J Mol Model (2007) 13:775–783
DOI 10.1007/s00894-007-0194-7

ORIGINAL PAPER

Vibrational analysis of a solvated green fluorescent protein chromophore

Tadeusz Andruniów

Received: 23 November 2006 / Accepted: 27 February 2007 / Published online: 23 March 2007

Abstract Resonance Raman (RR) spectra of green fluorescent protein (GFP) model chromophores in solution have been simulated with the CASSCF/MM methodology. Although several reports on vibrational analysis of GFP model chromophores have been recently published, the RR spectra were simulated for the first time in *explicit* solution with the inclusion of the counterion, as these effects are crucial for unambiguously reproducing the vibrational band assignment in the anionic form of the GFP chromophore. This strategy allows for a one-to-one correspondence of the calculated vibrational modes to the observed RR bands, concerning both the location and intensity pattern. In addition, these simulations were complemented with total energy distribution calculations to aid in the unambiguous assignment of the measured spectra. The current study helps to clarify some of the previous RR bands assignments as well as producing some new assignment for the anionic form of GFP chromophore. The *explicit* solvent simulations and PCM-based calculations are compared to the measured spectra, and these results demonstrate that *explicit* solvent simulations provide better agreement with experiment, both in terms of vibrational frequencies and intensity distribution.

Keywords GFP chromophore · Resonance Raman · Vibrational analysis · CASSCF/MM

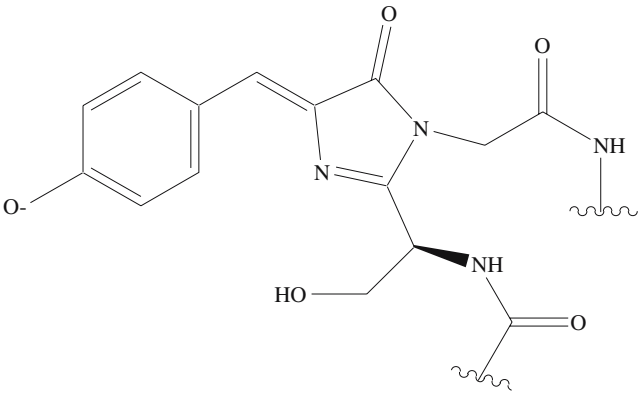
Introduction

The green fluorescent protein (GFP) was first discovered in jellyfish *Aequorea victoria* in 1962 as a companion protein to aequorin [1]. In this organism, GFP serves as the ultimate light emitter of the native bioluminescence system [2, 3]. The chromophore: 4-(p-hydroxybenzylidene) imidazolin-5-one [4, 5] (see Scheme 1), generated in an auto-catalytic posttranslational cyclization of residues Ser65-Tyr66-Gly67 [6], is located in the center of the folded protein [7, 8]. GFP and its mutants have found many applications in cell and molecular biology as a protein localization marker and can be used to monitor the expression of genes [9–11].

The absorption spectrum of wild-type GFP contains two absorption maxima at 395 and 475 nm that have been attributed to the neutral and anionic forms of the chromophore, respectively [12, 13]. These two protonation states are commonly known as states A (395 nm) and B (475 nm) [10, 14–16], respectively. Upon illumination of either A or B at room temperature the thermodynamically unstable intermediates are formed with almost identical emission maxima near 500 nm [14]. The photoconversion of A into B via A* is not an efficient process. In contrast, the excited state decay that converts A* into I* is a very rapid and efficient process that involves a proton transfer from the phenolic group of the chromophore along a hydrogen-bonding network to the carboxylate oxygen of the Glu222 residue. The green emission will follow to the I ground state, which can either revert back to the ground state A or rarely evolve to species B, although on a much longer timescale. As a result, the I state is structurally similar to both the ground state and excited state of species A, while electronically similar to the ground state and excited state of species B [10, 14–16].

The absorption maxima for the neutral and anionic forms of the model chromophore: ethyl 4′-hydroxybenzylidene-2-methyl-imidazolinone-3-acetate (HBMIA) [17, 18] and 4′-hydroxybenzylidene-2,3-dimethyl-imidazolinone (HBDI) [19, 20] (Scheme 2), resemble the maxima seen in the A and B forms of the protein. The fluorescence spectrum of GFP reveals only one peak, attributed within the three

T. Andruniów (✉)
Institute of Physical and Theoretical Chemistry,
Wroclaw University of Technology,
Wyb. Wyspianskiego 27,
Wroclaw 50-370, Poland
e-mail: ANDRUNIOW@MML.CH.PWR.WROC.PL

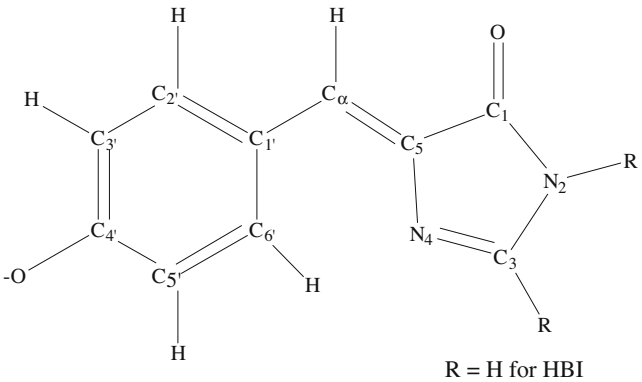


Scheme 1 4-(p-hydroxybenzylidene) imidazolin-5-one

state model [8, 10, 14], to the I state featuring the anionic form of the chromophore. In contrast to GFP, which is highly fluorescent, model chromophores do not fluoresce at room temperature, but become fluorescent at low temperatures (e.g., frozen ethanol solution at 77 K) [17].

