

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/265211122>

The Lineshape of the Electronic Spectrum of the Green Fluorescent Protein Chromophore, Part I: Gas Phase

ARTICLE *in* CHEMPHYSCHEM · OCTOBER 2014

Impact Factor: 3.42 · DOI: 10.1002/cphc.201402355

CITATIONS

3

READS

55

5 AUTHORS, INCLUDING:



Mehdi D. Davari

RWTH Aachen University

19 PUBLICATIONS 79 CITATIONS

SEE PROFILE



Francisco José Avila Ferrer

Italian National Research Council

28 PUBLICATIONS 310 CITATIONS

SEE PROFILE



Dmitry Morozov

University of Jyväskylä

5 PUBLICATIONS 10 CITATIONS

SEE PROFILE



Fabrizio Santoro

Italian National Research Council

141 PUBLICATIONS 2,931 CITATIONS

SEE PROFILE

The Lineshape of the Electronic Spectrum of the Green Fluorescent Protein Chromophore, Part I: Gas Phase

Mehdi D. Davari,^[c] Francisco J. Avila Ferrer,^[a, d] Dmitry Morozov,^[b] Fabrizio Santoro,^{*,[a]} and Gerrit Groenhof^{*,[b]}

In this work we present the vibrationally resolved optical absorption spectrum of *p*-hydroxybenzylidene-2,3-dimethylimidazolinone (HBDI), the green fluorescent protein (GFP) chromophore, computed at several levels of theory, including time-dependent DFT with various functionals and basis sets, CASSCF, CASPT2 and XMCQDPT2. We also investigated what happens to the spectrum if the ground- and excited-state geometries are optimized at different levels of theory (mixed approach), as has been used previously. The vibrationally resolved absorption spectra obtained by DFT, CASPT2 and XMCQDPT2 are very similar and consist of a main absorption peak and a shoulder that is $\sim 1500\text{ cm}^{-1}$ higher in energy. The vibrational progression increases moderately with temperature. These spectra are in

qualitative agreement with experimental action spectra, but much narrower and lack the long tail in the blue, even at high temperatures. Because our calculated emission spectra, which are equally narrow, are in good agreement with the emission of green fluorescent protein at 253 K, we argue that the action spectrum are too broad to be considered as the absorption spectrum. The CASSCF method and the mixed approaches overestimate the vibrational progressions with respect to CAM-B3LYP, CASPT2 and XMCQDPT2, due to inaccuracies in the geometric $S_0 \rightarrow S_1$ displacements. Finally, we computed the vibronic spectra of four chromophore analogues with different substitutions on the rings and found that these substitutions hardly affect the lineshape in vacuum.

1. Introduction

The green fluorescent protein (GFP) and its analogues have been widely used as fluorescent markers in molecular and cell biology.^[1,2] The absorption spectrum of GFP has two maxima, whose relative intensities are pH-dependent.^[3] These absorption peaks have been assigned to the deprotonated and protonated forms of a chromophore, which is formed autocatalytically inside the β -barrel of GFP. The bright fluorescence of GFP originates from the deprotonated chromophore. In some fluorescent proteins, the absorption peak associated with the deprotonated chromophore has a pronounced shoulder on the

blue side of the maximum, but not in others.^[1,2,4] As the shape of the emission peak in these FPs closely resembles the mirror image of the absorption, it has been suggested that the shoulder has a vibronic origin,^[5,6] but other explanations have been proposed as well.^[1,2,7]

In order to tailor the optical properties of FPs towards specific applications, it is important to understand the factors that control the position and shape of the absorption and emission spectra. However, it can be notoriously difficult to separate these factors by experiment. Therefore, in our work, we use computational methods to investigate the effect of intrinsic and environmental factors on the absorption spectrum of the GFP chromophore. As a first step, we examined whether we can reproduce the absorption spectrum of the chromophore (*p*-hydroxybenzylidene-2,3-dimethylimidazolinone, HBDI; Figure 1) in vacuum and focus on the intrinsic effects of chromophore modifications. In particular, we focus on the chromophore analogues created by Conyard et al., who observed striking differences in the absorption spectra between these analogues in water and ethanol.^[8] In the second step, which is described in a follow-up paper, we investigated the influence of a solvent environment on the spectra.^[9]

In this work, we have used various levels of theory to compute the absorption spectra of HBDI in vacuum, including time-dependent density functional theory, the multi-configurational CASSCF wave function method,^[10] and the multi-reference CASPT2^[11,12] and XMCQDPT2 methods.^[13] Our computations predict a spectrum composed of a main absorption band with a pronounced shoulder about 1500 cm^{-1} higher in


[a] Dr. F. J. A. Ferrer,[†] Dr. F. Santoro
CNR-Consiglio Nazionale delle Ricerche
Istituto di Chimica dei Composti Organo Metallici (ICCOM-CNR)
UOS di Pisa, Area della Ricerca
via G. Moruzzi 1, 56124 Pisa (Italy)
E-mail: fabrizio.santoro@iccom.cnr.it

[b] Dr. D. Morozov, Dr. G. Groenhof
Department of Chemistry and Nanoscience Center
University of Jyväskylä
P.O. Box 35 40014 Jyväskylä (Finland)
E-mail: gerrit.x.groenhof@jyu.fi

[c] Dr. M. D. Davari[†]
Theoretical and Computational Biophysics Department
Max Planck Institute for Biophysical Chemistry
Am Fassberg 11, Göttingen 37077 (Germany)

[d] Dr. F. J. A. Ferrer[†]
Physical Chemistry, Faculty of Science
University of Málaga, Málaga, 29071 (Spain)

[†] These authors made equal contributions.

 Supporting Information for this article is available on the WWW under <http://dx.doi.org/10.1002/cphc.201402355>.

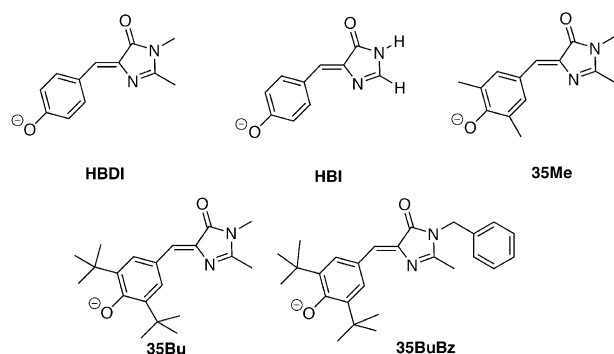


Figure 1. Chemical structure of GFP chromophore (HBDI) and its derivatives studied in this work.

energy. The shoulder is due to a combination of high-frequency modes with a significant Franck–Condon (FC) factor. Although the positions of the absorption bands agree with experimental action spectra, we found that even at higher temperatures (up to 300 K), the vibronic progression in the computed spectrum is much smaller. We suspect, therefore, that the action spectra are not identical to the absorption spectra and we provide additional arguments to support that. Finally, we demonstrate that the spectra are not very sensitive to substitutions on the rings.

Computational Methods

Vibronic spectra of HBDI were computed at the DFT, CASSCF CASPT2 and XMCQDPT2 levels of theory. At the DFT level, we used the Adiabatic Hessian (AH) approach^[14] to compute the vibronic spectra. In this approach both the S_1 and S_0 surfaces are expanded to second-order around their minimum energy configurations, so that geometric displacements, as well as frequency changes and Duschinsky rotations of the normal modes are taken into account. At all other levels of theory, however, this approach was too costly, and we used the simpler Adiabatic Shift (AS) approach instead. In this approach, the S_1 Hessian is approximated by the S_0 Hessian.^[14]

First, geometries of the ground state (S_0) and first singlet excited state (S_1) were optimized at the level of theory under consideration. The ground state was also optimized at the MP2 level. This was done because the MP2 normal modes were needed to compute the XMCQDPT2 spectrum, as explained below. To describe the excited state in the DFT calculations, we used the time-dependent formulation (TDDFT), for which analytical gradients are available. The optimizations at the CASPT2 and XMCQDPT2 levels were carried out with C_s symmetry, using numerical gradients.

