ISSN 0021-3640, JETP Letters, 2010, Vol. 91, No. 9, pp. 490–493. © Pleiades Publishing, Inc., 2010.
Original Russian Text © S.G. Ovchinnikov, L.Yu. Antipina, F.N. Tomilin, A.A. Kuzubov, 2010, published in Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, 2010, Vol. 91, No. 9, pp. 536–540.

Effect of Electron Correlations on the Structure of Photoprotein Substrates

S. G. Ovchinnikov[a, b], L. Yu. Antipina[b], F. N. Tomilin[a, b], and A. A. Kuzubov[b]

[a] *Kirensky Institute of Physics, Siberian Branch, Russian Academy of Sciences, Akademgorodok, Krasnoyarsk, 660036 Russia*

[b] *Siberian Federal University, Krasnoyarsk, 660041 Russia*

e-mail: antipinalyu@gmail.com

Received April 6, 2010

The electronic structure and total energy of various isomeric forms of coelenterazine and coelenteramide have been calculated by quantum chemistry methods both in the single-electron approximation and taking into account correlation effects. It has been shown that the inclusion of electron correlations makes it possible to obtain the structure of the coelenteramide close to the experimentally determined structure, as well as to choose the structure of the coelenterazine CLZ(1H) as the most probable isomeric form.

DOI: 10.1134/S0021364010090122

1. Photoproteins are responsible for the luminescence of many living organisms. The crystal structures of some photoproteins have recently been determined [1–3] and the mechanisms of the bioluminescence reaction have been presented [4, 5]. To understand the physical processes occurring at the active center of a photoprotein, it is necessary first of all to determine the real atomic structure of the central coelenterazine (CLZ) molecule among several possible isomeric forms. In this work, this problem is solved by means of microscopic quantum-mechanical calculations of the CLZ molecule. Various semiempirical and ib initio methods were used both in the Hartree–Fock approximation and taking into account correlation effects. The inclusion of correlation effects strongly changes the single-electron results. The single-electron semiempirical method PM3 and single-electron ab initio Hartree–Fock method provide the largest deviation from the experimental data. The inclusion of the correlation effects in the density functional theory (DFT) and in the second-order perturbation theory in the Coulomb interaction (the MP2 method) provides a good description of the geometry of the molecules. It was shown that the active center of photoproteins contains the CLZ(1H) structure rather than the CLZ(7H) structure, as was accepted previously.

2. The bioluminescence of photoproteins occurs through several stages (see Fig. 1). First, molecular oxygen is attached to the initial substrate of the coelenterazine with the formation of 2-hydroperoxycoelenterazine (GP-CLZ). Then, after the joining of calcium ions to the protein, the structure of the protein changes, and the reaction of oxidative decarboxilation occurs with the formation of the coelenteramide in the excited state.

The authors of [6] performed a detailed analysis of the pH effects and protic and aprotic solvents in an anaerobic medium on the spectral properties of the coelenterazine and some its analogs. They found that the absorption spectrum of the coelenterazine in an

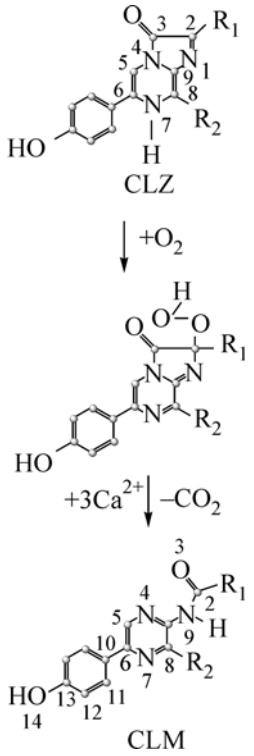

Fig. 1. Mechanism of the bioluminescence reaction from [4].

aprotic solvent (dimethyl sulphoxide) has a redshift with respect to methanol (454 and 435 nm, respectively). Different absorption maxima were attributed to different tautomeric forms CLZ(7H) and CLZ(2H), respectively. However, according to the crystal structure of the substrate, CLZ(2H) can be formed only inside the photoprotein obelin. Coelenterazine can in vitro exist in two tautomeric forms, CLZ(1H) and CLZ(7H) (see Fig. 2).