Recently, wild-type GFP, as well as its chromophore analogue HBDI in ethanol, has been the subject of a resonance Raman spectroscopic investigation [21], which is an extremely useful tool for probing electronic excitations in the Franck–Condon vicinity, since the intensity of the vibrational modes coupled to the electronic excitation are enhanced in relation to other modes. Moreover, RR spectra provide detailed information on both the ground and excited state molecular structures of the chromphore embedded in protein. In fact, Parson and his coworkers [19, 21] suggested that the initial relaxation out of the Franck–Condon region is due to the 1,556 cm^{-1} vibrational mode, which is strongly coupled to the electronic excitation.

A few theoretical studies have been published to date on the vibrational analysis of HBDI and HBMIA in different protonation states. On the basis of normal mode assignments in FTIR and Raman spectra of these GFP chromophore analogues, using density functional (DFT) and Hartree–Fock



Scheme 2 p-hydroxybenzylidene-imidazolin-5-one (HBI) and 4′-hydroxybenzylidene-2,3-dimethyl-imidazolinone (HBDI)

(HF) theories [19–23], it has been demonstrated that the most intensive modes correspond to an in-plane skeletal deformation of the chromophore structure, in particular normal coordinates dominated by C=C and C=N stretchings. The disagreement as to the assignment of some of the key modes may reflect the fact that calculations on model chromophores were performed in vacuo rather than in solution.

Solvent effects are vital for a reliable characterization of vibronic structures due to the fact that solvent–solute interactions may cause structural changes in ground state and excited states as well as a redistribution of charges. More recent studies by Altoe et al. [24] revealed that while solvent-induced effects, introduced through a polarizable continuum model (PCM) [25–27], have only minor influence on the Raman and RR spectra of neutral and cationic forms of HBDI, these effects have a rather dramatic impact on the off-resonance and resonance Raman spectra of anionic species. High level ab initio calculations (complete active space self-consistent field—CASSCF) combined with *implicit* treatment of the solvent, led to an improvement of the RR spectra as compared to gas-phase simulations and gave quite satisfactory agreement with the experimental spectra. There are still some major discrepancies, however, between the theoretical and experimental spectra which need to be addressed. On the other hand, it was recognized by He et al. [20] that calculated isotopic shifts for anionic HBDI in the gas phase do not match experimental values very well, and including the counterion and solvent effects is crucial for unambiguously reproducing the vibrational band assignment in the anionic form of the GFP chromophore.

In the present work, we report the vibrational analysis of the anionic form of the model chromophore p-hydroxybenzylidene-imidazolin-5-one (HBI) (Scheme 2), extending previous vacuo- and PCM-based models by embedding the HBI compound within an *explicit* water solvation box including a sodium counterion. Our ultimate goal is to revisit, with the aid of computational tools, the assignment of vibrational bands in the RR spectra of solvated HBI analogue of the GFP chromophore. By comparing the RR spectra calculated in water, represented by both *explicit* and *implicit* solvent models, and in the gas phase, we investigate the influence of the counterion and *explicit* solvent on the RR spectra of the GFP chromophore. We hope this analysis will also provide the means for an accurate and detailed interpretation of the vibrational spectra of the GFP protein in the future.

Materials and methods

In our previous papers, we have demonstrated that CASPT2//CASSCF/MM calculations in *explicit* solution

provide a quantitative evaluation of structural and spectroscopic parameters for 11-cis retinal [28] and the HBI model chromophore [29]. Also, CASSCF/MM preliminary studies on RR spectrum of HBI in *explicit* solution showed good agreement with the experimental spectrum [29]. Therefore, we have utilized CASSCF, an accurate but time-consuming computational method, to simulate the RR spectra of the HBI model system along with normal mode composition analysis and aid in the assignement of experimental bands.

The HBI model differs from the full chromophore in that hydrogen atoms replace the methyl groups connecting the chromophore to the protein. This truncation, however, retains the conjugated π-electron system of the chromophore, which is especially important in excited state calculations. In order to investigate a GFP model chromophore in *explicit* solvent, one needs to combine a high level ab initio description of the chromophore with a MM description of the solvent. Van der Waals parameters for the chromophore were taken from the GAFF force field [30]. MM charges account for polarization effects in a mean-field way and are the same for both ground and excited state computations. The ab initio QM calculations are based on the CASSCF/6-31G* level. The active space comprises the truncated π-system of HBI (12 electrons in 11 π-orbitals), since its complete π system (16 electrons in 14 π-orbitals) is out of reach for the CASSCF methodology at the moment. Accordingly, the nitrogen π-lone pair and the lowest occupied and highest unoccupied benzene π-orbitals were excluded from the active space. CASSCF/6-31G*/MM geometry optimization of the ground state and optically allowed excited state HBI's structures, as well as frequency calculations in the ground state, were carried out with the programs Gaussian 03 [31] and Tinker [32].

A model of the GFP chromophore in solution (water) was constructed by placing the ion-pair HBI (including sodium treated at the MM level as the counterion) in a rectangular box of TIP3P water molecules positioned within 10 Å from any given atom of the chromophore, using the xleap module of Amber package [33]; this procedure generated 701 water molecules. To neutralize the system, we added a sodium cation. The resulting system consisted of 2,125 atoms. The average ground state configuration of the solvent was determined according to the following scheme. The solvent was minimized for 2,000 steps using the steepest descent method while keeping the HBI solute fixed in its gas-phase configuration. The partial charges of the chromophore atoms employed in our QM/MM simulations were determined with Gaussian03, using a Restrained ElectroStatic Potential (RESP) [34] procedure at the HF/6-31G* level of theory. In the next step, we performed a CASSCF/6-31G*/MM geometry optimization to relax the coordinates of the QM chromophore, counterion, and all water molecules within 4.5 Å of any atom in HBI (45 water molecules forming the first and second water shells were optimized at the MM level). The positions of the remaining 656 solvent molecules, more distant from the chromophore, were kept frozen during the QM/MM calculations.

