Biophysical Chemistry 145 (2009) 1–6

Contents lists available at ScienceDirect

Biophysical Chemistry

journal homepage: http://www.elsevier.com/locate/biophyschem

On photoabsorption of the neutral form of the green fluorescent protein chromophore

Igor Topol [a,*], Jack Collins [a], Igor Polyakov [b], Bella Grigorenko [b], Alexander Nemukhin [b,c]

[a] *Advanced Biomedical Computing Center, Information Systems Program, SAIC-Frederick Inc., NCI-Frederick, Frederick, MD, 21702-1201, USA*
[b] *Chemistry Department, M.V. Lomonosov Moscow State University, 1/3 Leninskie Gory, Moscow, 119991, Russia*
[c] *N.M. Emanuel Institute of Biochemical Physics, Russian Academy of Sciences, 4 Kosygina, Moscow, 119334, Russia*

ARTICLE INFO

Article history:
Received 1 June 2009
Received in revised form 6 August 2009
Accepted 6 August 2009
Available online 14 August 2009

Keywords:
Green fluorescent protein
Photoabsorption spectra
Quantum calculations
Protein matrix influence

ABSTRACT

We present results of theoretical studies of the photoabsorption band corresponding to the vertical electronic transition S_0–S_1 between first two singlet states of the model chromophore from the green fluorescent protein (GFP) in its neutral form. Predictions of quantum chemical approaches including *ab initio* and semi-empirical approximations are compared for the model systems which mimic the GFP chromophore in different environments. We provide evidences that the protein matrix in GFP accounts for a fairly large shift of about 40 nm in the S_0–S_1 absorption band as compared to the gas phase.

1. Introduction

In spite of wide applications of the famous green fluorescent protein (GFP) in biology and medicine [1–3] its photophysical properties are still not completely understood. Multiple efforts, both experimental and theoretical, are being undertaken to study absorption and emission processes in the GFP-related model systems ranging from isolated chromophores in the gas phase to numerous wild-type and mutated proteins from the GFP family. The GFP chromophore apparently may appear in the protein and in solution in different protonation forms, two of which, anionic (Fig. 1a) and neutral (Fig. 1b), attract the greatest attention of researchers.

In the wild-type GFP, the neutral form of the chromophore absorbs light at ~400 nm and the deprotonated anionic form absorbs at ~480 nm. Excitation at both wavelengths leads to fluorescence emission at 510 nm. A large shift of emission after 400 nm excitation is explained by the excited state proton transfer along the hydrogen bond chains in the protein connecting the phenolic and imidazolinone sites in the chromophore molecule [4,5].

Estimates of the absorption spectra of biological chromophores in the gas phase provide important information on the role of protein environment in tuning chromophore's spectral properties. Following experimental studies performed at the heavy-ion storage ring, Andersen and co-authors reported the absorption maxima at 479 nm for the anionic GFP chromophore [6,7]. Also an absorption maxima at 406 nm was attributed for the cationic species in which (compared to the neutral form) the proton was attached to the imidazolinone nitrogen. Recently, the same group reported a maximum absorption at 415 nm for the synthetic chromophore called "neutral$^+$" [8] shown in Fig. 2. This charged molecule which carries a positive charge in the $-NH_3$ group, presumably well separated from the "neutral" fragment of the chromophore, was supposed to mimic spectral properties of the gas phase GFP chromophore in its neutral form. Upon correcting the measured value of 415 nm for the presence of the charged group by using the results of quantum chemistry calculations (TDDFT with the B3LYP/6-311++G(d) approximation at the MP3 optimized geometry configuration) for the "neutral$^+$" and the true neutral species the authors finally reported the absorption maximum for the gas phase neutral GFP chromophore at 399 nm [8]. Since both gas phase values (399 nm for neutral and 479 nm for anionic) were found to coincide with those known for the chromophore inside the protein matrix (400 nm and 480 nm, respectively) the conclusion was that "the absorption properties of the green fluorescent protein to a high degree are determined by the intrinsic chromophore properties" [8]. Therefore, the role of the protein environment was suggested to be negligible.

