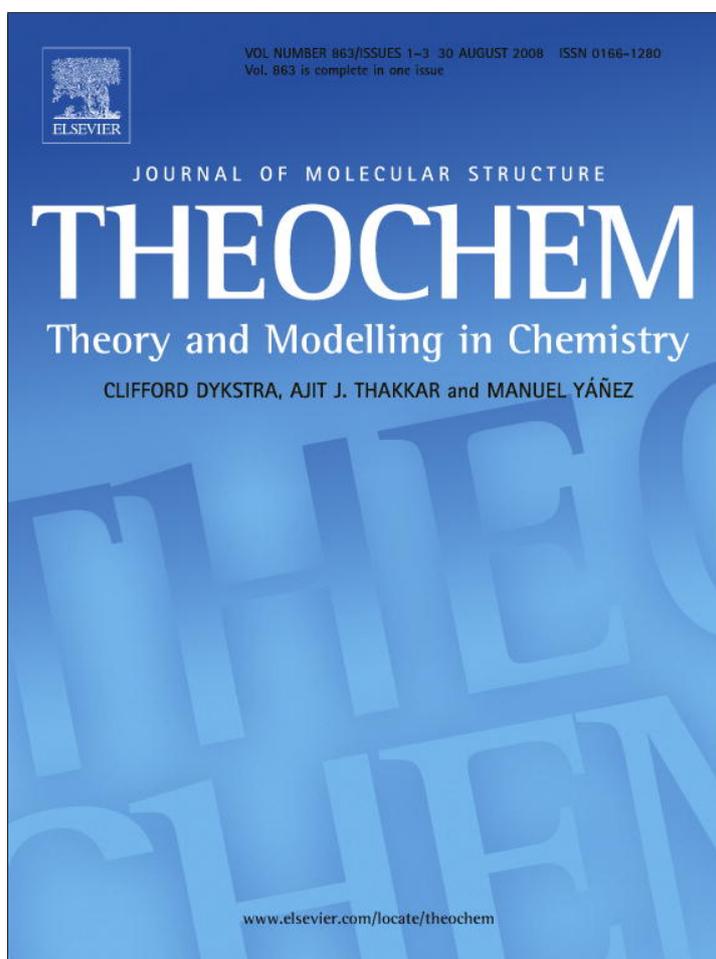


Provided for non-commercial research and education use.  
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

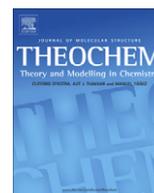
In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at ScienceDirect

## Journal of Molecular Structure: THEOCHEM

journal homepage: [www.elsevier.com/locate/theochem](http://www.elsevier.com/locate/theochem)Conformation dependence of  $pK_a$ 's of the chromophores from the purple asFP595 and yellow zFP538 fluorescent proteinsA.V. Nemukhin<sup>a,b,\*</sup>, I.A. Topol<sup>c</sup>, B.L. Grigorenko<sup>a</sup>, A.P. Savitsky<sup>d</sup>, J.R. Collins<sup>c</sup><sup>a</sup> Department of Chemistry, M.V. Lomonosov Moscow State University, Leninskie Gory, 1/3, Moscow 119991, Russian Federation<sup>b</sup> N.M. Emanuel Institute of Biochemical Physics, Russian Academy of Sciences, ul. Kosygina, 4, Moscow 119334, Russian Federation<sup>c</sup> Advanced Biomedical Computing Center, Advanced Technology Program, SAIC-Frederick Inc., NCI-Frederick, Frederick, MD 21702-1201, USA<sup>d</sup> A.N. Bach Institute of Biochemistry, Russian Academy of Sciences, Leninsky prospekt, 33, Moscow 119071, Russian Federation

## ARTICLE INFO

## Article history:

Received 17 April 2008

Received in revised form 13 May 2008

Accepted 13 May 2008

Available online 22 May 2008

## Keywords:

Biological chromophores

 $pK_a$ *cis-trans* isomerization

Fluorescent proteins

## 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  $pK_a$  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  $pK_a$  values were determined from the thermodynamic cycle based on B3LYP/6-311++G(2df,2p) calculations of the gas phase free energies of the molecules and the B3LYP/6-311++G(d,p) calculations of solvation energies. The results show that the  $pK_a$ '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.

© 2008 Elsevier B.V. All rights reserved.

## 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 *Anemonia sulcata*, asFP595 [2], and the yellow chromoprotein from button polyp *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 in macromolecules or for creating triggerable markers in living cells. 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 ( $O_{\text{phen}}$ ) and the imidazolinone nitrogen atom ( $N_{\text{imid}}$ ).