To quantify the effect of the solvent on the RR spectra, we have also performed CASSCF/6-31G* calculations with the inclusion of the PCM solvent model. Accordingly, we were able to compare the simulated RR spectra obtained within PCM and *explicit* solvation schemes with the experimental spectrum. The vibrational spectra were simulated in water. Resonance Raman intensities were simulated according to the Franck–Condon model, successfully used for a number of charged molecules [35–37]. In particular, this model assumes that the ground and excited state vibrational manifolds, formed by N harmonic symmetric oscillators, are represented by normal coordinates Q. These oscillators are displaced in the excited state relative to the ground state, but their frequencies remain the same in both. Within such an approach, we obtain, for each i^{th} totally symmetric mode, Franck–Condon displacement parameters relative to the $S_0 \rightarrow S_1$ transition using the following expression:

$$B_{x,i} = \mathbf{S}_i^{T}\mathbf{F}(S1, S0)\left(\frac{\mu_i \omega_i}{h}\right)$$

where $\mathbf{S}_i$ is a $3N$-dimensional vector containing the Cartesian nuclear displacements in the Q_i normal mode with the associated frequency ω_i and reduced mass μ_i. $\mathbf{F}$(X, G) is the $3N$-dimensional vector containing the differences of the nuclear position vectors in the first excited state and the ground state.

To improve the agreement between the calculated and experimental frequencies, we used scaling factor for CASSCF-based (0.90) frequencies to account for errors due to the incomplete treatment of electron correlation, basis set truncation and anharmonic effects. This scaling factor was adjusted to accurately reproduce the position of the most intense band located at ca. 1,555 cm^{-1} [21]. All spectra were obtained as superpositions of the Lorentzian curve with the line width 10 cm^{-1}. The interpretation of RR spectra requires a description of the normal modes, thus total energy distribution (TED) analysis was performed to derive normal mode composition.

Results and discussion

Explicit versus *implicit* solvent models in simulations of RR spectra for anionic HBI

Altoe et al. [24] have simulated both static and resonance Raman spectra for the anionic form of the HBDI model

chromophore with the inclusion of solvent effects through the PCM method (no counterion was considered). In an attempt to correlate the calculated vibrational modes with the observed Raman bands [20], they ascribed some of the computed vibrational modes to the observed Raman bands. Interestingly, these authors could not find any such correlation for the resonance Raman spectra as the quality of the simulated spectrum was not satisfactory enough to allow for a one-to-one correspondence with the measured spectrum.

Figure 1 shows calculated RR spectra for the anionic form of HBI calculated in vacuo and in solution compared to the experimental RR spectrum of HBDI in ethanol measured with an excitation energy of 368.9 nm [21]. Solvent effects were taken into account through the PCM method and by *explicitly* treating solvent at the MM level of theory. As one can see, the predicted RR spectrum simulated in vacuo is substantially different from the observed spectrum. The largest enhancement is predicted for the 1,351 cm^{-1} mode, in disagreement with experiment, and can be assigned to the exocyclic C-H rocking and CαC5 stretching as well as imidazolinone C-N stretching. Moreover, a rather intense mode at 1,580 cm^{-1} can be characterized by phenol CC stretchings coupled to exocyclic CαC5 stretch.

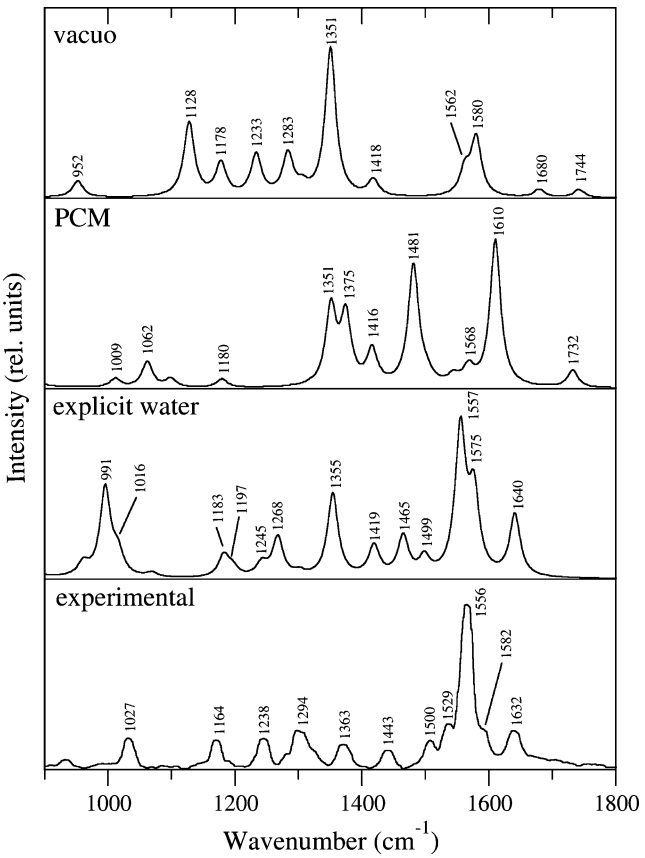


Fig. 1 Resonance Raman spectra for the anionic form of the HBI chromophore simulated (CASSCF/6-31G*/MM) in vacuo, in *implicit* water (with PCM), and in *explicit* water compared to the corresponding experimental spectrum [21]

