Biophysical Chemistry 158 (2011) 91–95

Contents lists available at ScienceDirect

Biophysical Chemistry

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

Computational strategy for tuning spectral properties of red fluorescent proteins

I. Topol [a,*], J. Collins [a], A. Savitsky [b,c], A. Nemukhin [b,d]

[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] A.N. Bach Institute of Biochemistry, Russian Academy of Sciences, 33 Leninskii prospekt, Moscow, 119071, Russia
[d] N.M. Emanuel Institute of Biochemical Physics, Russian Academy of Sciences, 4 Kosygina, Moscow, 119334, Russia

ARTICLE INFO

Article history:
Received 28 March 2011
Received in revised form 11 May 2011
Accepted 20 May 2011
Available online 27 May 2011

Keywords:
Red fluorescent proteins
Photoabsorption spectra
Quantum calculations
Electric field influence

ABSTRACT

Computational methods of quantum chemistry are used to characterize structures and vertical excitation energies of the S_0–S_1 optical transitions in the chromophore binding pockets of the red fluorescent proteins DsRed and of its artificial mutant mCherry. As previously shown, optimizing the equilibrium geometry configurations with B3LYP density functional theory, followed by ZINDO calculations of the electronic excitations, yields positions of the optical bands in good agreement with experimental data. These large scale quantum calculations elucidate the role of the hydrogen bonded network as well as point mutations in the absorption spectra of the DsRed and mCherry proteins. The effect of an external electric field applied to the fluorescent protein chromophores is examined and shows that such fields may result in large shifts in spectral bands. These strategies can be applied for rational design of the fluorescent proteins by site-directed mutagenesis.

1. Introduction

Fluorescent proteins (FP) are widely used as in vivo markers allowing one to carry out successive visual monitoring of numerous molecular and cell processes in intact living organisms [1]. To minimize tissue absorption and autofluorescence signals and therefore to achieve imaging deeper into tissues, application of red or near-infrared light for excitation is highly desirable [2]. Red FPs are also important partners in fluorescence resonance energy transfer (FRET) methods [3]. The primary target of this work is the modeling of photophysical properties of the red fluorescent proteins DsRed from *Discosoma* coral [4–6] and its artificial mutant of the so-called mFruits series, mCherry [7,8]. In spite of wide experimental studies of FPs proteins, theoretical modeling can be used to design different and more efficient markers, in particular, by in silico predictions of promising mutants with the improved photophysical properties.

Previous investigations [9,10] showed that by using combination of ab initio and semiempirical methods of quantum chemistry, one can accurately reproduce spectral properties of the GFP-like chromophores as well as of molecular clusters mimicking chromophore binding pockets of the proteins. In particular, important features of the green fluorescent protein (GFP) from the jellyfish *Aequorea victoria* [11] and of the monomeric teal fluorescent protein from *Clavularia* coral (mTFP1) [12,13] have been well reproduced [9,10] by using geometry optimization with density functional theory methods followed by ZINDO [14] calculations of the vertical electronic excitations. This computational strategy produces positions of the optical bands consistent with the experimental data.

In this work the computational strategy is expanded, this time focused on the properties of the red fluorescent proteins DsRed and mCherry by considering the effect of an external electric field. Such fields simulate an influence of polarized environmental molecular groups on the chromophore properties, and if its effect is noticeable, then a rational design of new variants of fluorescent proteins may be predicted, since the direction and strength of the field can be controlled by the position of charged or polar amino acid residues introduced into the protein matrix by site-directed mutagenesis.

Along with the bare chromophores, the properties of the chromophore binding pockets are analyzed by considering fairly large molecular clusters. Fig. 1 illustrates the chemical structures of the DsRed and mCherry bare chromophores in the anionic forms. Compared to the parent GFP chromophore their series of alternating single and double bonds is expanded by the additional CA1 N double bond (acylimine region), thus explaining the shifts to the red of both absorption and emission spectral bands. The main difference between the two chromophores is the presence of the sulfur containing fragment in the mCherry chromophore. The equilibrium geometry parameters of these chromophore moieties shown in Fig. 1 have been optimized using the B3LYP/6-31 + G(d,p) approximation.