At the DFT level, we tested several functionals and basis sets. First, we used a 6-31G(d) basis set with the following functionals, each with a different amount of exact exchange: BLYP,^[15,16] BHandHLYP (BHLYP),^[17] B3LYP,^[18] PBE0,^[19] ω B97X-D^[20] and CAM-B3LYP.^[21] Then, using the CAM-B3LYP functional, we tested Pople's basis set series 6-31G(d) to 6-31++G(d,p)^[22,23] and Dunning's cc-pVDZ^[24] basis. In all (TD)DFT calculations, the Hessians of the ground and excited states were obtained by numerical differentiation. In addition, also the first derivatives of the transition dipole, required for Herzberg–Teller (HT) simulations, were obtained numerically. HT terms were only included at the DFT level.

At the CASSCF level, we included all π electrons (16) and all π orbitals (14), expanded in the 6-31G(d) basis set, in the active space, i.e., CASSCF(16,14)/6-31G(d). The S_0 Hessian was obtained numerically.

At the CASPT2 level, gradients were obtained by numerical differentiating CASPT2/CASSCF(10,9)/6-31G(d) energies. Because evaluating the Hessian is too time-consuming, we used the CASSCF S_0 Hessian instead for computing the spectra. The $S_0 \rightarrow S_1$ displacement vector \mathbf{K} along the CASSCF normal modes was computed by subtracting the geometry displacements between the CASPT2 S_0 and S_1 minima and the CASSCF S_0 minimum: $\mathbf{K}[\text{PT2}] = \mathbf{K}_{10} - \mathbf{K}_{00}$, with $\mathbf{K}_{10} = \mathbf{K}[S_1(\text{CASPT2}) - S_0(\text{CASSCF})]$ and $\mathbf{K}_{00} = \mathbf{K}[S_0(\text{CASPT2}) - S_0(\text{CASSCF})]$.

At the XMCQDPT2/cc-pVDZ level, the ground state MP2/cc-pVDZ Hessian was used in all calculations. As for CASPT2, the $S_0 \rightarrow S_1$ displacements vector \mathbf{K} along the MP2 normal modes was obtained by subtracting the displacements between XMCQDPT2 S_0 and S_1 minima and the MP2 ground state geometry: $\mathbf{K}[\text{XMCQDPT2}] = \mathbf{K}_{10} - \mathbf{K}_{00}$, with $\mathbf{K}_{10} = \mathbf{K}[S_1(\text{XMCQDPT2}) - S_0(\text{MP2})]$ and $\mathbf{K}_{00} = \mathbf{K}[S_0(\text{XMCQDPT2}) - S_0(\text{MP2})]$.

The spectra at different levels of theory were computed within the Franck–Condon approximation (FC) at 0 K. The vibronic spectra of HBDI obtained with the CASPT2 and XMCQDPT2 methods were used to verify the validity of the TD-DFT results. Based on the good agreement with the correlated wavefunction methods, the CAM-B3LYP/6-31G(d) level of theory was considered to provide a good compromise between accuracy and efficiency. Therefore, we used this level of theory to compute the influence of temperature and the effect of including the Herzberg–Teller (FCHT) terms.

All spectra were computed with the *FCclasses* code.^[14,25–28] For simulating spectra at 0 K, the time-independent (TI) approach was used. A maximum number of 25 overtones for each mode and 20 combination bands on each pair of modes were included in the calculations. The maximum number of integrals to be computed for each class was set to 10^9 . The convergence in the total intensity is above 99%. For spectra at higher temperatures, we used the effective pre-screening time-dependent approach (TD).^[29] This approach is based on an analytical expression of the finite-temperature time-correlation function and it delivers fully converged spectra.^[30–33]

Spectra are reported as normalized absorption line shapes, where line shape is absorptivity divided by radiation frequency. Because similar spectra were obtained using a Lorentzian or a Gaussian function to convolute the stick spectra, the conclusions of our computations do not depend on the specific choice of the convolution function. We therefore used Gaussian function with a full width at half maximum (FWHM) of 320 cm^{-1} to model line broadening in our spectra.

All DFT computations were performed with Gaussian09 package,^[34] the CASSCF and CASPT2 computations were performed with Molcas 7.4.^[35] MP2/XMCQDPT2 calculations were performed with the Firefly 8.0.1 program package.^[36]

2. Results and Discussion

2.1. Effect of DFT Functional and Basis Set on the Absorption Lineshape of HBDI

To the best of our knowledge, no studies exist in which the influence of the method and basis set on the shape of the vibronic spectra of HBDI has been investigated systematically. In a recent paper^[37] authors present a study of performance of DFT functionals for computing absorption and fluorescence band shapes for a set of conjugated molecules not including GFP chromophore. Other previous benchmarks have focused only on the vertical or adiabatic 0–0 transition^[38,39] in HBDI. Therefore, we first compared the line shapes for HBDI in vacuum computed with various DFT functionals and basis sets. At all DFT levels tested, the chromophore has a C_s geometry in both S_1 and S_0 . Furthermore, all levels predict that the first strongly dipole allowed transition corresponds to the excitation of one electron from the HOMO to LUMO orbital (a π – π^* transition), which are illustrated in Figure 2.

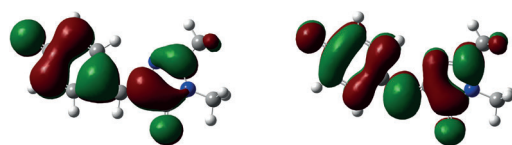


Figure 2. Highest molecular orbital (HOMO, π , left) and lowest unoccupied molecular orbital (LUMO, π^* , right) of the electronic ground state (S_0) of HBDI in the gas phase.

The simulated vibronic spectra at 0 K using various DFT functionals are shown in Figure 3, in which we have shifted the 0–0 transitions to the origin. All the examined functionals predict very similar vibronic spectra, with the exception of the pure BLYP functional that predicts a broad band in the 500–1000 cm^{-1} frequency range not found with the other functionals or higher levels of theory (see below). The position of the 0–0 transition, E_{00} reported in Table 1, varies in the range 24 000–26 000 cm^{-1} , and increases with the amount of Hartree–Fock exchange.

Using the CAM-B3LYP functional, we then computed the spectrum with different basis sets. Figure 3 shows that the vibronic spectra obtained at CAM-B3LYP level with various basis sets are very similar. As far as the absolute position of the

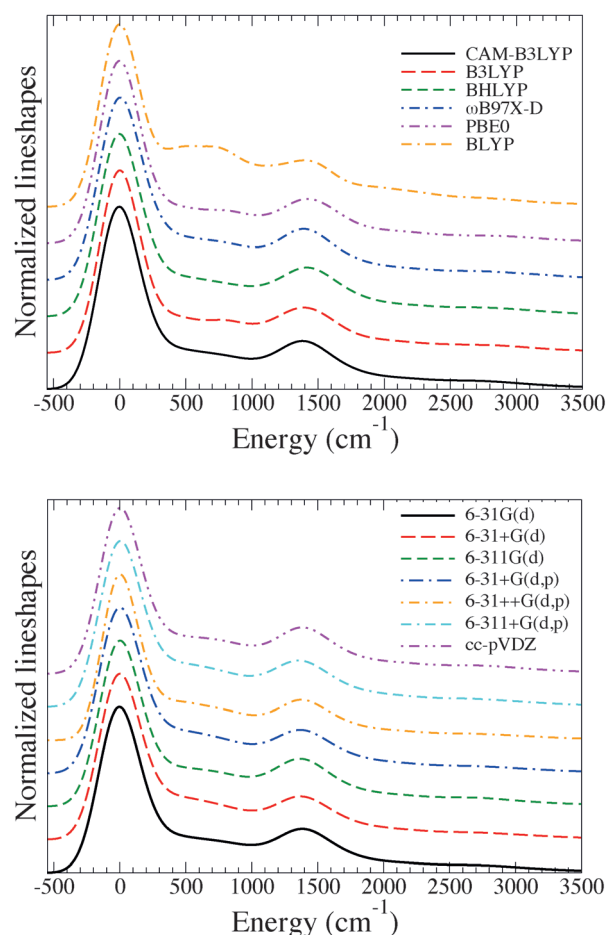


Figure 3. Vibrationally resolved absorption spectra of anionic HBDI in the gas phase at 0 K calculated using various DFT functionals and the 6-31G(d) basis sets (top), and various basis sets with the CAM-B3LYP functional (bottom); Gaussian convolution with $\text{FWHM} = 320 \text{ cm}^{-1}$. To make the comparison easier, the maxima of the spectra are shifted to 0. The 0–0 transition energies are listed in Table 1. The superimposed spectra is given in Figure S10 in Supporting Information.