An improvement of both intensities and vibrational frequencies in the RR spectrum is obtained upon including a PCM model into the CASSCF calculations. The most intense vibration simulated with the PCM approach is 1,610 cm^{-1}, which on the basis of TED analysis can be described as a mode dominated by C=N stretching. Interestingly, the intensity of the 1,568 cm^{-1} mode, having also a large contribution from the imidazolinone C=N bond stretching, is very weak. The second strongest vibration at 1,481 cm^{-1} can be ascribed to phenol CC stretchings coupled to the inter-ring CαC5 stretching deformation. The correspondence of calculated and experimental spectral patterns is also far from being perfect in the lower frequency region. In fact, the intensities of the dominant vibrations in the 1,300–1,450 cm^{-1} region are largely overestimated with respect to experiment. On the other hand, in the 1,100–1,300 cm^{-1} spectral region, the vibrations bear essentially no intensities, with the exception of a single feature at 1,180 cm^{-1}. This analysis suggests that the inclusion of solvent effects largely influences the calculated vibrational frequencies and intensities and also the shape of the associated normal coordinates.

Altoe et al. [24] have already shown that the inclusion of solvent effects considerably improves the agreement with the experimental spectrum [21] by dramatically changing exocyclic CC bond distances upon electronic $S_0 \rightarrow S_1$ excitation. Indeed, while in the gas phase the C1′Cα (single bond in the ground state) and CαC5 (double bond in the ground state) bond lengths change upon electronic excitation by 0.049 and 0.040 Å, respectively, in PCM-based calculations these bonds change their distances by 0.002 and 0.067 Å (−0.003 and 0.074 Å for the HBDI molecule [24]), respectively. This effect is even more pronounced in *explicit* solvent calculations which shorten the C1′Cα bond and lengthen the CαC5 bond upon the $S_0 \rightarrow S_1$ excitation by −0.030 and 0.080 Å, respectively. Consequently, various environments (gas phase, *implicit* and *explicit* water) do influence, to a different extent, the exocyclic CC bond lengths upon excitation, implying different RR intensity enhancements. Vibrational modes enhanced most strongly in the RR spectra are those which have a large contribution of the CαC5 bond stretching coordinate.

Consistently with previous results [24], an analysis of PCM-based Mulliken charges (Fig. 2) indicates that ca. 35% of the negative charge initially located on the phenolic $C–O^-$ is transfered to inter-ring region (5%) and imidazolinone ring (30%) upon excitation, leading to a large change in dipole moment ($\Delta\mu$=9.3 D). Similarly, in *explicit* solvent calculationsca. 37% of the negative charge moves away from the phenol group to inter-ring region (15%) and imidazolinone ring (21%). A slightly different

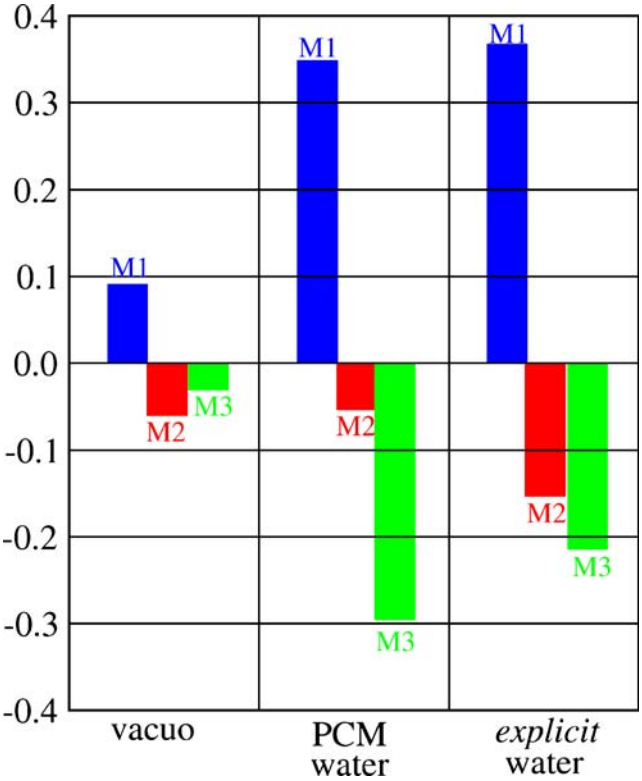


Fig. 2 Magnitude of the charge transfer induced by the $S_0 \rightarrow S_1$ transition evaluated as the change in the Mulliken charges of the p-hydroxyphenate (*M1*), inter-ring region (*M2*) and imidazolinone ring (*M3*). Results obtained from in vacuo and PCM-corrected CASSCF/6-31G* calculations as well as from *explicit* solvent CASSCF/6-31G*/MM calculations

distribution of charges in the HBI molecule is obtained upon $S_0 \rightarrow S_1$ excitation with the *explicit* and PCM models, resulting in the fact that the former method reproduces the experimental change in dipole moment much more accurately. In fact, agreement with the experimental value ($\Delta\mu$= 6.8 D) [38] is remarkable, as the calculated $\Delta\mu$ (6.6 D) is only 0.2 D off the measured value. In complete contrast, in vacuo $\Delta\mu$ is only 1.0 D ($\Delta\mu$=2.0 D for HBDI) due to a very limited charge transfer upon $S_0 \rightarrow S_1$ excitation (9%).