Strong geometric differences between the structures of CLZ(1H) and CLZ(7H) are absent; for this reason, it is difficult to experimentally determine which form exists under certain conditions. In this work, the structures are theoretically investigated by the quantum chemical methods. The single-electron calculations for the case of sufficiently complicated structures (i.e., a large number of diverse bonds), which contain atoms with unshared electron pairs (nitrogen and oxygen), can provide a large error. For this reason, calculations with the inclusion of electron correlations are preferable for these systems.

3. The Hartree–Fock method describes the motion of each electron in the averaged field of the other electrons. The difference between the exact (nonrelativistic) energy and Hartree–Fock energy of a multielectron system is called the correlation energy,

$$E_{\text{correl}} = E_{\text{exact}} - E_{HF}. \quad (1)$$

The spatial correlation of the motions of the electrons in the molecule has three causes:

(i) the location of the electrons near each other is energetically unfavorable owing to the Coulomb repulsion forces between them,

(ii) two electrons with parallel spins cannot be at the same spatial point owing to the Pauli exclusion principle, and

(iii) the correlation effects depend on the spatial symmetry properties (point symmetry groups).

The Möller–Plesset perturbation theory (usually the second and fourth orders, MP2 and MP4, respectively), configuration interaction methods, and the density functional theory [7–11] are widely used to calculate the correlation energy.

In the configuration interaction method, the multielectron wavefunction is expanded into the Slater determinants, each describing the system in one of the possible electronic states. Each state corresponds to a certain electronic configuration: their set represents the possible transitions of the electron from an occupied molecular orbital to various unoccupied (virtual) orbitals. This means that each of such determinant is constructed from spin orbitals corresponding to the ground or one of the excited single-electron states of the molecule.

The Möller–Plesset theory is based on the representation of the solution of the complete multielectron problem in the form of the perturbation of the Hartree–Fock solution. The main contribution to the correlation energy comes from the second-order term in the case of sufficiently small correlations.

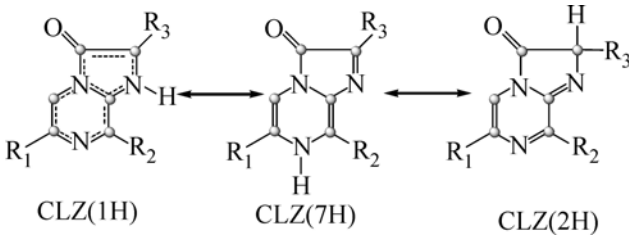

Fig. 2. Isomeric transition between the forms of the coelenterazine.

The main idea of the density functional theory is that the energy of the electron system is certainly determined by the electron density. Thus, instead of the search for the wavefunction, which is a function of $3N$ variables (where N is the number of electrons in the system), the problem is reduced to the determination of the electron density of the system, which is a function of three spatial coordinates.

At present, the so-called hybrid density functional methods are widely used in the calculations of molecules and clusters. In these methods, the exchange energy is written as a linear combination of exchange functionals LSDA, GGA, and Hartree–Fock exchange energy. The correlation energy is written as a linear combination of correlation functionals. The currently most popular functional B3LYP [12] is given by the expression

$$E_{XC}^{\text{B3LYP}} = (1-a)E_X^{\text{LSDA}} + aE_X^{\text{HF}} + b\Delta E_X^{\text{B88}} + E_X^{\text{VWN}} + c\Delta E_C^{\text{LYP}}. \quad (2)$$

The parameters a, b, and c are optimized in order to reproduce the experimental data.