The latter conclusion may be challenged following the results of careful quantum chemical calculations for a series of model systems which include the anionic and neutral forms of the GFP chromophore. It should be noted that accurate *ab initio* treatment of the vertical S_0–S_1 excitation in the GFP-like chromophores is still a hard task for quantum chemistry, and it is difficult to expect theoretical errors less than 15 nm in this range of 400–500 nm [9]. We do not concentrate in this paper on the anionic species which is a subject of many thorough theoretical studies [9–17], but mention that the agreement between

---

* Corresponding author. Tel.: +1 301 846 4275; fax: +1 301 846 5762.
E-mail addresses: topol@ncifcrf.gov, topoli@mail.nih.gov (I. Topol).

0301-4622/$ – see front matter
doi:10.1016/j.bpc.2009.08.002

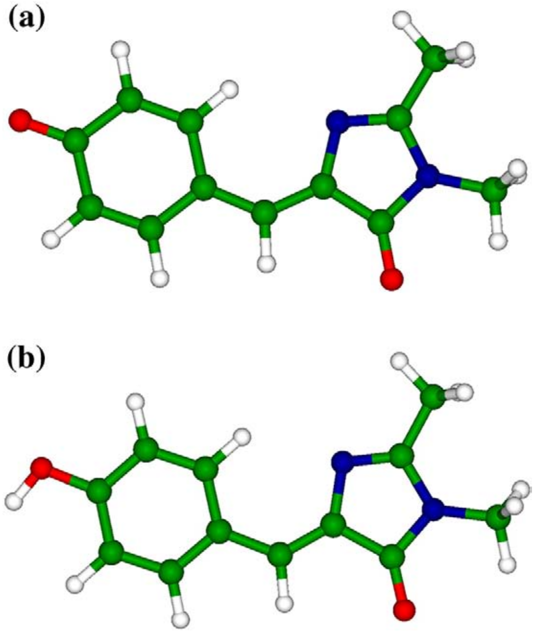

Fig. 1. Anionic (a) and neutral (b) forms of the GFP chromophore in the cis-conformation. Here and in other figures, carbon atoms are shown in green, oxygen in red, nitrogen in blue. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the best theoretical estimates for the gas phase GFP chromophore and the experimental estimate of 479 nm [7,8] are consistent within the claimed error bars of ±15 nm. However, this is not the case for the neutral chromophore. In this case we compare the previously and newly obtained theoretical values for the several model systems which mimic the neutral GFP chromophore and provide evidences that the protein matrix in GFP accounts for a fairly large shift in the S_0–S_1 absorption band compared to the gas phase.

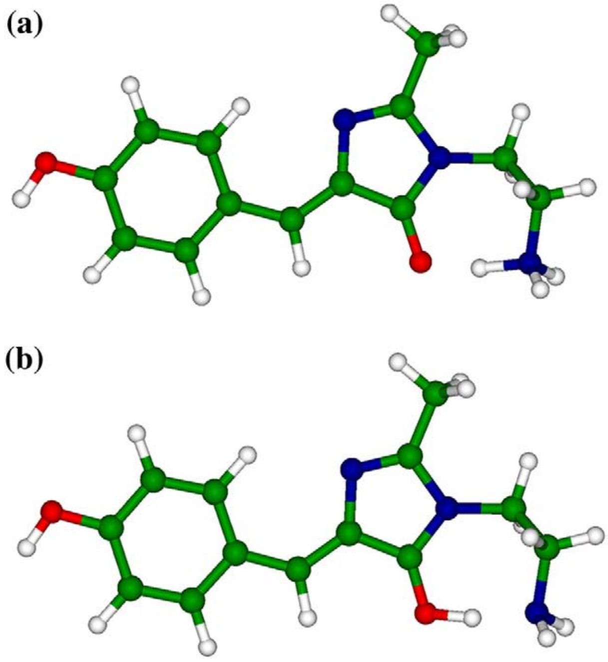

Fig. 2. Structures of the "neutral$^+$" chromophore. Panel (a) — the structure assumed in Ref. [8], panel (b) — the lower energy structure with the proton re-located from the NH_3^+ group to the carbonyl oxygen.