To evaluate the performance of our *explicit* solvent model in relation to the PCM-based solvent model in simulating RR spectra, we decided to compare both sets of calculated vibrational frequencies with the experimental RR transitions. Least-square fits were used to correlate the experimental [21] and theoretical vibrational frequencies. The linear correlations between CASSCF and experimental RR vibrational frequencies are shown in Fig. 3. The values of slope and intercept (ideal correlation would be with a unit slope and zero intercept) reflect the quality of the approach used to reproduce the experimental vibrational frequencies. It is evident from looking at Fig. 3 that the correlation obtained with the *explicit* solution is better than the one from the PCM *implicit* solution. Our *explicit* calculations produce a regression coefficient far closer to unity than the corresponding PCM method. Consequently, our calculated vibrational frequencies match experimental values quite well. In fact, out of nine modes depicted in Fig. 3 seven of them are within 20 cm^{-1} (five of them are within 10 cm^{-1}) of the experimental values and only two modes corresponding to the observed bands at 1,238 and 1,443 cm^{-1} display larger shifts, 30 and 24 cm^{-1}, respectively.

The PCM solvent model does not seem to work as well as the *explicit* model. In fact, there are numerous vibrational frequencies with very significant discrepancies between experiment [21] and theory. The largest deviations, 68 and 65 cm^{-1}, are found for the modes corresponding to the 1,443 and 1,164 cm^{-1} experimental bands. Other modes were calculated with smaller, but still significant, errors between 12–58 cm^{-1}, and only the 1,500 cm^{-1} vibration is in perfect agreement with its experimental counterpart. Also, as will be shown later, the *explicit* solvent model outperforms the *implicit* solvent model in reproducing the experimental intensity pattern.

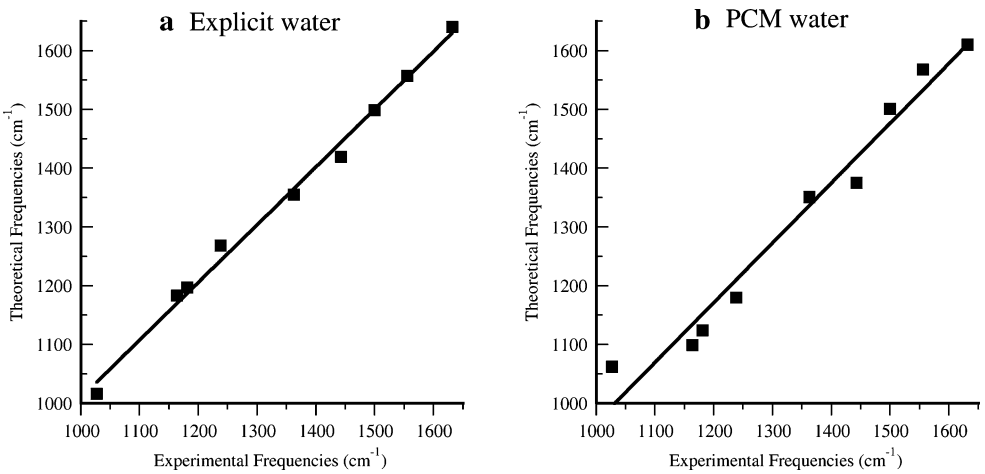


Fig. 3 **a** Correlation of *explicit* hydration calculations (CASSCF/6-31G*/MM) for the HBI model chromophore and experimental RR data [21]; slope=0.982, intercept=27.210 and regression coefficient= 0.997. **b** Correlation of *implicit* PCM calculations (CASSCF/6-31G*) for the HBI model chromophore and experimental RR data [21], slope=1.017, intercept=−48.838 and regression coefficient=0.984

This rather modest performance of the PCM method compared to our *explicit* model is not surprising when bearing in mind that one of the well known shortcomings of dielectric continuum methods is that they cannot account for strong short-range interactions between ionic solutes and solvents, for example hydrogen bonding [39, 40]. To overcome this problem and improve the accuracy of PCM or any other *implicit* calculations, one would have to use a combined strategy by incorporating several *explicit* solvent molecules in the continuum calculations (in some cases one *explicit* water molecule is enough to improve the results) [39, 40].

Resonance Raman spectra of anionic HBI

Figure 1 compares the experimental RR spectrum [21] with those calculated in vacuo and solution. Due to photodegradation, the Raman spectra were not measured in resonance. For this reason, in principle all the bands observed in the RR spectrum are also observed in the off-resonance Raman spectrum. However, the pattern of relative intensities is slightly different. As one can see in Fig. 1 the calculated bands are in one-to-one accord with those in the experiment, although there are minor differences in the positions and intensities of the calculated CASSCF/MM and experimental RR bands. The inspection of the associated normal-mode contributions, included in Table 1, provides an assignment of the vibrational features in the GFP chromophore.

The highest frequency mode of the RR spectrum presented in Fig. 1 is a skeletal stretch corresponding to the C=O bond. Experimentally, this transition is strongly IR-active and weakly Raman-active, with a band at 1,665 cm^{-1}, and is not observed in the RR spectrum [21]. The calculated C=O stretching frequency is found at 1,740 cm^{-1} and thus is considerably higher than the above-mentioned experimental value. This discrepancy may partly reflect the reduced active space in the CASSCF calculation and consequently unbalanced electron correlation as well as the modest basis set. Indeed, the off-resonance Raman spectrum simulated on the basis of B3LYP/MM calculations, which in contrast to the CASSCF scheme include dynamic correlation energy, reveals this feature at 1,637 cm^{-1} (unpublished data) in reasonable agreement with the experimental value of 1,665 cm^{-1} [21].

