Fedorov, D. G. (2019). Solvent screening in zwitterions analyzed with the fragment molecular orbital method. *Journal of Chemical Theory and Computation, 15*, 5404–5416.