2. Computational methods

We considered either the isolated chromophore species as shown in Figs. 1 and 2, or the molecular clusters constructed on the base of the coordinates of heavy atoms from the crystal structure PDBID:1EMG [18]. To prepare the cluster systems (as one illustrated in Fig. 3) we surrounded the chromophore molecule by the nearest amino acid residues and water molecules, and kept the coordinates of their C_α atoms frozen as in the crystal structure upon optimization of all other coordinates.

In majority of calculations, the equilibrium geometry parameters of the model systems in the ground electronic state S_0 were obtained in the density functional theory (DFT) approach. For the isolated chromophore molecules (Fig. 1) we used either the B3LYP/6-31+G(d,p) or PBE0/cc-pVDZ approximation. When considering the "neutral$^+$" structure (Fig. 2) the MP2/6-311+G(d,p) approximation was also applied. Ground state geometry parameters of the molecular clusters were optimized by using the B3LYP/6-31+G(d,p) method.

To compute vertical excitation energies from the respective minima on the S_0 potential surfaces of the model systems we considered several quantum chemical approaches. Vast amount of the results presented below was obtained with the semi-empirical ZINDO method [19]. As shown here and elsewhere [16,20] a strategy to estimate the S_0–S_1 excitation energies at the DFT-optimized equilibrium geometry parameters performs perfectly for these model systems. Another inexpensive approach is the time-dependent DFT (TDDFT) approximation [21]. From the *ab initio* quantum chemical side, it is feasible to perform calculations of the vertical $\pi\pi^*$ transition energies for the isolated chromophore molecule by using a very expensive and highly correlated approach, the state-averaged CASSCF wave functions augmented by perturbative corrections: multireference second-order Møller–Plesset perturbation theory (MRMP2) [22] and versions of the multiconfigurational quasidegenerate perturbation theory (MCQDPT2) [23–25]. These techniques, however, are computationally demanding, and their execution requires advanced skills and extreme care, as the application of the method involves: (i) a careful selection of a large number of active space orbitals in fairly large basis sets; (ii) converging the state- veraged CASSCF solutions corresponding to the $\pi\pi^*$ transition, especially in realistic basis sets; (iii) a careful and often ambiguous treatment of perturbative corrections to the reference CASSCF solutions.

Here, the B3LYP, MP2, TDDFT, ZINDO calculations have been carried out with the Gaussian03 program [26]. Calculations in the PBE0, MRMP2 and various versions of MCQDPT2 have been performed with the PC GAMESS program [27].

3. Results and discussion

(i) Ground state geometry configurations.

Computed ground state equilibrium geometry parameters of the anionic and neutral forms of the GFP chromophore (Fig. 1a,b) are consistent with the results of previous theoretical studies [14–17,28]. In the original paper of Lammich et al. [8] describing the gas phase "neutral$^+$" species (Fig. 2a), the authors reported the results of MP3 and B3LYP/6-31+G(d) calculations of its structural parameters showing, in particular, a strong hydrogen bonding between the carbonyl oxygen of the imidasole ring and one of the ammonium hydrogens. More accurate calculations performed in this work apparently indicate that another structure with the proton attached to oxygen, but not to ammonium nitrogen (Fig. 2b) in fact refers to the lowest energy potential energy minimum. According to the MP2/6-311+G(d,p) results the energy of the latter (Fig. 2b) is 0.5 kcal/mol lower than that of the initially suggested structure "neutral$^+$" (Fig. 2a). Apparently, the excitation energy of the structure shown