The 1,640 cm^{-1} mode likely corresponds to the band observed at 1,632 cm^{-1} and can be assigned to a mode that involves the stretching of the exocyclic CαC5 bond and phenol C–O and CC bonds stretches. Experimental spectrum reveals a shoulder (1,582 cm^{-1}) to the high-energy side of the strongest RR transition, which has its counterpart in the simulated spectrum at 1,575 cm^{-1}. The main internal coordinates contributing to this mode, called 'phenol 2' by Esposito et al. [19], are the CαC5 exocyclic stretch and the C–O stretch in phenol, as well as the C=N stretch in the imidazolinone group. Thus, the 1,575 cm^{-1} peak is delocalized over two rings, which may be the reason for the lack of sensitivity observed in the isotope labeling off-resonance Raman experiments performed by He et al. [20].

In accordance with experiment, the most prominent RR band is found at 1,557 cm^{-1} (observed at 1,556 cm^{-1} in the experimental spectrum). Careful inspection of the normal-mode displacement vector associated with this mode and TED analysis (Table 1) reveals that this mode is delocalized over both rings and inter-ring region and involves mainly the stretching of the C=N bond and to a lesser extent phenol CC stretching motions. This finding is in line with the isotope labeling off-resonance Raman studies by Tonge group [20].

In the high frequency region between 1,000 and 1,650 cm^{-1} most of the peaks of the calculated spectrum match the peaks in the experimental RR spectrum. Intense features calculated at 1,640, 1,575, 1,499, 1,419, 1,355, 1,268, 1,245, 1,197, 1,183 and 1,016 cm^{-1} seem to correspond to bands seen in the HBDI experimental RR spectrum at 1,632, 1,582, 1,500, 1,433, 1,363, 1,294, 1,238, 1,181, 1,164 and 1,027 cm^{-1} [21], respectively. The intensities of the calculated peaks are much stronger relative to the 1,556 cm^{-1} band than in the experimental spectrum.

The intensity pattern in the 1,400–1,650 cm^{-1} region is nicely reproduced for modes dominated by phenol stretchings except for the 1,529 cm^{-1} feature, for which it seems to be underestimated. Moreover, the 1,529 cm^{-1} mode is upshifted by ca. 30 cm^{-1} in the simulated spectrum, reflecting the exaggerated contribution of the C=O coordinate. In addition, we found an involvement of the phenolic ring in the following modes: the moderately intense feature at 1,499 cm^{-1} is dominated by phenolic ring stretchings while the band at 1,419 cm^{-1} involves the C–O stretch and C–H in-plane bendings in agreement with an earlier assignment [18]. On the other hand, the intensities of some features, characterized by a large contribution of the C=N stretching mode, are overestimated, e.g., 1,575 and 1,355 cm^{-1}.

The spectral region 1,200–1,400 cm^{-1} reveals three rather intense bands at 1,355, 1,268, and 1,245 cm^{-1}, which probably have their experimental counterparts at 1,363, 1,294, and 1,238 cm^{-1}, respectively. The 1,363 cm^{-1} band has been reported before but not assigned [20]. Our calculations show that it has a noticeable contribution from the Cα=C5-N4-C3 fragment of the HBI molecule and may be assigned to a normal mode having major contribution from inter-ring vibrations, mainly Cα-H bending motion and CαC5 stretching. Similarly, the peak at 1,268 cm^{-1}, which may either correspond to the 1,238 cm^{-1} or 1,298 cm^{-1} feature in the measured spectrum, does arise solely from inter-ring and imidazolinone stretching motions. This assignment seems to be in agreement with

Table 1 CASSCF/6-31G*/MM vibrational modes between 600 and 1,800 cm^{-1} for anionic HBI

Frequency[a] (cm^{-1})	Description[b]
1,740	C1=O im str (77)[c]
1,640	C4′-O phe str (22), CαC5 inter str (19), C2′C3′ phe str (10), C5′C6′ phe str (10)
1,575	CαC5 inter str (30), C4′-O phe str (15), C3=N4 im str (11)
1,563	C4′-O phe str (23), C5′C6′ phe str (12), C3=N4 im str (10), C2′C3′ phe str (10)
1,557	C3=N4 im str (42)
1,499	C2′-H phe i.p. bend (14), C1′C6′ phe str (13), C3′C4′ phe str (12), C1′C2′ phe str (11), C4′C5′ phe str (11)
1,465	C2′C3′ phe str (32), C5′C6′ phe str (26), C3′-H phe i.p. bend (12)
1,419	C4′-O phe str (23), C2′-H phe i.p. bend (15), C3′-H phe i.p. bend (13), C5′-H phe i.p. bend (13), C6′-H phe i.p. bend (10)
1,395	N-H im i.p. bend (68), C3-N2 im str (18)
1,355	Cα-H inter i.p. bend (48), C5-N4 im str (18), CαC5 inter str (10)
1,301	C3-H im i.p. bend (36), C3=N4 im str (16)
1,288	C3-H im i.p. bend (15), C2′-H phe i.p. bend (11), C6′-H phe i.p. bend (10), C5′-H phe i.p. bend (10)
1,268	C1′Cα inter str (28), C1–C5 im str (19)
1,245	C3′-H phe i.p. bend (20), C3′C4′ phe str (17), C4′C5′ phe str (15)
1,197	C5′-H phe i.p. bend (26), C6′-H phe i.p. bend (25), C5′C6′ phe str (17)
1,183	C2′-H phe i.p. bend (29), C3′-H phe i.p. bend (20), C2′C3′ phe str (10), C1′Cα inter str (10)
1,112	C1′C2′ phe str (18), C6′-H phe i.p. bend (10), C5′-H phe i.p. bend (10)
1,069	C3–N2 Im str (53), N2-H im i.p. bend (13), C3-H im i.p. bend (13)
1,030	C6′-H phe o.p. bend (100)
1,016	C1–N2 im str (32), C5–N4 im str (16), C3–N2 im str (10)
991	C2′-H phe o.p. bend (70), C3′-H phe o.p. bend (23)
960	Cα-H inter o.p. bend (90)
941	Phe i.p. def (38), C2-H phe i.p. bend (15), C1′C2′ phe str (11)
897	C3-H im o.p. bend (98)
876	im i.p. def (49), CαC5 inter str (24)
838	C2′-H phe o.p. bend (48), C3′-H phe o.p. bend (48)
814	C1N2 im str (14), im i.p. def (15), C5′-H phe o.p. bend (11), C1′C6′ phe str (10), C1′C2′ phe str (10)
808	C3′C4′ phe str (24), C4′C5′ phe str (17), C-H phe o.p. bend (16)
805	C5′-H phe o.p. bend (14), C3′-H phe o.p. def (14)
782	C5–N4 im str (31), C1′CαC5 inter i.p. bend (12)
746	C1=O im o.p. bend (68), im o.p. def (26)
731	Phe o.p. def (65), C4′-O phe o.p. bend (23)
702	N2-H im o.p. bend (76)
660	im i.p. def (32)
632	im o.p. def (55)
626	phe i.p. def (21), C4′-O phe i.p. bend (16), im i.p. def (10)
581	phe i.p. def (24), C1=O im i.p. bend (22)