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

\* Corresponding author. Address: Department of Chemistry, M.V. Lomonosov Moscow State University, Leninskie Gory, 1/3, Moscow 119991, Russian Federation. Tel.: +7 4949391096; fax: +7 4959390283.

E-mail address: [anem@lcc.chem.msu.ru](mailto:anem@lcc.chem.msu.ru) (A.V. Nemukhin).

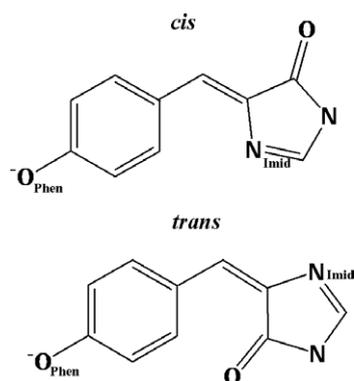


Fig. 1. Structural formula of the GFP chromophore, 4-hydroxybenzylidene-imidazolinone, in the *cis*- (top) and *trans*- (bottom) conformations. Two important protonation sites are designated as  $O_{\text{Phen}}$  and  $N_{\text{Imid}}$ .

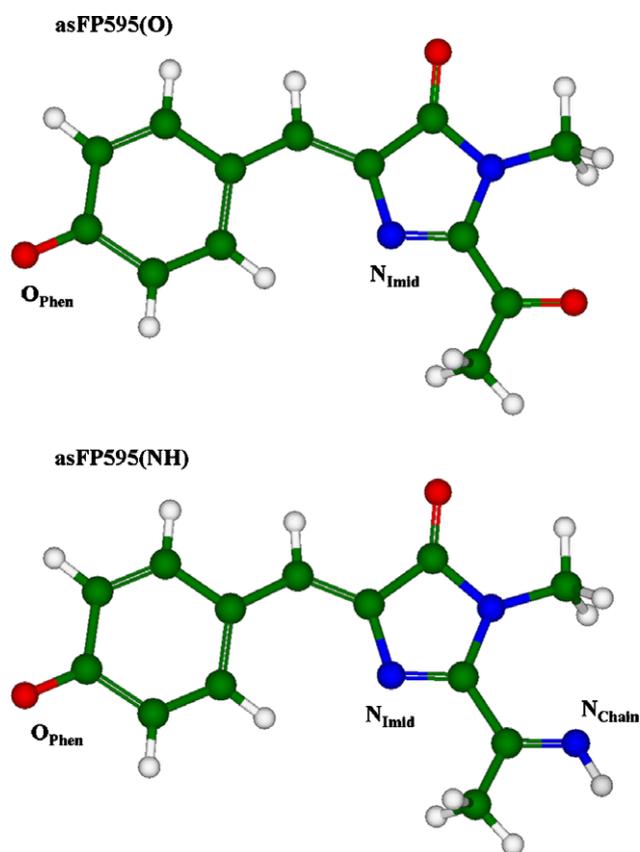


Fig. 2. Two possible structures of the chromophore of the asFP595 protein, asFP595(O) (top) and asFP595(NH) (bottom), in the *cis*-anionic form. In all figures carbon atoms are designated by green, oxygen by red and nitrogen by blue. (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

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

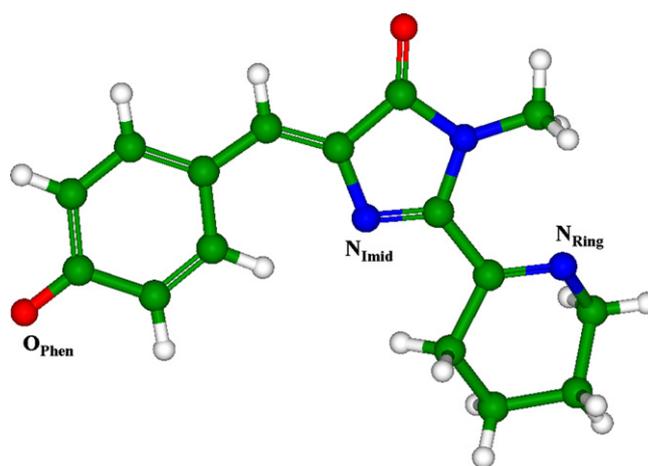


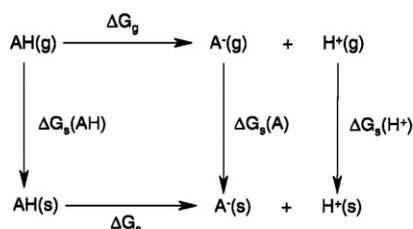
Fig. 3. The chromophore of the zFP538 protein in the *cis*-anionic form.