4. The structure of the coelenteramide (CLM) is experimentally known, whereas the structure of CLZ is unknown. For this reason, we first calculated the structure of the coelenteramide and compared it with the experimental data. Nitrogen atoms with unshared electron pairs are present in the structure of CLM, where two nitrogen atoms are involved in the formation of both binary bonds with the carbon atom and single bonds (N4 and N7) and the third nitrogen atom has only single bonds (N1). The structure also contains the oxygen atom with unshared electron pairs. The presence of these atoms leads to correlation effects; for this reason, we decided to test the atomic structure of the coelenteramide in the single-electron approximations (PM6 and Hartree–Fock methods) taking into account the correlation corrections (PM6 CI, DFT, MP2). Since this structure is an emitter of a fluorescence reaction, the knowledge of the exact geometry of the structure becomes particularly impor-

Table 1. Deviations of the calculated bond lengths of the CLM structure from the experimental data (in Å). The ordinal numbers of atoms *ri* and *rj* correspond to Fig. 1

ri–*rj*	PM6 CI	HF	DFT	MP2
1–2	0.077	0.023	0.034	0.034
1–9	0.071	0.052	0.054	0.047
2–3	–0.022	–0.036	–0.011	–0.003
4–5	–0.008	–0.039	–0.026	–0.016
4–9	0.025	–0.032	–0.004	0.000
5–6	0.041	0.005	0.022	0.020
6–7	0.004	–0.037	–0.018	–0.012
7–8	0.012	–0.033	–0.012	–0.004
8–9	0.032	–0.001	0.013	0.008
6–10	–0.035	–0.020	–0.025	–0.034
10–11	0.024	0.006	0.019	0.016
11–12	–0.021	–0.033	–0.022	–0.021
12–13	–0.004	–0.021	–0.008	–0.011
13–14	0.029	0.005	0.021	0.027
σ	0.0333	0.0272	0.0246	0.0227
σ^2	0.00111	0.00074	0.00060	0.00052

Table 2. Energy of the tautomeric forms of the coelenterazine (in kilojoule per mole)

	CLZ(1H)	CLZ(7H)	CLZ(2H)
PM6	–470891	–470873	–470876
PM6 CI	–470900	–470894	–470885
HF	–3635586	–3635645	–3635675
DFT	–3656046	–3656071	–3656071
MP2	–3647264	–3647270	–3647415

tant. Even small changes in the atomic structure result in sufficiently large shifts in the absorption spectra. For this reason, it is necessary to determine the degree of the manifestation of the correlation effects in this structure and whether they should be taken into account or only single-electron approximations can be used, which require much shorter time and smaller computer resources.

Table 1 presents the differences between the bond lengths of the coelenteramide structure obtained by various methods and the respective experimental data. The ordinal numbers of the bonds correspond to Fig. 1. The standard deviation σ and variance σ^2 are also given in the table.

According to the presented data, the semiempirical calculations with the use of the PM6 CI procedure are worse than the ab initio calculations. In the system of bonds with N(1), the semiempirical results are strongly different from the experimental structure. However, the calculations of this bond by all of the methods also significantly differ from the experimental data (by 0.05 Å or more). This is probably explained by the fact that oxygen O(3) attracts the electron density and, thus, distorts the structure. It is worth noting that the Hartree–Fock method provides an error in the calculated lengths of the bonds of nitrogen and oxygen (r_{2-3}, r_{6-7}, and r_{7-8}). The semiempirical method somewhat better describes these bonds.

Table 2 presents the energies of tautomeric forms of the coelenterazine in vacuum obtained by various methods. The comparison of the energies obtained by the semiempirical method with and without the configuration interaction indicates that the configuration interaction reduces the energy of the structure.