[a] CASSCF frequencies were uniformly scaled by 0.9
[b] *phe* Phenyl, *im* imidazolanone, *inter* inter-ring, *str* stretching, *i.p.* in-plane, *o.p.* out-of-plane, *def* deformation, *bend* bending mode
[c] The number in parentheses refers to the percent contribution each mode has on the frequency. Only modes contributing more than 10% are listed.

earlier studies by He et al. [20], who have shown that peak at ca. 1,240 cm^{-1} has a contribution from stretchings of bridging bonds. In other reports [18, 41], however, these authors proposed this band to be attributed to the C–O stretching mode, which seems to contradict with our findings since we did not detect any vibration in the 1,200–1,300 cm^{-1} region having a noticable contribution of the C–O stretching coordinate.

A pair of experimental bands at 1,164 and 1,181 cm^{-1} can readily be attributed to the simulated bands at 1,183 and 1,197 cm^{-1}. Although their calculated vibrational frequencies are underestimated in comparison to the corresponding RR bands by 19 and 16 cm^{-1}, respectively, their energy separation of 17 cm^{-1} is in excellent agreement with the experimental value (14 cm^{-1}). Also, the relative intensity pattern of these two bands in the simulated spectrum is reversed with respect to the experimental spectrum. On the basis of computational data, the 1,164 and 1,181 cm^{-1} bands can be confidently assigned to the phenol C–H in-plane bendings weakly coupled to the phenol CC stretches. The former band has also some contribution from the exocyclic CαC5 bond stretching.

The vibrational mode calculated at 1,016 cm^{-1} and probably corresponding to the RR band located at

1,033 cm^{-1} involves substantial contributions from C–N stretching modes, in contradiction with the experimental studies by He et al. [20] who characterized this mode as an inter-ring C1′Cα stretching coupled with carbonyl bending, based on the fact that this band is shifted by 10 cm^{-1} to lower wavenumbers upon 1-^{13}C isotope labeling.

Below 1,000 cm^{-1}, the spectrum is dominated by out-of-plane modes as well as by modes being more delocalised and thus difficult to characterize in terms of local coordinates. The calculated strong feature at 991 cm^{-1} involves phenol out-of-plane C–H bending motions. Features near 800 cm^{-1} are ascribed to in-plane phenol deformations that are heavily mixed with CαC5 bridge stretching. A distinctive feature in this spectral region, visible at 814 cm^{-1}, is due to imidazolinone phenol stretchings mixed with deformations of the two rings. The next-highest-frequency mode at 838 cm^{-1} can be ascribed to out-of-plane C-H bendings. On the other hand, the 808 cm^{-1} mode can be described as phenolic ring 'breathing'. A 626 cm^{-1} mode, found in the simulated spectrum of the HBI molecule and matching the 601 cm^{-1} experimental RR band, can be described as an in-plane deformation of both rings. The activity of this mode does support the hypothesis of Tozzini and Nifosi [22], who proposed that stretching modes are coupled to the low-frequency angular deformation of both rings. This would explain the vibrational feature observed at 591 cm^{-1} in ultrafast pump-and-probe experiments [42].

Concluding remarks

The present work provides the first vibrational analysis of the solvated green fluorescent protein chromophore in *explicit* water with the addition of a counterion. CASSCF/MM *explicit* solvent vibrational frequency calculations complemented with total energy distribution analysis allowed for a detailed and accurate assignment of RR peaks and outperformed (at least for the anionic GFP chromophore) the standard PCM approach. This study revisited previously assigned RR bands and proposed some new assignments for the anionic chromophore of GFP.

In conclusion, we have demonstrated that our CASSCF/MM *explicit* solvent methodology was successful in reproducing the RR spectra, as well as absorption and fluorescence spectra (absorption and fluorescence spectra calculations were reported in our previous paper [29]) of model GFP chromophores in solution and may prove useful in extending current knowledge on how the protein matrix modulates the spectral and photophysical properties of the chromophore.

Acknowledgment The author acknowledges Prof. M. Olivucci for many insightful discussions and critical reading of the manuscript.