2. Models and methods

This investigation examines the isolated chromophore species from the red fluorescent proteins as shown in Fig. 1 and the molecular

* Corresponding author.
E-mail address: topoli@mail.nih.gov (I. Topol).

0301-4622/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.bpc.2011.05.016

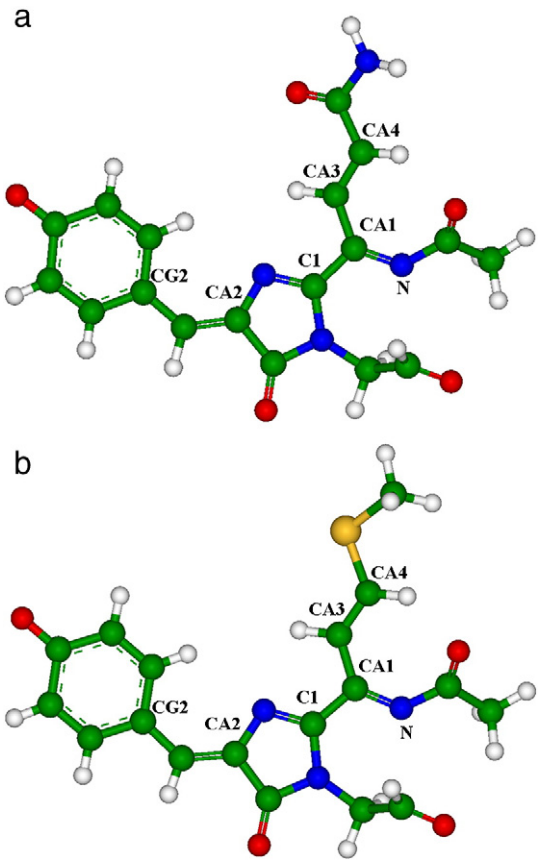

Fig. 1. Chemical structures of the DsRed (left) and of the mCherry (right) fluorescent protein chromophores. Here and in other figures, carbon atoms are shown in green, oxygen in red, nitrogen in blue, sulfur in yellow.

clusters constructed on the basis of the coordinates of heavy atoms from the available crystal structures. The immediate chromophore environment for the DsRed model included 12 residues (Gln42, Pro63, Gln64, Phe65, Ser69, Lys70, Arg95, Ser146, Glu148, Ile161, His163, Glu215) and 6 water molecules. The mCherry cluster model was built using 15 residues (Gln42, Pro63, Gln64, Phe65, Ser69, Lys70, Arg95, Ser146, Glu148, Ile161, Gln163, Ile197, Leu199, Gln213 and Glu215) and 7 water molecules. The broken bonds of surface cluster atoms with the atoms outside the cluster were saturated by hydrogen atoms. The equilibrium geometries of the isolated chromophores in the ground singlet electronic state S_0 were obtained using the B3LYP/6-31 + G(d,p) density functional theory (DFT) approximation. Coordinates of heavy atoms from the structures deposited to the Protein Data Bank were used as starting coordinates for the heavy atoms in the molecular clusters. Hydrogen atoms were then added to the model systems and the final coordinates for the clusters were optimized using the B3LYP/6-31 G(d) approximation. During this geometry optimization, the positions of the C_α atoms were frozen while all other coordinates were allowed to vary. Positions of water molecules were taken initially from the crystal pdb structures and further were completely optimized without any constrains.

The semiempirical ZINDO method [14] was used to compute vertical excitation energies from the respective minima on the S_0 potential surfaces of every model. As shown previously [9,10], this strategy to estimate the S_0–S_1 excitation energies at the DFT-optimized equilibrium geometry parameters performs quite well for these model systems. All simulations have been carried out with the Gaussian03 program [15]. No changes of the default parameters have been introduced in calculations.