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  $pK_a$  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  $pK_a$ 's.

Previously, the protonation sites at the phenolic oxygen atom and the imidazolinone 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 ( $pK_a$ ). El Yazal et al. [17] calculated the theoretical  $pK_a$  values for all possible protonation sites of the chromophore in denatured GFP chromophores. Scharnagl and Raupp-Kossmann [18] also reported theoretical  $pK_a$ 's for GFP from hybrid quantum-classical calculations with an average error of 0.8 units. In particular, the authors computed the constant  $pK_a = 8.3$  for neutral/anion equilibria in excellent agreement with the experimental data ( $pK_a = 8.2$ ) on solution-phase acid–base dissociation constants of the model GFP chromophore [13]. No comparison of  $pK_a$ '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  $pK_a$  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  $pK_a$ 's were reported. (2) The seeming  $pK_a$  constants of the zFP538 chromophore in solution were evaluated from the absorption and fluorescent spectra giving the values  $pK1 \approx 5.6$ – $5.7$  and  $pK2 \approx 7.4$ – $7.8$  [20]. From theoretical side, the only estimate of the zwitterionic imidazolinone NH proton of the asFP595 chromophore in aqueous solution by using the density functional theory calculations which resulted in the  $pK$  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  $pK_a$  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  $pK_a$  values accurate within half of a  $pK_a$  unit [21,22].



**Scheme 1.** Thermodynamic cycle for proton abstraction reaction in the gas-phase (g) and in solvent (s).

## 2. Calculation details

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

$$pK_a = (2.303RT)^{-1} \{ \Delta G_g + \Delta G_s(A^-) - \Delta G_s(AH) + \Delta G_s(H^+) \} \quad (1)$$

In Scheme 1 and formula (1),  $\Delta G_g$  and  $\Delta G_s$  are free energies of deprotonation in the gas phase and solvent, respectively.  $\Delta G_s(A^-)$ ,  $\Delta G_s(AH)$  and  $\Delta G_s(H^+)$  are solvation free energies of anion  $A^-$ , protonated compound 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/6-31++G(2d,2p) 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  $\Delta G_s(A^-)$  and  $\Delta G_s(AH)$  were estimated at the B3LYP/6-311++G(d,p) level by using the Jaguar program [24].

The proton solvation free energy  $\Delta G_s(H^+)$ , 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  $\Delta G_s(H^+)$ . 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**  
Computed  $pK_a$  values for the selected protonation sites of the chromophores shown in Figs. 1–3

Species	O <sub>Phen</sub>	N <sub>Imid</sub>	N <sub>Ring</sub>	N <sub>Chain</sub>
GFP, <i>cis</i>	8.53	5.30	–	–
GFP, <i>trans</i>	7.27	3.58	–	–
Anionic zFP538, <i>cis</i>	9.06	4.52	7.02	–
Anionic zFP538, <i>trans</i>	9.09	4.38	7.39	–
Neutral zFP538, <i>cis</i>	–	–0.18	4.13	–
Neutral zFP538, <i>trans</i>	–	2.33	4.21	–
asFP595(O), <i>cis</i>	6.96	1.29	–	–
asFP595(O), <i>trans</i>	8.38	4.63	–	–
asFP595(NH), <i>cis</i>	8.14	3.22	–	7.57
asFP595(NH), <i>trans</i>	7.69	4.74	–	6.77

The *cis*–*trans* isomerization of a model GFP chromophore, 4-hydroxybenzylidene-1,2-dimethylimidazolinone (HBDI), 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,  $1.3 \div 1.7$   $pK_a$  units for the phenolic oxygen and the imino nitrogen of the imidazolinone 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 with the protonated O<sub>Phen</sub>. These values are shown in the rows 'Neutral zFP538' in Table 1. The nitrogen site of the imidazolinone ring turns out to be the most sensitive to conformational changes ( $-0.2$  vs  $2.3$ ). One of the goals of the present modeling of the zFP538 chromophore is to provide interpretation of the experimental measurements of the apparent constants of the chromophore in solution ( $pK_1 \approx 5.6$ – $5.7$  and  $pK_2 \approx 7.4$ – $7.8$  [20]). The computed  $pK_a$ 's for the nitrogen sites ( $4.4$ – $4.5$  and  $7.0$ – $7.4$ ) correlate well with these experimental estimates.