spectra is concerned, Table 1 shows that it is moderately dependent on the basis set. The blue-most spectrum is obtained with the small 6-31G(d) basis set, while extending the basis set causes the position of the spectrum to vary within $\sim 1100 \text{ cm}^{-1}$, mostly due to the introduction of diffuse functions. Thus, although the size of the basis set has an effect on the 0–0 transition, it has only a minor effect on the vibronic progression. Therefore, for the purpose of the current work, it is sufficient to use the relatively small 6-31G(d) basis set.

Next, we investigated the influence of Duschinsky mixing on the spectra by computing the CAM-B3LYP spectrum within the Adiabatic Shift (AS) approximation and comparing it to the spectrum obtained with the Adiabatic Hessian (AH) approach. The AS spectra, which are shown in Figure S2 of the Supporting Information are very similar to the AH spectra, demonstrating that the Duschinsky effects are very small.

Finally, we checked the contribution of non-Condon effects by including the Herzberg–Teller terms. As shown in Figure S3, the FC and FCHT spectra are essentially the same, suggesting that non-Condon effects are negligible for HBDI in the gas

Table 1. Effect of the basis set and DFT functional in the energy of the 0–0 vibronic transition for HBDI in gas phase.			
CAM-B3LYP/Basis Set	E_{00} [cm^{-1}]	Functional/6-31G(d)	E_{00} [cm^{-1}]
6-31G(d)	25 020	CAM-B3LYP	25 020
6-31 + G(d)	23 940	B3LYP	24 020
6-311G(d)	24 800	BHLYP	26 050
6-31 + G(d,p)	23 890	ω B97X-D	24 960
6-31 + + G(d,p)	23 940	PBE0	24 760
6-311 + G(d,p)	23 950	BLYP	22 250
cc-pVDZ	24 700		

phase. This conclusion is in agreement with the finding by Karmarchik et al.^[40] for one photon absorption of HBDI. As a result, FC spectra within the harmonic approximation capture the main features of the vibronic spectra of the HBDI anion. Therefore, in the remainder we focus on FC spectra only.

2.2. Comparison to Spectra Obtained with Wave-Function-Based Methods

The spectra computed at CASSCF, CASPT2 and XMCQDPT2 levels of theory are shown in Figure 4. The CAM-B3LYP/6-31G(d) spectrum is in good agreement with the CASPT2 and XMCQDPT2 spectra. A small difference is seen in the position of the shoulder at $\sim 1500\text{ cm}^{-1}$, which is slightly more blue-shifted at the XMCQDPT2 and (a bit more) at the CASPT2 level of theory. The height of this peak is also slightly different, with the CAM-B3LYP height in between the XMCQDPT2 and CASPT2 heights. These differences are most likely caused by using different normal modes and frequencies in our calculations. Furthermore, a larger active space was used for optimizing at the XMCQDPT2 level (i.e., (14,13)) than at the CASPT2 level (i.e., (10,9)).

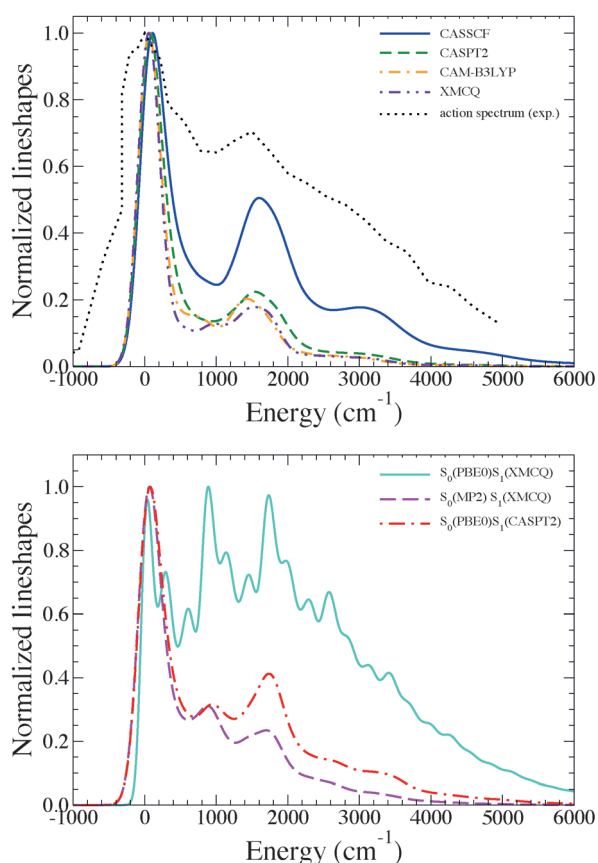


Figure 4. Vibrational resolved absorption spectra of HBDI in the gas phase at various levels; they are compared with the experimental action spectrum (black)^[52] (top panel) and with spectra obtained using different levels of theory for S_0 and S_1 geometry optimization. The 0–0 transitions are shifted to zero. Gaussian convolution with $\text{FWHM} = 320\text{ cm}^{-1}$, except for the $S_0(\text{PBE0})/S_1(\text{XMCQ})$ spectrum convoluted with $\text{FWHM} = 160\text{ cm}^{-1}$ in order to show a resolution comparable to the one of Figure 10 in ref. [56].

Based on the good agreement of the DFT spectra with the CASPT2 and XMCQDPT2 spectra, we conclude that DFT is sufficiently accurate for computing vibronic spectra of HBDI. Furthermore, the results are not very sensitive to basis set or functional (except BLYP). Because the CAM-B3LYP functional yields a good compromise between efficiency and accuracy in related systems,^[38,39] and is size-independent with respect to the overestimation of excitation energies,^[41–43] we will be using the CAM-B3LYP functional in combination with the 6-31G(d) basis set in what follows.

2.3. Assignment of Vibronic Transitions in HBDI in the Gas Phase

The gas phase FC absorption spectrum of anionic HBDI at 0 K computed at CAM-B3LYP/6-31G(d) level with the AH approach, is shown in Figure 5, both as a stick spectrum (Franck–Condon Factors or FCFs) and as a convolution of this stick spectrum with Gaussians. The convoluted spectrum exhibits a maximum and a major shoulder with relative intensities of 1.00 and ~ 0.25 at 0 K. Several vibronic transitions contribute to the vibrational progression and we have assigned the main transitions in the stick spectrum.

Although the 0–0 transition has the largest individual FCF, the high density of peaks associated with blending of low vibrational frequency modes in this region, causes a blue shift of the lowest energy peak.