The model system for simulations of the DsRed protein was constructed on the basis of the PDBID:2VAD [6]. The corresponding molecular cluster is shown in Fig. 2. The model system for simulations of the mCherry protein was constructed following the motifs of the crystal structure PDBID:2H5Q [8], as shown in Fig. 3.

2.1. Results for isolated chromophores

The series of calculations were performed for the chromophore molecules shown in Fig. 1. Compared to the denatured GFP chromophore, 4'-hydroxybenzylidene-2,3-dimethylimidazolinone (HBDI), the chromophores depicted in Fig. 1, possess an enlarged π-conjugated system and correspondingly are characterized by longer wavelength absorption and emission spectral bands.

It should be noted that the experimental absorption maxima for the DsRed and mCherry proteins are observed at 558 nm [16] and 587 nm [8], respectively. The results of the ZINDO/DFT approximation for the isolated chromophore molecules for the S_0–S_1 excitation energies correspond to slightly longer wavelengths; 625 nm for DsRed and 613 for mCherry. The position**s** of the emission band**s** of the chromophore molecules from the DsRed and mCherry proteins were also estimated by optimizing the geometry in the excited electronic state S_1 in the configuration interaction method with single excitations (CIS) approximation as implemented in Gaussian03 [15]. The corresponding energy gaps between the S_0–S_1 states were calculated in the ZINDO approach. As expected the fluorescence bands lie at the longer wavelength (630.7 and 624.5 nm for DsRed and mCherry chromophores, respectively) compared to the corresponding absorption bands from the ground state S_0. Both computed absorption and emission bands are characterized by large oscillator strengths of 1.2.

Therefore, the chromophore molecules in the anionic form in the gas phase should possess even longer absorption wavelengths than the chromophores inside the proteins. This suggests that the molecular groups in the protein are responsible for shortening the wavelengths of the optical transitions.

2.2. Anisotropic effect of the external electric field on the chromophore excitations

If the molecular groups, and therefore the electric field, around the chromophores significantly affect its spectra, this may yield a powerful predictive strategy to tune photophysical properties of

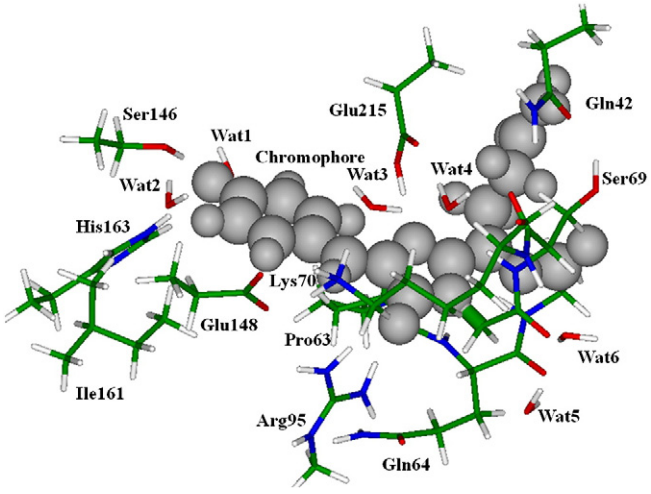

Fig. 2. Molecular cluster modeling the chromophore binding pocket of the DsRed fluorescent protein. The chromophore species is shown in the space-filled representation.

