Conformation dependence of pK\textsubscript{a}'s of the chromophores from the purple asFP595 and yellow zFP538 fluorescent proteins

A.V. Nemukhin, I.A. Topol, B.L. Grigorenko, A.P. Savitsky, J.R. Collins

\textsuperscript{a} Department of Chemistry, M.V. Lomonosov Moscow State University, Leninskie Gory, 1/3, Moscow 119991, Russian Federation
\textsuperscript{b} N.M. Emanuel Institute of Biochemical Physics, Russian Academy of Sciences, ul. Kosygina, 4, Moscow 119334, Russian Federation
\textsuperscript{c} Advanced Biomedical Computing Center, Advanced Technology Program, SAIC-Frederick Inc., NCI-Frederick, Frederick, MD 21702-1201, USA
\textsuperscript{d} A.N. Bach Institute of Biochemistry, Russian Academy of Sciences, Leninsky prospekt, 33, Moscow 119071, Russian Federation

\begin{abstract}
Two members of the green fluorescent protein family, the purple asFP595 and yellow zFP538 proteins, are perspective fluorescent markers for use in multicolor imaging and resonance energy-transfer applications. We report the results of quantum based calculations of the solution pK\textsubscript{a} values for selected protonation sites of the denatured asFP595 and zFP538 chromophores in the trans- and cis-conformations in order to add in the interpretation of photo-physical properties of these proteins. The pK\textsubscript{a} values were determined from the thermodynamic cycle based on B3LYP/6-311++G(\textit{d},\textit{p}) calculations of gas phase free energies of the molecules and the B3LYP/6-311++G(\textit{d},\textit{p}) calculations of solvation energies. The results show that the pK\textsubscript{a}'s of the protonation sites of the chromophore from asFP595 noticeably depend on the isomer conformation (cis- or trans-), while those of zFP538 are much less sensitive to isomerization.
\end{abstract}

1. Introduction

The properties of chromophores from the extended family of the green fluorescent protein (GFP) are being intensively investigated because of great importance of GFP-like biomarkers in living cells [1]. In this work we focus on two members of the family: the purple chromoprotein from the sea anemone \textit{Anemonia sulcata}, asFP595 [2], and the yellow chromoprotein from button polyp \textit{Zoanthus}, zFP538 [3], which are perspective fluorescent markers for use, in particular, in multicolor imaging and resonance energy-transfer applications.

The fluorescent protein asFP595, which is the Ala143Gly mutant of the wild type chromoprotein asCP, is famous for its unique photo-physical properties. The protein asFP595 is initially nonfluorescent, but in response to intense green light irradiation at 568 nm, it becomes brightly fluorescent (kindles) with emission at 595 nm. Irradiation of the emitting protein form with blue light quenches fluorescence. The photoswitching properties of this kindling fluorescent protein may be also useful for information storage applications. A large amount of experimental data [4–9] indicate that the mechanism of kindling may be related to the trans- to cis-photoisomerization of the chromophore. The most relevant direct observations are due to Andresen et al. [7] who presented the crystal structures of the Ala143Ser mutant of asCP. In the dark state of these crystals, the chromophore was in the trans-form, but after irradiation the chromophore could be visualized in the cis-configuration.

Unlike asFP595, the protein zFP538 is not known as a photoswitching species. Absorption and fluorescence of zFP538 containing a three-ring chromophore [3] correspond to the wavelengths intermediate between fluorescence of green proteins and absorption of red proteins. This property allows one to use zFP538 for the resonance energy transfer processes important in analytical applications.

The chromophores of asFP595 and zFP538 may be considered as derivatives of the chromophore from the parent protein GFP, 4-hydroxybenzylidene-imidazolinone (HBI), shown in Fig. 1 in the anionic form. Two most important protonation sites of the molecule refer to oxygen of the phenolic ring (O\textsubscript{phen}) and the imidazoline nitrogen atom (N\textsubscript{imid}).

The molecules of chromophores from asFP595 (Fig. 2) and zFP538 (Fig. 3) possess extra \pi-electron conjugation accounting for absorption and emission shifts to longer wavelengths compared to GFP. The structures of the chromophores from asFP595 and zFP538 are basically known from the corresponding crystals, however, some uncertainties still remain. In particular, an alternative
interpretations of X-ray experiments resulted either in the carbonyl group (as shown in the upper panel of Fig. 1) [5] or the iminic group (lower panel of Fig. 2) [6,7] in the asFP595 chromophore.