The shoulder appears at about 1500 cm^{-1} on the blue side of the maximum. It arises from a combination of several transitions (assigned with respect to excited-state normal modes); the most intense ones lie at 1390 cm^{-1} (49^1 , the fundamental excitation of mode 49, Figure 6) and 1474 cm^{-1} (49^13^1 , a binary combination band of modes 3 and 49), respectively. Mode 49 is assigned to the in-phase stretching of C=C and C–C (vinyl) bonds, which are assumed to be optically active (Figure 6).

The vibrational modes at 84 cm^{-1} (3^1), 168 and 252 cm^{-1} (3^2 and 3^3 , overtones of band 3) are responsible for the other small peaks in the spectrum (Figure 6).

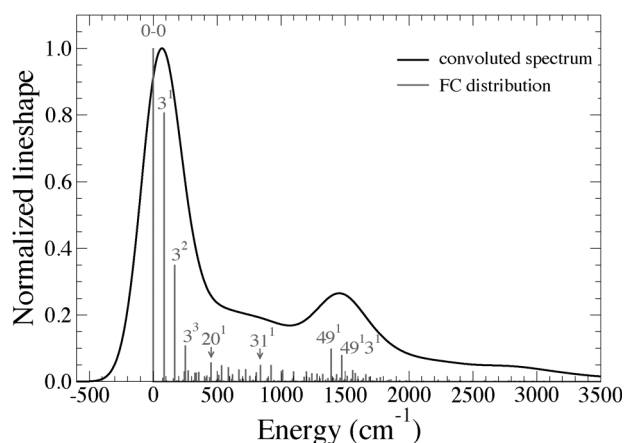


Figure 5. Vibrational resolved absorption spectra of anionic HBDI in the gas phase at 0 K; Gaussian convolution with $\text{FWHM} = 320\text{ cm}^{-1}$. Energies are given with respect to 0–0 transition calculated at CAM-B3LYP/6-31G(d) level of theory (see Table 1 for absolute 0–0 energies).

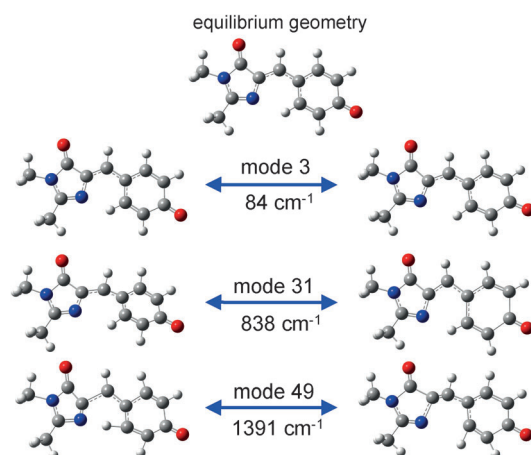


Figure 6. Most active normal modes in the absorption spectrum of HBDI. Because the modes are not visible in the vector representation, two geometries in opposite phase for each mode are presented. Mode 49 is in-phase stretching of C=C and C–C (vinyl) bonds and mode 31 is a phenyl-breathing coupled to the central bridge (–C=CH–C–) bending. Animations of these modes are available as Supporting Information.

A number of transitions with moderate intensity exist in the range 500–1000 cm^{-1} , which include the transition involving mode 31 (838 cm^{-1}) that we assign to a breathing of the phenyl ring coupled with the central bridge (–C=CH–C–) bending (Figure 6).

Previous experimental and computational studies also indicated that the vibrational progression in this frequency region is dominated mainly by in-plane stretching modes that are localized over the imidazolinone or phenolate portions of chromophore, rather than being delocalized over the entire chromophore.^[44–48] The C=O stretchings are distributed in three modes with frequencies in the range 1600–1700 cm^{-1} and have very small activity in the absorption spectrum.

2.4. Comparison with Experimental Spectra

There is no consensus on the shape of the absorption spectrum of HBDI in the gas phase.^[49–51] In previous studies^[40,52] vibronic spectra of anionic HBDI have been compared with the gas phase action spectrum^[52] at 77 K. In Figure 4 we compare this experimental action spectrum to the FC spectrum computed at CAM-B3LYP/6-31G(d) level of theory at 0 K. Our vibronic spectrum correctly predicts the existence of a shoulder blue-shifted by $\sim 1500 \text{ cm}^{-1}$ with respect to the maximum (0–0) peak. However, its relative intensity with respect to the main peak is remarkably lower compared to the experimental action spectrum. Furthermore, the blue tail in our spectrum loses intensity much faster than the action spectrum.

Since the experimental action spectrum was measured at 77 K, it may contain excitations in low frequency modes which could contribute to the broadening. To investigate this possibility we have computed the vibronic spectra in gas phase at several temperatures: 20, 100, 200 and 300 K. These spectra, which are shown in Figure 7, reveal a progressive broadening of the spectrum with increasing temperature. However, com-

paring Figure 4 and Figure 7 (top left panel) suggests that the long tail observed in the action spectra cannot be explained by the temperature.

In summary, while the main spectral features observed in the action spectrum are qualitatively reproduced in our simulation, remarkable differences exist, and they mostly concern the blue-wing of the spectrum.

These discrepancies imply that either our computations, or, as suggested by Chingin et al.,^[51] the gas-phase action spectra do not represent the absorption spectrum of the anionic chromophore in vacuum. To resolve this issue, we first turn to the protein spectra. Because absorption by the protein is due to both neutral and anionic forms of the chromophore,^[53] the spectrum is a mixture. Instead, emission is due to the anionic chromophore only. Therefore, we focus on the latter and will compare the emission spectrum computed in vacuum, to the emission spectrum of GFP measured at 253 K.

The GFP emission spectrum at 253 K (Figure 8) shows a much smaller vibrational progression than the vacuum action spectrum.^[54] We consider it reasonable to assume that above the protein glass transition at $\sim 200 \text{ K}$,^[55] the fluctuating protein environment would lead to a broader spectrum of the chromophore than in vacuum at 77 K. Therefore, we take the observation of the sharper spectrum in the protein as a clear support for the suggestion of Chingin et al.^[51] that the action spectrum^[52] is not the same as the absorption spectrum.

Our computed emission spectrum is shown in Figure 8 and is in excellent agreement with the emission spectrum of GFP.^[54] Therefore, we conclude that the smaller vibrational progression in both our absorption and emission spectra are accurate and that the discrepancy with the action spectrum are due to the fact that the latter is not identical with the one-photon absorption spectrum. Furthermore, we also infer that the effect of the protein environment on the vibronic transitions of the chromophore is rather limited. The same conclusion was also drawn by Bochenkova et al. on the basis of QM/MM spectra calculations.^[56]

We speculate that the differences between the intrinsic absorption spectrum and the action spectrum might be connected to an incomplete deconvolution of the ionization and auto-ionization channels. We also notice that the action spectra given by Forbes et al.^[52] are different from the action spectra measured by Chingin et al.^[51] even with respect to the position of the maxima (480 nm in ref. [52] and 440 nm in ref. [51]). Chingin et al.^[51] have suggested that this might be due to differences in laser power, so that the spectrum recorded by Forbes et al.^[52] might also reflect two-photon absorption processes. Nevertheless, although resolving these controversies is beyond the scope of our work, we consider that there is sufficient ground to doubt whether the action spectrum represents the absorption spectrum.