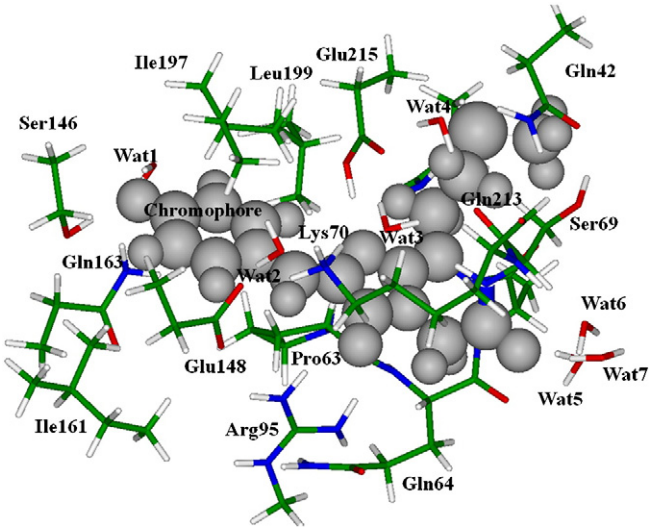

Fig. 3. Molecular cluster modeling the chromophore binding pocket of the mCherry fluorescent protein. The chromophore species is shown in the space-filled representation.

photoreceptor proteins. More specifically, if changes in the spectra can be directly associated with changes in the electric field around the chromophores, site-directed mutations may be proposed to reproduce this electric field.

It is too expensive to directly calculate excitation energies for the chromophores in the external field, but it is feasible to estimate the field effect on the HOMO–LUMO gap. Basing on the simple quantum chemistry consideration we expect that the HOMO–LUMO gaps for the chromophore molecules should correlate with the total energy differences between the excited and ground states. To verify this assumption we plot on Fig. 4 the HOMO–LUMO energy gap versus the vertical energy gap between S_0 and S_1 for initial chromophore structures depicted in Fig. 1 and as well as for their resonance counterparts. The linear fit through the calculated values for the DsRed and mCherry chromophores depicted in Fig. 4 clearly demonstrates that changes in the HOMO–LUMO gap are directly correlated with changes in the absorption energy. The same linear correlation between the excitation energy and the HOMO–LUMO gap was previously observed for the mTFP1 chromophore [10].

A series of calculations was performed by applying the external electric field of different strength and in different directions to the chromophore molecules; estimating the typical values of the interior electric field inside the bare chromophore species. According to these calculations the electronic distribution in these molecules without an external field is responsible for the absolute field values on the heavy atoms within the range from 0.003 to 0.04 a.u. Therefore external field strengths up to 0.01 a.u. should not create an excessive perturbation from the exterior of the chromophore molecule. In calculations, we imposed the external electric field by placing point charges as described in the Gaussian03 [15] manual.

The main conclusion from the estimates illustrated in Fig. 5 and Table 1 is that the field applied in certain specific directions may account for very large spectral shifts up to 200 nm to the red. This effect is worth further studies due to its potential importance in practical applications.

It should be taken into consideration that the protein-structural dynamics produce fluctuating electric fields because polar groups throughout the protein are moving, and these fluctuating fields are responsible for observable properties of cofactors [16]. However, as shown for example in [17] the correlation between the calculated transition energies and the structural distortions in both the chromophore and the surrounding protein environment is mainly responsible for the temperature-dependent broadening of the electronic transitions seen in the visible spectra, but not for positions of the bands.

2.3. Results for molecular clusters

Fig. 6 shows the most important fragments of the computed model cluster mimicking the chromophore binding pocket of the DsRed protein. Included in this figure are the optimized distances between heavy atoms in angstroms (the values in parentheses refer to the crystal structure 2VAD). The molecular groups near the oxygen and nitrogen atoms in the conjugated system of the chromophore moiety may have the largest effect on the spectra. The hydrogen bond network over the negatively charged phenolic oxygen including side