Many details of the fluorescent properties of these proteins remain unknown in spite of substantial efforts to understand the chemistry and physics of chromophore excitations. Specifically, the role of trans–cis isomerization of the asFP595 chromophore inside the protein is still under debate [7,10–12], although the working hypothesis assumes this trans–to cis-transformation. If such isomerization stimulated by photoexcitation actually occurs in the protein, it should proceed with an active involvement of the neighboring amino acid side chains. In particular, protons from these residues may be attached to electronegative sites of the chromophore modulating its spectral properties. Therefore it is important to compare proton affinities and respective pK\textsubscript{a} values of different sites of the chromophore in various conformations. The trans- and cis-forms of the chromophore are the first candidates for studies of conformational dependence of pK\textsubscript{a}'s.

Previously, the protonation sites at the phenolic oxygen atom and the imidazoline nitrogen atom of the model chromophore (Fig. 1) of the parent GFP have been experimentally [13] and theoretically [14–18] studied for their relative acidity (pK\textsubscript{a}). El Yazal et al. [17] calculated the theoretical pK\textsubscript{a} values for all possible protonation sites of the chromophore in denatured GFP chromophores. Scharnagl and Raupp-Kossmann [18] also reported theoretical pK\textsubscript{a}'s for GFP from hybrid quantum-classical calculations with an average error of 0.8 units. In particular, the authors computed the constant pK\textsubscript{a} = 8.3 for neutral/anion equilibria in excellent agreement with the experimental data (pK\textsubscript{a} = 8.2) on solution-phase acid–base dissociation constants of the model GFP chromophore [13]. No comparison of pK\textsubscript{a}'s for the original cis- and hypothetical trans-forms of the GFP chromophore were performed.

The extended chromophores (Figs. 2 and 3) of the primary interest of this work contain additional protonation sites and one might expect certain changes in the pK\textsubscript{a} values compared to those of GFP. The experimental data for asFP595 and zFP538 are rare: (1) The denatured asFP595 chromophore in aqueous solution was studied in the work of Yampolsky et al. [19], but no experimentally based estimates of pK\textsubscript{a}'s were reported. (2) The seeming pK\textsubscript{a} constants of the zFP538 chromophore in solution were evaluated from the absorption and fluorescent spectra giving the values pK\textsubscript{1} = 5.6–5.7 and pK\textsubscript{2} = 7.4–7.8 [20]. From theoretical side, the only estimate of the zwitterionic imidazoline NH proton of the asFP595 chromophore in aqueous solution by using the density functional theory calculations which resulted in the pK\textsubscript{a} values 4.7 and 9.1 for the cis- and trans-forms, respectively, was described in Ref. [11].

In order to provide better understanding the photo-physical phenomena with the fluorescent proteins asFP595 and zFP538 we performed accurate quantum chemistry calculations of the solution pK\textsubscript{a} values for selected protonation sites of the asFP595 and zFP538 chromophores in the trans- and cis-conformations. We follow the strategy combining the correlated quantum chemistry technique and continuum solvation models through the thermodynamic cycle, which is capable to determine the pK\textsubscript{a} values accurate within half of a pK\textsubscript{a} unit [21,22].
2. Calculation details

Our calculations use the thermodynamic cycle shown in Scheme 1, leading to the well-known formula

\[
pK_a = (2.303RT)^{-1} \left\{ \Delta G_x + \Delta G_y(A^-) - \Delta G_y(\text{AH}) + \Delta G_y(H^+) \right\}
\]

(1)

In Scheme 1 and formula (1), \(\Delta G_x\) and \(\Delta G_y\) are free energies of deprotonation in the gas phase and solvent, respectively. \(\Delta G_y(\text{AH})\) and \(\Delta G_y(H^+)\) are solvation free energies of anion \(A^-\), protonated compound \(\text{AH}\), and proton \(H^+\), respectively. Optimization of equilibrium geometry parameters and calculations of harmonic frequencies for the gas-phase species were performed in the density functional theory B3LYP/6-311+G(d,p) approximation by using the Gaussian-03 program [23]. Corrections to the free energy changes for the gas-phase part of the reaction in Scheme 1 were obtained in the single point calculations in the B3LYP/6-311++G(2df,2p) approximation. Solvation energies \(\Delta G_y(\text{AH})\) and \(\Delta G_y(H^+)\) were estimated at the B3LYP/6-311++G(d,p) level by using the Jaguar program [24].

The proton solvation free energy \(\Delta G_y(H^+)\), the last term in Eq. (1) is not known with high precision. Based on previous calculations of pK\(_a\) values of substituted imidazoles [21], the value \(-262.5\) kcal/mol [22] was for a long time considered as the best estimate. However, recently Palascak and Schields [25] suggested an even lower value \(-264.0\) kcal/mol. In this work we used both estimates and found that although the absolute pK\(_a\)s may differ by up to one pK\(_a\) unit, the results are qualitatively independent of \(\Delta G_y(H^+)\). Below we cite only the values calculated using the value \(-264.0\) kcal/mol.

3. Results

The results of calculations for all model species are summarized in Table 1.