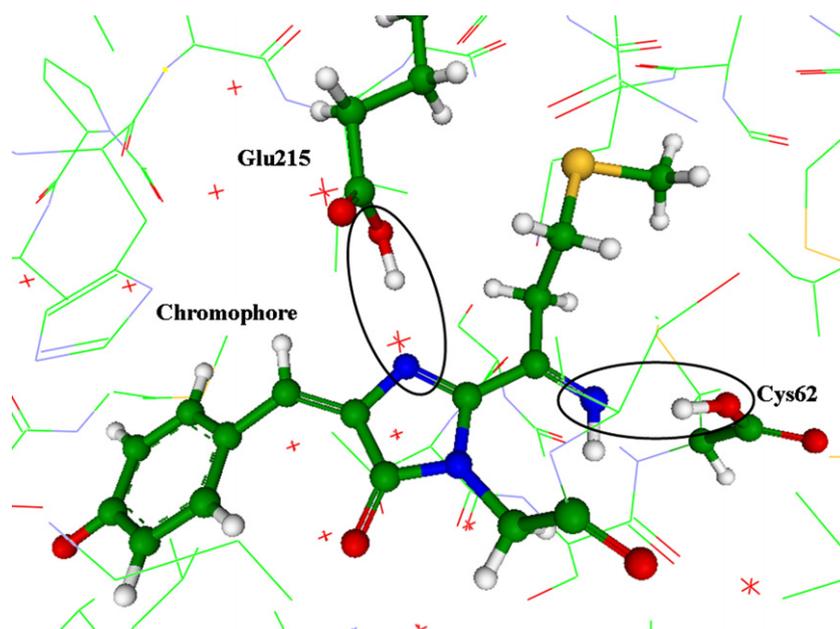
### 3.3. asFP595 Chromophore

As mentioned in Section 1, Quillin et al. [5] reported the chromophore structure for the asFP595 protein in which the conjugated  $\pi$ -system in the GFP-like moiety 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 oxygen (lower panel in Fig. 2). In all crystal structures corresponding to the dark state of the protein, the chromophore appears in the *trans*-conformation [5–7]. It hard to assume formation of the asFP595 chromophore with the protonated O<sub>Phen</sub> site in the protein; therefore, only anionic species are considered in this work. The results of calculations of  $pK_a$ 's shown in Table 1 demonstrate larger sensitivity of the protonation constants of the asFP595 chromophore compared to those of the zFP538 chromophore.

## 4. Discussion and conclusions

The results of calculations demonstrate that the  $pK_a$ 's of the protonation sites of the chromophore from the asFP595 chromophore noticeably depend on the isomer conformation (*cis*- or *trans*-). The most striking difference of more than three  $pK$  units is obtained for the imidasole nitrogen of the asFP595(O) chromophore: 1.29 for the *cis*-form vs 4.63 for the *trans*-form. Such change in the local structure of the chromophore upon isomerization may be related, among other factors, to the kindling fluorescent properties of this protein, the phenomenon which is often attributed to the occurrence of the photo-induced *trans*–*cis* isomerization of the chromophore.

Fig. 4 illustrates a fragment of the asFP595(NH) protein with the *trans*-form of the chromophore as obtained in quantum mechanical – molecular mechanical (QM/MM) calculations [10]. Clearly, formation of the zwitterionic form of the chromophore may occur in the protein matrix when protons from the nearby amino acid side chains are transferred to nitrogen atoms N<sub>Imid</sub> (or N<sub>Chain</sub> in case of the asFP595(NH) form). The present calculations show that



**Fig. 4.** Geometry configuration of the asFP595(NH) chromophore and the nearby amino acid side chains of Glu215 and Cys62 inside the protein matrix. Two possible regions of proton transfers either to the  $N_{\text{imid}}$  or  $N_{\text{chain}}$  protonation sites of the chromophore are indicated by black ovals.

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  $pK_a$  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  $pK_a$ 's for the *trans*- and *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  $pK_a$  constants evaluated directly from the absorption and fluorescent spectra of the zFP538 chromophore in solution are as follows:  $pK1 \approx 5.6$ – $5.7$  and  $pK2 \approx 7.4$ – $7.8$  [20]. Apparently, the computed  $pK_a$ '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.