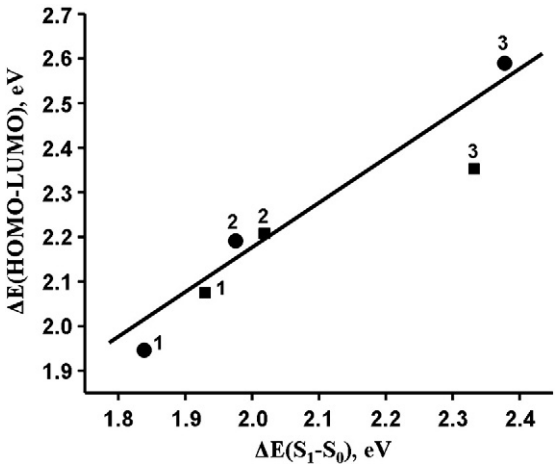

Fig. 4. Correlation of the vertical energy gap between S_0 and S_1 and the HOMO–LUMO energy difference for the variants of DsRed chromophore (circles) and the mCherry chromophore (squares). Points 1 correspond to chromophores depicted in Fig. 1. Points 2 and 3 correspond to the variants of chromophores caused by changing the CA3–CA4 to a double bond (Points 2) or by changing the CA1–CA3 single bond to a double bond with protonation of N atom in the acylimine group (Points 3). The linear fit is performed through the entire set of values for both systems.

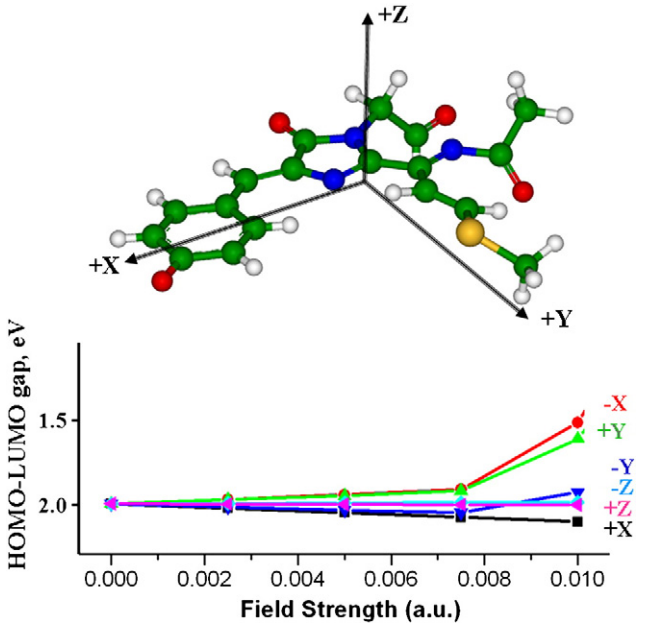

Fig. 5. Dependence of the HOMO–LUMO gap in the mCherry chromophore on the magnitude and direction of the applied electric field.

Table 1
Dependence of the HOMO–LUMO gap on the electric field (0.01 a.u.) applied in different directions (see Fig. 5).

Field applied	HOMO–LUMO energy gap, eV	Band shift, nm
+X	2.46	−110
−X	1.52	+205
+Y	1.62	+154
−Y	2.08	−16
+Z	2.19	−47
−Z	2.23	−57

chains of Ser146, His163 and the water molecule (designated as Wat2042 in the crystal structure 2VAD) is well reproduced in calculations. The same holds for the relative positions of the Arg95 side chain and the oxygen atom as well as for the Glu215 side chain and the nitrogen atom of the imidazole ring of the chromophore.

For the molecular cluster mimicking the immediate chromophore environment in DsRed, Fig. 6, our calculations predict the absorption maximum at 560 nm which is in excellent agreement with the observations at 558 nm [18]. We discuss below how this calculation result depends on the protonation state of the Glu215 residue and its involvement into the hydrogen bond network.

Calculations for the molecular cluster mimicking the chromophore binding pocket in mCherry (Fig. 7) actually correspond to the largest quantum calculations performed at this level of modeling (compare, for instance, to ref.[19]). The system included over 250 atoms from the chromophore and 15 amino acid residues. Like in the case of the DsRed model these calculation results (574 nm for the absorption maximum) are in the good agreement with the experimental data (587 nm [8]). Calculations correctly predicted a noticeable red shift for this system (the calculated wavelength 574 nm) compared to the DsRed model (560 nm).