3.1. Model GFP chromophore (Fig. 1)

In the protein, the cis-conformation of the GFP chromophore occurs and no evidences of the trans-form have been reported so far.

Table 1

<table>
<thead>
<tr>
<th>Species</th>
<th>(pK_a)</th>
<th>(N_{\text{side}})</th>
<th>(N_{\text{main}})</th>
<th>(N_{\text{chain}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>GFP, cis</td>
<td>8.53</td>
<td>5.30</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>GFP, trans</td>
<td>7.27</td>
<td>3.58</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Anionic zFP538, cis</td>
<td>9.06</td>
<td>4.52</td>
<td>7.02</td>
<td>-</td>
</tr>
<tr>
<td>Anionic zFP538, trans</td>
<td>9.09</td>
<td>4.38</td>
<td>7.39</td>
<td>-</td>
</tr>
<tr>
<td>Neutral asFP595, cis</td>
<td>-</td>
<td>-0.18</td>
<td>4.13</td>
<td>-</td>
</tr>
<tr>
<td>Neutral asFP595, trans</td>
<td>-</td>
<td>2.33</td>
<td>4.21</td>
<td>-</td>
</tr>
<tr>
<td>asFP595(NH), cis</td>
<td>6.96</td>
<td>1.29</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>asFP595(NH), trans</td>
<td>8.38</td>
<td>4.63</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>asFP595(NH), NH(_i)</td>
<td>8.14</td>
<td>3.22</td>
<td>7.57</td>
<td>-</td>
</tr>
<tr>
<td>asFP595(NH), NH(_i), trans</td>
<td>7.69</td>
<td>4.74</td>
<td>-6.77</td>
<td>-</td>
</tr>
</tbody>
</table>

The cis-trans isomerization of a model GFP chromophore, 4-hydroxybenzylidene-1,2-dimethylimidazoline (HBDL), was studied experimentally in aqueous solution [26].

Our computed pK\(_a\) value on the phenolic oxygen site in the cis-conformation (8.5) agrees with the experimental (8.2 [13]) and previous theoretical (8.2 [17], 8.3 [18]) results within the generally accepted uncertainties of pK\(_a\) calculations. Here, we notice the obtained changes in pK\(_a\)’s due to cis-trans isomerization, 1.3 \(\pm\) 1.7 pK\(_a\) units for the phenolic oxygen and the imino nitrogen of the imidazoline ring.

3.2. zFP538 Chromophore (Fig. 3)

In calculations for the zFP538 chromophore, we first added protons to the different sites of anionic species; the corresponding pK\(_a\)’s are indicated in Table 1 in the rows ‘Anionic zFP538’. Remarkably, the computed constants are practically the same for both (cis- and trans-) conformers. Next set of calculations was carried out for the possible neutral chromophore structure for the asFP595 chromophore in which the conjugated \(\pi\)-system of the chromophore is extended by the chain with the carbonyl group (upper panel in Fig. 2). However, Wilmann et al. [6] and Andresen et al. [7] placed the imino group instead of the carbonyl group into the protein matrix when protons from the nearby amino acid side chains are transferred to nitrogen atoms \(N_{\text{side}}\) (or \(N_{\text{chain}}\) in case of the asFP595(NH) form). The present calculations show that...
the proton affinity of these particular chromophore atoms strongly depends on the isomer conformation affecting the location of the proton along the corresponding hydrogen bond. Even recognizing that the pKₐ values inside the protein may noticeably differ from those in solution the tendencies in both condense phases are believed to be similar. Study of the mechanism of kindling in asFP595 including the trans-cis chromophore isomerization by using the QM/MM approach is a subject of current research in our laboratories.

Remarkably, calculations for the chromophore from zFP538 do not show noticeable differences in solution pKₐ’s for the trans–cis–forms. Only if the phenolic oxygen is protonated the properties of the imidazole nitrogen are affected (~0.2 for the cis-form vs 2.3 for the trans-form). As mentioned in Section 1, no evidence has been reported that the trans–cis isomerization affects the fluorescent properties of this protein. The only experimental data for the seeming pKₐ constants evaluated directly from the absorption and fluorescent spectra of the zFP538 chromophore in solution are as follows: pK₁ ≈ 5.6–5.7 and pK₂ ≈ 7.4–7.8 [20]. Apparently, the computed pKₐ’s for the nitrogen sites (4.4–4.5 and 7.0–7.4) correlate well with these experimental estimates.

Acknowledgements

This work was partly supported by the Science and Innovation Federal Russian Agency (project #02.522.11.2002), by the Program for Molecular and Cell Biology from Russian Academy of Sciences, and by the grant from the Russian Foundation for Basic Research (project #07-03-00059). We thank the staff and administration of the Advanced Biomedical Computing Center for their support of this project. This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under contract N01-CO-12400. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organization imply endorsement by the U.S. Government.