2.5. Comparison with Previously Simulated Spectra

Absorption spectra of anionic HBDI in gas phase have been computed previously and we now compare our CAM-B3LYP, CASPT2 and XMCQDPT2 spectra obtained with AS model to

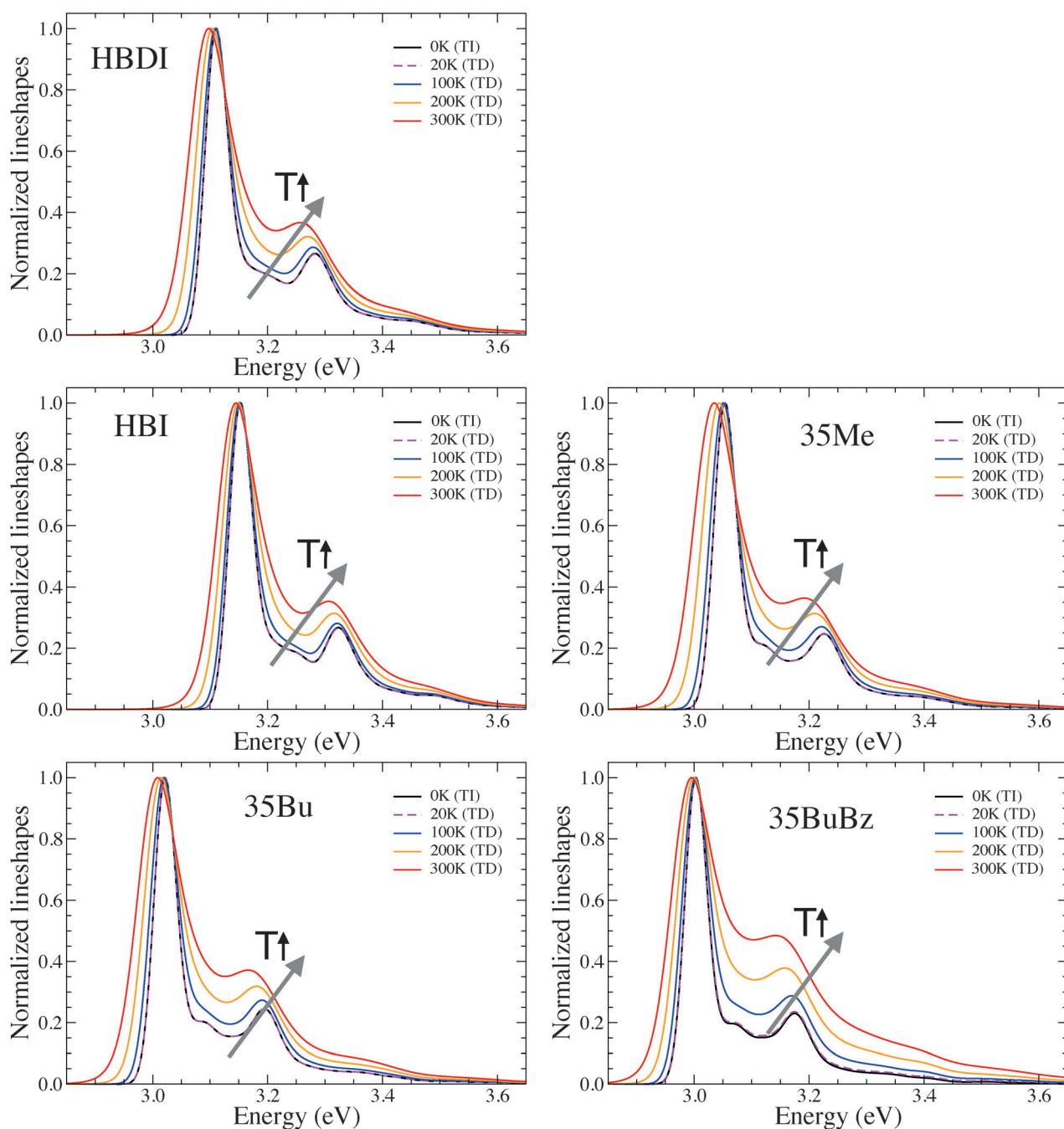


Figure 7. Vibrationally resolved absorption spectra of anionic HBDI and its derivatives in the gas phase at 0–300 K; Gaussian convolution with FWHM = 320 cm⁻¹. TD: time-dependent, TI: time-independent approach.

those reported in the literature. Ai et al.^[57] computed the AS vibronic spectrum at CASSCF(12,11)/6-31G(d) level of theory, and found that it agrees better with the gas-phase action spectrum measured by Forbes et al.^[52] than their DFT spectra. Therefore, these authors concluded that DFT/TD-DFT is not able to correctly capture the differences between the S_0 and S_1 potential energy surfaces of the chromophore. However, based on the agreement between our DFT spectrum on the one hand and the CASPT2 and XMCQDPT2 spectra on the other hand, we doubt that conclusion.

To support our case, we also computed the CASSCF spectrum of HBDI at 0 K (Figure 4). Our CASSCF spectrum is similar to previous CASSCF spectra,^[57,58] but not completely equivalent. Most probably, the discrepancies are caused by the different active spaces: Ai and co-workers used a reduced active space,^[57] whereas we included all π electrons and orbitals. In agreement with the results of Ai et al.,^[57] CASSCF predicts a significantly larger vibrational progression than CAM-B3LYP. In particular, the ratio of the intensities of the shoulder and main peaks is 0.55 at the CASSCF level, but only 0.2 at the CAM-

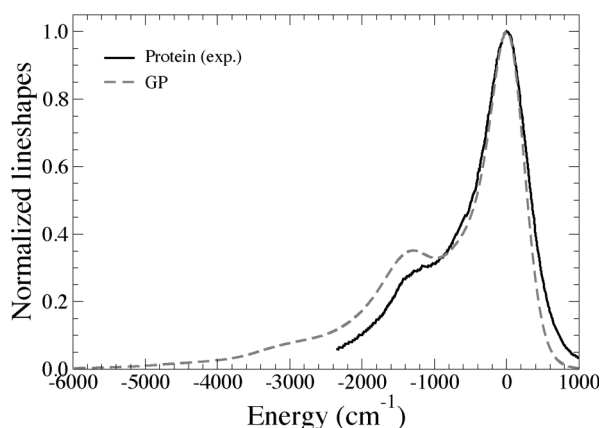


Figure 8. Vibrationally resolved emission spectrum of anionic HBDI at 253 K using the CAM-B3LYP/6-31G(d) method. A comparison is made with the protein emission spectrum reported in ref. [54].

B3LYP level. However, since our CASPT2 and XMCQDPT2 spectra are also much narrower than the CASSCF spectrum and have a smaller intensity ratio between the shoulder and main peak as well, we conclude that the larger progression in the CASSCF spectra is an artifact due to the neglect of dynamic electron-electron correlation.

We found that neglecting dynamic correlation causes small inaccuracies in the geometries. Comparing the geometric shifts between S_1 and S_0 at CASSCF, CAM-B3LYP, CASPT2 and XMCQDPT2 levels of theory in Figure 9 reveals that the CASSCF displacements are the largest. In particular, the $C_{ph}-C$

bond is significantly more displaced at the CASSCF level. Appreciable differences also appear on the $C-O^-$ internuclear distances. CASPT2 and XMCQDPT2 predict very similar displacements, which, in most of the cases, are intermediate between CASSCF and CAM-B3LYP displacements.

The absorption spectrum of anionic HBDI was also calculated by Kamarchik et al.^[40] and by Bochenkova,^[56] using a parallel normal mode approximation, which is equivalent to what we defined the Adiabatic Shift (AS) approach.^[14] In contrast to our calculations, however, both groups have optimized the geometry in the excited state state at a different level than the ground state. Presumably to avoid heavy computations, Kamarchik et al.^[40] used MP2 for S_0 and CIS(D) for S_1 , while Bochenkova et al.^[56] used PBE0 for S_0 and XMCQDPT2/CASSCF-(14,13) for S_1 . Both calculations predict absorption spectra that are broader than our spectra. Although the broadening brings these spectra in closer agreement with the experimental action spectra than ours, we believe that the broadening is caused by inaccuracies in the displacements when different levels of theory are used for S_1 and S_0 .