It is important to examine the role of the conservative amino acid residue Glu215, which is present in all GFP-like fluorescent proteins. In most species it is protonated, and this proton potentially can be transferred to an imidazole nitrogen of the chromophore leading to the zwitterionic structure of the latter. Such transfer is not observed in our simulations, however, the protonation status of Glu215 and hydrogen bonding involving Glu215 strongly affects the absorption spectrum: for instance, if Glu215 is unprotonated in DsRed, the calculated band is at 544 nm; if Glu215 is protonated and forms the hydrogen bond with the chromophore, the calculated band is at

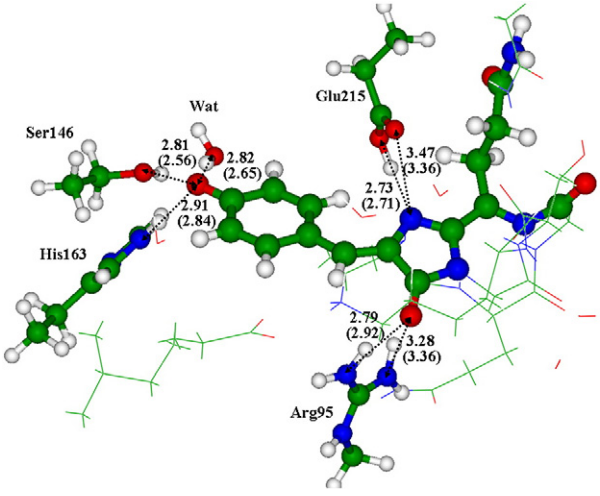

Fig. 6. Equilibrium structure of the model cluster from the DsRed chromophore binding pocket. The key groups are shown in balls and sticks, other groups in lines. The distances between selected heavy atoms (in Å) without parentheses are obtained in calculations, and those in parentheses refer to the crystal structure PDBID:2VAD.

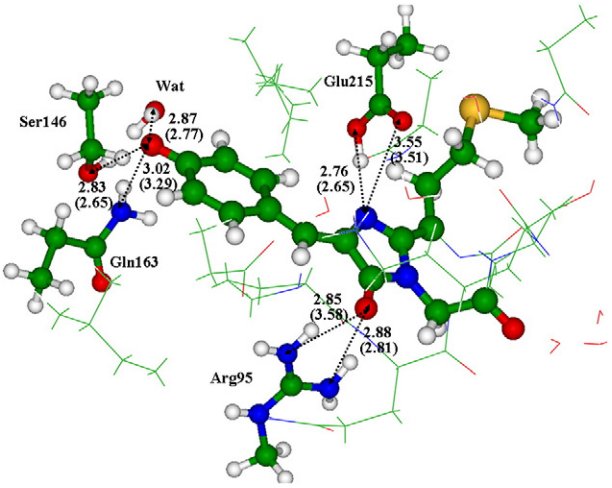

Fig. 7. Equilibrium structure of the model cluster from the mCherry chromophore binding pocket. The key groups are shown in balls and sticks, other groups in lines. The distances between selected heavy atoms (in Å) without parentheses are obtained in calculations, and those in parentheses refer to the crystal structure PDBID:2H5Q.

560 nm, and if it forms the hydrogen bond with the neighboring water molecule, the band is at 556 nm.

The specific orientation of water chains near the chromophore also affects the excitation wavelength. Fig. 8 shows the fragment of the molecular cluster for the mCherry model (illustrated in full in Fig. 3 or Fig. 7) focusing on the hydrogen bond network over the water molecules in the close proximity of the chromophore. This structure, which corresponds to the equilibrium geometry configuration of the model system, produces an excellent prediction of the absorption band at 574 nm. However, a different orientation of the water molecules designated Wat2 and Wat3 in Fig. 8 can be obtained in another local minimum on the potential energy surface which results in the computed absorption band at 565 nm. Therefore the spectral tuning of fluorescent proteins can be significantly affected by slight changes in the hydrogen bond network around the chromophore. We also considered few point mutations in the clusters by changing polar residues by non-polar ones and vice a verse. In particular, mutations of His163 or Ser146 by Met in the DsRed protein were examined, but according to our calculations these modifications did not lead to noticeable changes in the absorption bands.