References

1. Shimomura O, Johnson FH, Saiga Y (1962) J Cell Comp Physiol 59:223–239
2. Morin JG, Hastings JW (1971) J Cell Physiol 77:313–318
3. Johnson FH, Shimomura O, Saiga Y, Gershman LC, Reynolds GT, Waters JR (1962) J Cell Comp Physiol 60:85–103
4. Prasher DC, Eckenrode VK, Ward WW, Prendergast FG, Cormier MJ (1992) Gene 111:29–233
5. Cody CW, Prasher DC, Westler WM, Prendergast FG, Ward WW (1993) Biochemistry 32:1212–1218
6. Cubitt AB, Heim R, Adams SR, Boyd AE, Gross LA, Tsien RY (1995) Trends Biochem Sci 20:448–455
7. Yang F, Moss LG, Phillips Jr GN (1996) Nature Biotech 14:1246–1251
8. Brejc K, Sixma TK, Kitts PA, Kain SR, Tsien RY, Ormo M, Remington SJ (1997) Proc Natl Acad Sci USA 94:2306–2311
9. Tsien RY (1998) Annu Rev Biochem 67:509–544
10. Zimmer M (2002) Chem Rev 102:759–781
11. Tsien RY, Prasher DC, Chalfie M, Kain SE (eds) (1998) Wiley–Liss, New York, pp 97–118
12. Ward WW, Prentice HJ, Roth AF, Cody CW, Reeves SC (1982) Photochem Photobiol 35:803–808
13. Heim R, Prasher DC, Tsien RY (1994) Proc Natl Acad Sci USA 91:12501–12504
14. Chattoray M, King BA, Bublitz GU, Boxer SG (1996) Proc Natl Acad Sci USA 93:8362–8367
15. Lossau H, Kummer A, Heinecke R, Pollinger-Dammer F, Kompa C, Bieser G, Jonsson T, Silva CM, Yang MM, Youvan DC, Michel-Beyerle ME (1996) Chem Phys 213:1–16
16. Weber W, Helms V, McCammon JA, Langhoff PW (1999) Proc Natl Acad Sci USA 96:6177–6182
17. Niwa H, Inouye S, Hirano T, Matsuno T, Kojima S, Kubota M, Ohashi M, Tsuji FI (1996) Proc Natl Acad Sci USA 93:13617–13622
18. Bell AF, He X, Wachter RM, Tonge PJ (2000) Biochemistry 39:4423–4431
19. Esposito AP, Schellenberg P, Parson WW, Reid PJ (2001) J Mol Struct 569:25–41
20. He X, Bell AF, Tonge PJ (2002) J Phys Chem B 106:6056–6066
21. Schellenberg P, Johnson E, Esposito AP, Reid PJ, Parson WW (2001) J Phys Chem B 105:5316–5322
22. Tozzini V, Nifosi R (2001) J Phys Chem B 105:5797–5803
23. Yoo HY, Boatz JA, Helms V, McCammon JA, Langhoff PW (2001) J Phys Chem B 105:2850–2857
24. Altoe P, Bernardi F, Garavelli M, Orlandi G, Negri F (2005) J Am Chem Soc 127:3952–3963
25. Cossi M, Barone V, Cammi R, Tomasi J (1996) Chem Phys Lett 255:327–335
26. Mennucci B, Tomasi J (1997) J Chem Phys 106:5151–5158
27. Miertus S, Scrocco E, Tomasi J (1981) Chem Phys 55:117–123
28. Andruniów T, Ferré N, Olivucci M (2004) Proc Natl Acad Sci USA 101:17908–17913
29. Sinicropi A, Andruniow T, Ferré N, Basosi R, Olivucci M (2005) J Am Chem Soc 127:11534–11535
30. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) J Comp Chem 25:1157–1173
31. Gaussian 03, Revision A.1, Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery Jr JA, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ,

Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, P. Gill MW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA (2003) Gaussian, Pittsburgh, PA

32. Ponder JW (2004) TINKER: Software Tools for Molecular Design, Version 4.2. Washington University School of Medicine, St. Louis
33. Case DA, Pearlman DA, Caldwell JW, Cheatham III TE, Wang J, Roos WS, Simmerling CL, Darden TA, Merz KM, Stanton RV, Cheng AL, Vicent JJ, Crowley M, Tsui V, Gohlke H, Radmer RJ, Duan Y, Pitera J, Massova I, Seibel GL, Singh UC, Weiner PK, Kollman PA (2002) AMBER 7, University of California, San Francisco
34. Fox T, Kollman PA (1998) J Phys Chem B 102:8070–8079
35. Andruniow T, Zborowski K, Pawlikowski M (1996) Chem Phys Lett 259:193–198
36. Andruniow T, Pawlikowski M (2000) Chem Phys Lett 321:485–490
37. Andruniow T, Pawlikowski M, Zgierski MZ (2000) J Phys Chem A 104:845–851
38. Bublitz G, Brett AK, Boxer SG (1998) J Am Chem Soc 120:9370–9374
39. Chipman DM (2002) J Phys Chem A 106:7413–7422
40. Kelly CP, Cramer CJ, Truhlar DG (2005) J Chem Theory Comput 1:1133–1152
41. Bell AF, Stoner-Ma D, Wachter RM, Tonge PJ (2003) J Am Chem Soc 125:6919–6926
42. Cinelli RAG, Tozzini V, Pellegrini V, Beltram F, Cerullo G, Zavelani-Rossi M, De Silvestri D, Tyagi M, Giacca M (2001) Phys Rev Lett 86:3439–3442