To support our hypothesis, we recomputed the spectrum of HBDI using the same approach as Bochenkova et al.^[56] Our $S_0(\text{PBE0/cc-pVDZ}) \rightarrow S_1(\text{XMCQDPT2/cc-pVDZ})$ spectrum is shown in Figure 4 and is very similar to their spectrum.^[56] The small differences are probably due to the fact that these authors used a basis set augmented with a set of diffuse functions (aug-cc-pVDZ). Indeed, the $S_0(\text{PBE0}) \rightarrow S_1(\text{XMCQDPT2})$ spectrum is much broader than the both PBE0 (Figure 3) and XMCQDPT2 spectra (Figure 4).

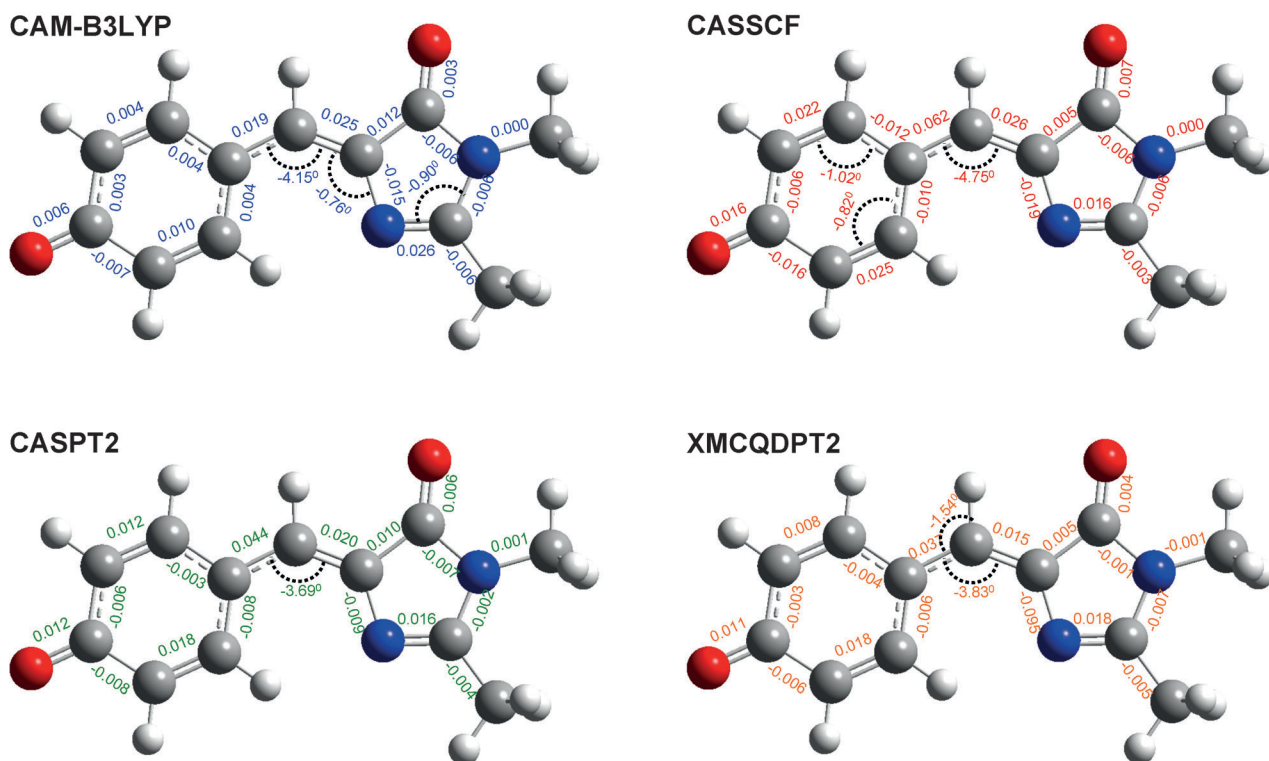


Figure 9. Comparison of geometrical shifts for HBDI in $S_0 \rightarrow S_1$ transition computed using CAM-B3LYP (top left), CAS(16,14) (top right), CASPT2(10,9) (bottom left) and XMCQDPT2 (bottom right) methods with the 6-31G(d) basis set. Angular shift shorter than 0.70° not shown.

The main vibrational progression in the $S_0(\text{PBE0/cc-pVDZ}) \rightarrow S_1(\text{XMCQDPT2/cc-pVDZ})$ spectrum is due to mode 30 (PBE0), a phenolate-breathing mode coupled to central bridge with frequency $\sim 850 \text{ cm}^{-1}$ (in combination with a low-frequency mode, equivalent to 3 in Figure 6). This mode is analogous to the CAM-B3LYP S_1 mode 31 in Figure 6 but exhibits a much larger displacement in the PBE0/XMCQDPT2 hybrid approach, namely 1.24 (dimensionless units). For comparison, the displacement along this mode is 0.27 if PBE0 is used to optimize both S_0 and S_1 , 0.25 at CAM-B3LYP and 0.34 at XMCQDPT2 level. Further analysis reveals that more than half of this larger displacement (1.24) is due to an intrinsic large displacement (0.76) between the PBE0 and XMCQDPT2 S_0 geometries along this breathing mode. Because the FC intensities are proportional to the square of the displacement, the artificial enhancement of the displacement overestimates the vibrational progression of the $S_0(\text{PBE0}) \rightarrow S_1(\text{XMCQDPT2})$ spectrum.

Further evidence that using different approaches to optimize the S_0 and S_1 geometries can lead to inaccurate spectra, is provided by a larger progression if we use PBE0 for S_0 and CASPT2 for S_1 (Figure 4). Although the 'pure' $S_0(\text{PBE0}) \rightarrow S_1(\text{PBE0})$ and $S_0(\text{CASPT2}) \rightarrow S_1(\text{CASPT2})$ spectra are extremely similar (compare Figure 3 and Figure 4), the mixed $S_0(\text{PBE0}) \rightarrow S_1(\text{CASPT2})$ spectrum is much broader. Also in this situation, the phenolate-breathing mode at $\sim 850 \text{ cm}^{-1}$ has a much larger displacement in the mixed approach (0.56) than in the pure PBE0 (0.27) or CASPT2 (0.27) calculations. The origin of the much larger displacement at the $S_0(\text{PBE0}) \rightarrow S_1(\text{CASPT2})$ level is that the CASPT2 geometries of both S_1 and S_0 are displaced by 0.27 with respect to the PBE0 geometries.

Thus, the displacement along the phenolate-breathing mode is extremely sensitive to using different levels of theory to optimize the ground and excited state. As a further illustration, we consider the displacement if we use MP2 to optimize the ground state, rather than XMCQDPT2. Despite the very high similarity of the MP2 and XMCQDPT2 structures ($\Delta E \approx 100 \text{ cm}^{-1}$, for details see Figure S8 in the Supporting Information), the displacement is 0.23 along the phenolate-breathing mode, which is caused by small differences in the bonding angles in the phenolate ring. When simulating the $S_0(\text{MP2}) \rightarrow S_1(\text{XMCQDPT2})$ absorption spectrum, this displacement is added to the intrinsic displacement of 0.34 that is coupled to the optical transition at the XMCQDPT2 level. Therefore, the total displacement is 0.57. Because of the quadratic dependence of the intensities on the displacements, the larger displacement causes an artificial increase of the absorption at 850 cm^{-1} by a factor of $(0.57/0.34)^2 \approx 3$ (Figure 4).

The sensitivity of the spectra to spurious displacements is also illustrated in Figure S4 of the Supporting Information, where we computed "fictitious" $S_0 \rightarrow S_0$ and $S_1 \rightarrow S_1$ spectra adopting the displacements between the geometries of the same electronic state optimized at different levels of theory. Indeed, these displacements give rise to vibronic progressions, along the same modes involved in the $S_0 \rightarrow S_1$ optical transition (like mode 31), and with intensities comparable to the $S_0 \rightarrow S_1$ spectra.