#### References

- [1] M. Zimmer, Chem. Rev. 102 (2002) 759.
- [2] K.A. Lukyanov, A.F. Fradkov, N.G. Gurskaya, M.V. Matz, Y.A. Labas, A.P. Savitsky, M.L. Markelov, A.G. Zaraisky, X. Zhao, Y. Fang, W. Tan, S.A. Lukyanov, J. Biol. Chem. 275 (2000) 25879.
- [3] S.J. Remington, R.M. Wachter, D.K. Yarbrough, B. Branchaud, D.C. Anderson, K. Kallio, K.A. Lukyanov, Biochemistry 44 (2005) 202.
- [4] D.M. Chudakov, A.V. Feofanov, N.N. Mudrik, S. Lukyanov, K.A. Lukyanov, J. Biol. Chem. 278 (2003) 7215.
- [5] M.L. Quillin, D.M. Anstrom, X. Shu, S. O'Leary, K. Kallio, D.M. Chudakov, S.J. Remington, Biochemistry 44 (2005) 5774.
- [6] P.G. Wilmann, J. Petersen, R.J. Devenish, M. Prescott, J. Rossjohn, J. Biol. Chem. 280 (2005) 2401.
- [7] M. Andresen, M.C. Wahl, A.C. Stiel, F. Gräter, L.V. Schäfer, S. Trowitzsch, G. Weber, C. Eggeling, H. Grubmüller, S.W. Hell, S. Jakobs, Proc. Natl. Acad. Sci. USA 102 (2005) 13070.
- [8] T.A. Schüttigkeit, T. von Feilitzsch, C. Kompa, K.A. Lukyanov, A.P. Savitsky, A.A. Voityuk, M.-E. Michel-Beyerle, Chem. Phys. 323 (2006) 149.
- [9] J.N. Henderson, S.J. Remington, Physiology 21 (2006) 162.
- [10] B.L. Grigorenko, A.P. Savitsky, I.A. Topol, S.K. Burt, A.V. Nemukhin, J. Phys. Chem. B 110 (2006) 18735.
- [11] L.V. Schäfer, G. Groenhof, A.R. Kligen, G.M. Ullmann, M. Boggio-Pasqua, M.A. Robb, H. Grubmüller, Angew. Chem. Int. Ed. 46 (2007) 530.
- [12] L.V. Schäfer, G. Groenhof, M. Boggio-Pasqua, M.A. Robb, H. Grubmüller, PLOS Comput. Biol. 4 (2008) e1000034.
- [13] A.F. Bell, X. He, R.M. Wachter, P.J. Tonge, Biochemistry 39 (2000) 4423.
- [14] A.A. Voityuk, M.-E. Michel-Beyerle, N. Rösch, Chem. Phys. Lett. 272 (1997) 162.
- [15] A.A. Voityuk, M.-E. Michel-Beyerle, N. Rösch, Chem. Phys. 231 (1998) 13.
- [16] C. Scharnagl, R.A. Raupp-Kossmann, S.F. Fischer, Biophys. J. 77 (1999) 1839.
- [17] J. El Yazal, F.G. Prendergast, D.A. Shaw, Y.-P. Pang, J. Am. Chem. Soc. 122 (2000) 11411.
- [18] C. Scharnagl, R.A. Raupp-Kossmann, J. Phys. Chem. B 108 (2004) 477.
- [19] I.V. Yampolsky, S.J. Remington, V.I. Martynov, V.K. Potapov, S. Lukyanov, K.A. Lukyanov, Biochemistry 44 (2005) 5788.
- [20] N. Zubova, Ph.D. thesis, Department of Chemistry, M.V. Lomonosov Moscow State University, 2006.
- [21] I.A. Topol, G.J. Tawa, S.K. Burt, A.A. Rashin, J. Phys. Chem. A 101 (1997) 10075.
- [22] G.J. Tawa, I.A. Topol, S.K. Burt, R.A. Caldwell, A.A. Rashin, J. Chem. Phys. 109 (1998) 4852.
- [23] Gaussian 03, Revision C.02, M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, J.A. Montgomery Jr., T. Vreven, K.N. Kudin, J.C. Burant, J.M. Millam, S.S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G.A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J.E. Knox, H.P. Hratchian, 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, P.Y. Ayala, K. Morokuma, G.A. Voth, P. Salvador, J.J. Dannenberg, V.G. Zakrzewski, S. Dapprich, A.D. Daniels, M.C. Strain, O. Farkas, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J.V. Ortiz, Q. Cui, A.G. Baboul, S. Clifford, J. Cioslowski, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, C. Gonzalez, J.A. Pople, Gaussian, Inc., Wallingford, CT, 2004.
- [24] Jaguar, Schrodinger, LLC: Portland, OR, 1991–2005.
- [25] M.W. Palascak, G.C. Schields, *J. Phys. Chem. A* 108 (2004) 3692.
- [26] X. He, A.F. Bell, P.J. Tonge, *FEBS Lett.* 549 (2003) 35.