As requested by one of the reviewers, we carried out calculations of the spectral bands of these model systems at the CIS and CIS + D levels, as well as by the TDDFT method. In accord with the previous knowledge (e.g., refs. [20,21]), the TDDFT excitation energies overestimate experimental values by ~0.5 eV. The CIS and CIS + D data also are inferior to the results of ZINDO calculations. For example, for

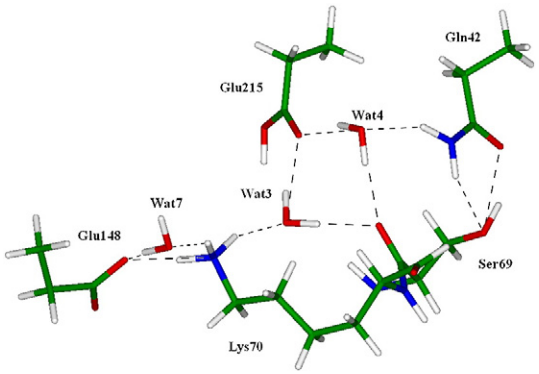

Fig. 8. The hydrogen bond network over the chain of water molecules near the mCherry chromophore (see Figs. 3 and 7 for a complete system).

the free DsRed and mCherry chromophores the CIS energies for $S_0 \rightarrow S_1$ transition were calculated as 385 and 379 nm, respectively. The inclusion of double excitations (CIS + D) improved results (558 and 551 nm, respectively) but they are still far below from ZINDO excitation energies (625 and 613 nm, respectively). The cluster data for DsRed showed the same tendency: S_0–S_1 transition energy was substantially overestimated at the CIS level—372 nm versus 560 nm (ZINDO-B3LYP cluster) and 558 nm (mDsRed protein).

3. Conclusion

This work confirms the value of modern tools of quantum chemistry for accurate calculations of structures and optical band positions in the spectra of biological chromophores in protein matrices. When modeling bare chromophores as well as the chromophore binding pockets in the protein matrices for the DsRed and mCherry variants we are able to optimize equilibrium geometry parameters and to accurately compute energy differences between the excited and ground state energies, thus estimating the bands in absorption spectra of fluorescent proteins. Application of this strategy for red fluorescent proteins permits the identification of functional states of the chromophores and to understand effects of the nearby amino acid residues and elucidate the role of point mutations. The effect of the external electric field to the fluorescent protein chromophores was modeled and demonstrated that application of the field along certain directions can account for fairly large shifts in spectral bands. Such effect can be used for rational design of the fluorescent proteins by site-directed mutagenesis.

Acknowledgments

We thank the staff and administration of the Advanced Biomedical Computing Center for their support of this project. We thank Dr. Brian Luke for very helpful comments. This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under contract number HHSN261200800001E. 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. This work is partly supported by the Russian Foundation for Basic Research (project 10-03-00085) and the Program of Molecular and Cell Biology from the Russian Academy of Sciences.