Finally, our conclusion that CASSCF and mixed approaches overestimate broadening, is also supported by the absorption spectrum of HBDI in DMSO, which is structureless and has a FWHM $\sim 2400 \text{ cm}^{-1}$.^[59] If we assume that the underlying vibrational structure is the same as in gas phase and that the effect of this aprotic solvent is only to inhomogeneously broaden the absorption spectrum, we would need a gaussian with a FWHM of at least $\sim 1600 \text{ cm}^{-1}$ to remove the vibrational structure from our CASSCF gas-phase spectrum (See Figure S9 in Supporting Information), obtaining a total spectrum whose FWHM is $\sim 3300 \text{ cm}^{-1}$. Since this is much larger than the FWHM in the experiment, we conclude that the CASSCF spectrum is too broad. On the other hand, to broaden our CAM-B3LYP spectra, a FWHM of $\sim 1600 \text{ cm}^{-1}$ would be required, which is much more reasonable, since it gives rise to a total spectrum with FWHM $\sim 2150 \text{ cm}^{-1}$.

Our analysis revealed that systematic errors in geometries do not cancel if different levels of theories are used for the ground and excited states. As a consequence, the displacements predicted by such approaches can be highly inaccurate and cause artificial broadening of the spectra. To avoid such spurious absorptions in the calculated spectra, we recommend that the ground and excited state geometries are optimized at the same level of theory instead.

2.6. Effect of Substitution on the Absorption Lineshapes of HBDI in the Gas Phase

Having established the validity of CAM-B3LYP/6-31G(d), we used this level of theory to investigate the effect of substitutions on the vibronic spectra of the HBDI analogues (Figure 1). The molecular orbitals involved in the optical transitions of the analogues, sketched in Figure S6 of the Supporting Information, are very similar to those of HBDI and consequently their vibronic spectra closely resemble that of the parent molecule, with a main absorption peak and a major shoulder at higher energy, whose intensity is about four times lower than the main peak (Figure 7 and Figure S7 in Supporting Information). In Supporting Information we report the assignments of the main bands of 35Bu showing that, as for HBDI, Duschinsky mixings have a weak effect on the spectrum and HT effects are negligible. We further confirmed that also for 35Bu different DFT hybrid functionals predict very similar spectra. From the assignment of the modes, we conclude that the FC active modes that dominate the vibrational progression are not affected by the substitutions and that the modes contributing to the shoulder are also almost identical in all analogues. Therefore, the substitutions have only a minor effect on the absorption lineshape in vacuum.

Although the substitutions hardly affect the lineshape, the absolute position of the spectra depends strongly on the substitution pattern. The position of the absorption maximum increases with the size of the substitution and we find the following trend: HBI ($\sim 3.15 \text{ eV}$) > HBDI ($\sim 3.10 \text{ eV}$) > 35 Me ($\sim 3.04 \text{ eV}$) > 35 Bu (3.01 eV) \approx 35 BuBz ($\sim 3.0 \text{ eV}$). This trend is in qualitative agreement with the experimental trend observed in

ethanol: HBDI (~ 2.80 eV) > 35 Me (~ 2.60 eV) > 35 Bu (2.44 eV) \approx 35 BuBz (~ 2.40 eV).^[8]

The effect of the temperature is also very similar for all HBDI derivatives. Increasing the temperature causes a progressive broadening of all peaks. Therefore, at higher temperatures, the relative intensity of the shoulder (~ 1500 cm⁻¹) increases with respect to that of the maximum (Figure 7). Most of the broadening can be attributed to vibrational excitations in the ground state of modes that are similar to mode 3 of HBDI (depicted in Figure 6). Furthermore, the temperature dependent broadening of the main peaks increases with the size of the alkyl substituents. Whereas the valley between the main peak and the shoulder are clearly visible at all temperatures for the smallest analogue HBI, it completely disappears in the largest analogue 35BuBz, which also has the most intense shoulder at the highest temperature considered here (300 K). One explanation for these observations is that the dimensionless displacement along the relevant mode (i.e., mode 3 in HBDI) increases with substituent size. For example, in HBDI it is 1.3, whereas in 35Bu, where this is mode 4, it has increased to 1.7 and thus leads to a larger vibrational progression (compare Figure 5 and Figure S5 in the Supporting Information). These modes correspond to an in-plane bending of the methine bridge. Although the equilibrium angles in S_0 (132.2 degrees) and S_1 (128.1 degrees) are the same in both chromophores, the displacement is larger along this mode in 35Bu, because bending this angle also induces a linear displacement of the bulky *t*Bu substituents. We observed very similar effects for the other low-frequency modes, but because these modes are only weakly FC active, they do not affect the spectrum.

3. Conclusions

We have computed, analysed and compared the vibronic spectra of the isolated GFP chromophore at various levels of ab initio theory and with several approaches. We found that CAM-B3LYP with a 6-31G(d) basis set offers the best trade-off in terms of accuracy and computational cost. The CAM-B3LYP spectrum is in good agreement with the spectrum computed at the CASPT2 and XMCQDPT2 levels of theory. We demonstrated that optimizing the structures of the ground and excited state at different levels of theory, or using a non-correlated method, such as CASSCF, introduces inaccuracies in the $S_0 \rightarrow S_1$ geometric displacements. We showed that even small errors in the displacements can cause artificial absorption peaks, which broaden the spectrum. The apparent agreement of previous spectra with the experimental action spectra, which are much broader than our CAM-B3LYP, CASPT2 and XMCQDPT2 spectra, may therefore have been fortuitous. Further evidence that the action spectrum is too broad to be considered an absorption spectrum, is provided by the good agreement between the measured emission spectra in the protein at 253 K and our computed emission spectra in vacuum, both of which are much narrower. Finally we investigated the effect of substitutions at various positions on the HBDI chromophore, and found that these have only very minor effects on the vibronic absorption spectra.

Acknowledgements

M.D.D. thanks Mr. Kaveh Haghighi Mood (RWTH University Aachen) and F.S. thank Professor Nadia Rega (University of Naples) for fruitful discussions. Professors Stephen R. Meech, Kyril M. Solntsev, Rebeca Jockusch, Anna Krylov and Dr. Anastasia Bochenkova are acknowledged for providing the experimental and simulated spectra in the electronic form. This work was partially supported by Italian MIUR (PRIN 2010–2011 2010ERFKXL). F.A. acknowledges support from the Marie Curie COFUND programme U-Mobility, co-financed by the University of Malaga, the European Commission FP7 under GA No. 246550, and Ministerio de Economía y Competitividad (COFUND2013-40259) G.G. is supported by the Academy of Finland.

Keywords: absorption • chromophores • computational chemistry • green fluorescent protein • vibrational spectroscopy

- [1] R. Y. Tsien, *Annu. Rev. Biochem.* **1998**, 67, 509–544.
- [2] M. Zimmer, *Chem. Rev.* **2002**, 102, 759–781.
- [3] L. M. Tolbert, A. Baldrige, K. Janusz, K. M. Solntsev, *Acc. Chem. Res.* **2012**, 45, 171–181.
- [4] K. Nienhaus, F. Renzi, B. Vallone, J. Wiedenmann, G. U. Nienhaus, *Bio-phys. J.* **2006**, 91, 4210–4220.
- [5] R. M. Wachter, M. Elsliger, K. Kallio, G. T. Hanson, S. J. Remington, *Structure* **1998**, 6, 1267–1277.
- [6] V. Adam, K. Nienhaus, D. Bourgeois, G. U. Nienhaus, *Biochemistry* **2009**, 48, 4905–4915.
- [7] K. Suto, H. Masuda, Y. Takenaka, F. I. Tsuji, H. Mizuno, *Genes Cells* **2009**, 14, 727–737.
- [8] J. Conyard, M. Kondo, I. A. Heisler, G. Jones, A. Baldrige, L. M. Tolbert, K. M. Solntsev, S. R. Meech, *J. Phys. Chem. B* **2011**, 115, 1571–1577.
- [9] F. J. Avila Ferrer, M. D. Davari, D. Morozov, G. Groenhof, F. Santoro, *ChemPhysChem* **2014**, 15, 3246.
- [10] B. O. Roos, P. R. Taylor, *Chem. Phys.* **1980**, 48, 157–173.
- [11] K. Andersson, P. Malmqvist, B. O. Roos, A. J. Sadlej, K. Wolinski, *J. Phys. Chem.* **1990**, 94, 5483–5488.
- [12] K. Andersson, P. Malmqvist, B. O. Roos, *J. Chem. Phys.* **1992**, 96, 1218–1226.
- [13] A. A. Granovsky, *J. Chem. Phys.* **2011**, 134, 214113–214115.
- [14] F. J. A. Ferrer, F. Santoro, *Phys. Chem. Chem. Phys.* **2012**, 14, 13549–13563.
- [15] A. D. Becke, *Phys. Rev. A* **1988**, 38, 3098–3100.
- [16] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, 37, 785–789.
- [17] A. D. Becke, *J. Chem. Phys.* **1993**, 98, 1372–1377.
- [18] A. D. Becke, *J. Chem. Phys.* **1993**, 98, 5648–5652.
- [19] C. Adamo, V. Barone, *J. Chem. Phys.* **1999**, 110, 6158–6170.
- [20] J. Chai, M. Head-Gordon, *Phys. Chem. Chem. Phys.* **2008**, 10, 6615–6620.
- [21] T. Yanai, D. P. Tew, N. C. Handy, *Chem. Phys. Lett.* **2004**, 393, 51–57.
- [22] W. J. Hehre, R. Ditchfield, J. A. Pople, *J. Chem. Phys.* **1972**, 56, 2257–2261.
- [23] P. C. Hariharan, J. A. Pople, *Theor. Chem. Acc.* **1973**, 28, 213–222.
- [24] T. H. Dunning, *J. Chem. Phys.* **1989**, 90, 1007–1023.
- [25] F. Santoro, *FCclasses*, a Fortran 77 code: available at <http://village.pi.ic-com.cnr.it/Software>, last accessed May 10, 2014.
- [26] F. Santoro, A. Lami, R. Improta, J. Bloino, V. Barone, *J. Chem. Phys.* **2008**, 128, 224311–224317.
- [27] F. Santoro, A. Lami, R. Improta, V. Barone, *J. Chem. Phys.* **2007**, 126, 184102–1841011.
- [28] F. Santoro, R. Improta, A. Lami, J. Bloino, V. Barone, *J. Chem. Phys.* **2007**, 126, 084509.
- [29] F. J. Avila Ferrer, J. Cerezo, F. Santoro, in preparation **2014**.
- [30] R. Ianculescu, E. Pollak, *J. Phys. Chem. A* **2004**, 108, 7778–7784.
- [31] Q. Peng, Y. Niu, C. Deng, Z. Shuai, *Chem. Phys.* **2010**, 370, 215–222.
- [32] R. Borrelli, A. Capobianco, A. Peluso, *J. Phys. Chem. A* **2012**, 116, 9934–9940.

- [33] A. Baiardi, J. Bloino, V. Barone, *J. Chem. Theory Comput.* **2013**, *9*, 4097–4115.
- [34] Gaussian 09 (Revision C.3), M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Men-
nucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian,
A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara,
K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O.
Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro,
M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Ko-
bayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyen-
gar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B.
Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann,
O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin,
K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg,
S. Dapprich, A. D. Daniels, Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski,
D. J. Fox, Gaussian Inc. Wallingford CT, **2009**.
- [35] F. Aquilante, L. De Vico, N. Ferré, G. Ghigo, P.-A. Malmqvist, P. Neogrády,
T. B. Pedersen, M. Pitonák, M. Reiher, B. O. Roos, L. Serrano-Andres, M.
Urban, V. Veryazov, R. Lindh, *J. Comput. Chem.* **2010**, *31*, 224–247.
- [36] A. A. Granovsky, Firey version 8.0.0.
- [37] A. Charaf-Eddin, A. Planchat, B. Mennucci, C. Adamo, D. Jacquemin, *J.
Chem. Theory Comput.* **2013**, *9*, 2749–2760.
- [38] D. Jacquemin, E. Bremond, A. Planchat, I. Ciofini, C. Adamo, *J. Chem.
Theory Comput.* **2011**, *7*, 1882–1892.
- [39] M. Uppsten, B. Durbec, *J. Comput. Chem.* **2012**, *33*, 1892–1901.
- [40] E. Kamarchik, A. I. Krylov, *J. Phys. Chem. Lett.* **2011**, *2*, 488–492.
- [41] N. H. List, J. M. Olsen, T. Rocha-Rinza, O. Christiansen, J. Kongsted, *Int. J.
Quantum Chem.* **2012**, *112*, 789–800.
- [42] E. Stendardo, F. A. Ferrer, F. Santoro, R. Improta, *J. Chem. Theory
Comput.* **2012**, *8*, 4483–4493.
- [43] F. J. A. Ferrer, J. Cerezo, E. Stendardo, R. Improta, F. Santoro, *J. Chem.
Theory Comput.* **2013**, *9*, 2072–2082.
- [44] P. Schellenberg, E. Johnson, A. P. Esposito, P. J. Reid, W. W. Parson, *J.
Phys. Chem. B* **2001**, *105*, 5316–5322.
- [45] X. He, A. F. Bell, P. J. Tonge, *J. Phys. Chem. B* **2002**, *106*, 6056–6066.
- [46] P. Altoe, F. Bernardi, M. Garavelli, G. Orlandi, F. Negri, *J. Am. Chem. Soc.*
2005, *127*, 3952–3963.
- [47] A. P. Esposito, P. Schellenberg, W. W. Parson, P. J. Reid, *J. Mol. Struct.*
2001, *569*, 25–41.
- [48] M. Almasian, J. Grzetic, G. Berden, B. Bakker, W. Jan, J. Oomens, *Int. J.
Mass Spectrom.* **2012**, *330–332*, 118–123.
- [49] M. W. Forbes, A. M. Nagy, R. A. Jockusch, *Int. J. Mass Spectrom.* **2011**,
308, 155–166.
- [50] T. Tanabe, M. Saito, K. Noda, *EPJ D* **2011**, *62*, 191–195.
- [51] K. Chingin, R. M. Balabin, V. Frankevich, K. Barylyuk, R. Nieckarz, P. Sagu-
lenko, R. Zenobi, *Int. J. Mass Spectrom.* **2011**, *306*, 241–245.
- [52] M. W. Forbes, R. A. Jockusch, *J. Am. Chem. Soc.* **2009**, *131*, 17038–
17039.
- [53] G. Bublit, B. A. King, S. G. Boxer, *J. Am. Chem. Soc.* **1998**, *120*, 9370–
9371.
- [54] S. S. Stavrov, K. M. Solntsev, L. M. Tolbert, D. Huppert, *J. Am. Chem. Soc.*
2006, *128*, 1540–1546.
- [55] D. Vitkup, D. Ringe, G. A. Petsko, M. Karplus, *Nat. Struct. Mol. Biol.* **2000**,
7, 34–38.
- [56] A. V. Bochenkova, L. H. Andersen, *Faraday Discuss.* **2013**, *163*, 297–319.
- [57] Y. Ai, G. Tian, Y. Luo, *Mol. Phys.* **2013**, *111*, 1316–1321.
- [58] M. E. Martin, F. Negri, M. Olivucci, *J. Am. Chem. Soc.* **2004**, *126*, 5452–
5464.
- [59] J. Dong, K. M. Solntsev, L. M. Tolbert, *J. Am. Chem. Soc.* **2006**, *128*,
12038–12039.

Received: May 20, 2014

Published online on September 1, 2014