References

[1] B.N.G. Giepmans, S.R. Adams, M.H. Ellisman, R.Y. Tsien, The fluorescent toolbox for assesting protein location and function, Science 312 (2006) 217–224.
[2] I.T. Li, E. Pham, K. Truong, Protein biosensors based on the principle of fluorescence resonance energy transfer for monitoring cellular dynamics, Biotechnol. Lett. 28 (2006) 1971–1982.
[3] K. Truong, M. Ikura, The use of FRET imaging microscopy to detect protein-protein interactions and protein conformational changes in vivo, Curr. Opin. Struct. Biol. 11 (2001) 573–578.
[4] M.V. Matz, A.F. Fradkov, Y.A. Labas, A.P. Savitsky, A.G. Zaraisky, M.L. Markelov, S.A. Lukyanov, Fluorescent proteins from nonbioluminescent *Anthozoa* species, Nat. Biotechnol. 17 (1999) 969–973.
[5] R.E. Campbell, O. Tour, A.E. Palmer, P.A. Steinbach, G.S. Baird, D.A. Zacharias, R.Y. Tsien, A monomeric red fluorescent protein, Proc. Natl. Acad. Sci. U.S.A. 99 (2002) 7877–7882.
[6] D.E. Strongin, B. Brooke, N. Khuong, M.E. Downing, R.L. Strack, K. Sundaram, B.S. Glick, R.J. Keenan, Structural rearrangements near the chromophore influence the maturation speed and brightness of DsRed variants, Protein Eng. Des. Sel. 20 (2007) 525–534.
[7] N.C. Shaner, R.E. Campbell, P.A. Steinbach, B.N. Giepmans, A.E. Palmer, R.Y. Tsien, Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein, Nat. Biotechnol. 22 (2004) 1567–1572.
[8] X. Shu, N.C. Shaner, C.A. Yarbrough, R.Y. Tsien, S.J. Remington, Novel chromophores and buried charges control color in mFruits, Biochemistry 45 (2006) 9639–9647.
[9] I. Topol, J. Collins, I. Polyakov, B. Grigorenko, A. Nemukhin, On photoabsorption of the neutral form of the green fluorescent protein chromophore, Biophys. Chem. 145 (2009) 1–6.
[10] I. Topol, J. Collins, A. Nemukhin, Modeling spectral tuning in monomeric teal fluorescent protein mTFP1, Biophys. Chem. 149 (2010) 78–82.
[11] M. Zimmer, Green fluorescent protein (GFP): applications, structure, and related photophysical behavior, Chem. Rev. 102 (2002) 759–781.
[12] H.W. Ai, J.N. Henderson, S.J. Remington, R.E. Campbell, Directed evolution of a monomeric, bright and photostable version of *Clavularia* cyan fluorescent protein: structural characterization and applications in fluorescence imaging, Biochem. J. 400 (2006) 531–540.
[13] H.W. Ai, S.G. Olenych, P. Womg, M.W. Davidson, R.E. Campbell, Hue-shifted monomeric variants of *Clavularia* cyan fluorescent protein: identification of the molecular determinants of color and applications in fluorescence imaging, BMC Biol. 6 (2008) 13.
[14] M.C. Zerner, in: K.B. Lipkowitz, D.B. Boyd (Eds.), Rev. Comput. Chem., 2, VCH Publishing, New York, 1991, pp. 313–366.
[15] 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 03, Gaussian, Inc, Wallingford, CT, 2004.
[16] M.D. Fayer, Fast protein dynamics probed with infrared vibrational echo experiments, Ann. Rev. Phys. Chem. 52 (2001) 315–356.
[17] N.V. Prabhu, S.D. Dalosto, K.A. Sharp, W.W. Wright, J.M. Vanderkooi, Optical spectra of Fe(II) cytochrome *c* interpreted using molecular dynamics simulations and quantum mechanical calculations, J. Phys. Chem. B 106 (2002) 5561–5571.
[18] G.S. Baird, D.A. Zacharias, R.Y. Tsien, Biochemistry, mutagenesis and oligomerization of DsRed, a red fluorescent protein from coral, Proc. Natl. Acad. Sci. U.S.A. 97 (2000) 11984–11989.
[19] N. Taguchi, Y. Mochizuki, T. Nakano, S. Amari, K. Fukuzawa, T. Ishikawa, M. Sakurai, A. Tanaka, Fragment molecular orbital calculations on red fluorescent proteins (dsRed and mFruits), J. Phys. Chem. B 113 (2009) 1153–1161.
[20] B. Grigorenko, A. Savitsky, I. Topol, S. Burt, A. Nemukhin, *trans* and *cis* chromophore structures in the kindling fluorescent protein asFP595, Chem. Phys. Lett. 424 (2006) 184–188.
[21] A.V. Nemukhin, I.A. Topol, S.K. Burt, Electronic excitations of the chromophore from the fluorescent protein asFP595 in solutions, J. Chem. Theory Comput. 2 (2006) 292–299.