The energy difference between the CLZ(1H) and CLZ(7H) structures decreases when correlations are taken into account (18 kJ/mole vs. 6 kJ/mole without and with correlations). This is due to the fact that the CLZ(1H) and CLZ(7H) structures are geometrically similar in contrast to the CLZ(2H) structure, where C(2) is in the sp^3 hybridization and the substituent R_1 deviates from the plane of the pyrazine ring by 125°.

A similar situation is also observed in the ab initio methods. The Hartree-Fock calculations provide a rather large difference between the energies of the CLZ(1H) and CLZ(7H) structures up to 59 kJ/mole and acquire a small difference between the energies of the CLZ(7H) and CLZ(2H) structures up to 30 kJ/mole. In the MP2 calculations, the difference between the energies of the CLZ(1H) and CLZ(7H) structures decreases to 6 kJ/mole, but the difference between the energies of the CLZ(7H) and CLZ(2H) structures increases to 145 kJ/mole. The density-functional calculations do not provide the energy difference between the CLZ(7H) and CLZ(2H) structures. However, the CLZ(2H) structure in the MP2 method is much more stable than the CLZ(1H) and CLZ(7H) structures. This means that the electron correlation stabilizes the CLZ(2H) structure, i.e., the coelenterazine structure protonized at the C(2) atom position.

It is worth noting that the CLZ(1H) structure is more stable according to the semiempirical calculations, whereas the CLZ(7H) structure is more stable according to the ab initio calculations. When the electron correlation is taken into account by the MP2 method, the energy difference between the CLZ(1H) and CLZ(7H) structures is small. For this reason, these isomers in protic solvents can be transformed from one form to another; this transformation was observed experimentally [6].

5. These structures in solvents were also calculated with the use of the polarization continuum model [13–16]. The calculations show that a solvent eliminates the difference between the structures. The difference between the energies of the CLZ(1H) and CLZ(7H) structures in the Hartree–Fock method decreases to 37 kJ/mole. When the correlation is taken into account by the density-functional method, the

Table 3. Deviations of the calculated bond lengths of the CLZ(1H) structure from the experimental data (in Å). The ordinal numbers of atoms *ri* and *rj* correspond to Fig. 1

ri–*rj*	PM6 CI	HF	DFT	MP2
1–2	0.028	0.035	0.029	0.014
2–3	–0.022	–0.081	–0.049	–0.033
3–4	0.144	0.104	0.115	0.095
4–5	0.047	0.014	0.018	0.023
5–6	–0.016	–0.047	–0.021	–0.017
6–7	0.057	0.025	0.029	0.029
7–8	–0.033	–0.069	–0.031	–0.011
8–9	–0.038	–0.056	–0.075	–0.090
9–1	0.003	–0.054	–0.020	–0.010
4–9	0.080	–0.007	0.031	0.046
σ	0.0582	0.0583	0.0536	0.0496
σ^2	0.00339	0.00339	0.00288	0.00246

Table 4. Deviations of the calculated bond lengths of the CLZ(7H) structure from the experimental data (in Å). The ordinal numbers of atoms *ri* and *rj* correspond to Fig. 1

ri–*rj*	PM6 CI	HF	DFT	MP2
1–2	–0.030	–0.085	–0.038	–0.009
2–3	0.043	0.036	0.015	–0.007
3–4	0.097	0.017	0.060	0.069
4–5	0.060	0.051	0.038	0.032
5–6	–0.040	–0.070	–0.041	–0.032
6–7	0.091	0.055	0.060	0.048
7–8	0.042	0.037	0.027	0.018
8–9	–0.106	–0.153	–0.121	–0.108
9–1	0.036	0.007	–0.009	–0.020
4–9	0.119	0.062	0.072	0.062
σ	0.0702	0.0729	0.0599	0.0529
σ^2	0.00493	0.00532	0.00358	0.00280

difference decreases strongly to 1–2 kJ/mole in dependence on the solvent. For this reason, it is impossible to determine which structure appears in a certain solvent. To determine this, we compared the bond lengths for the CLZ(1H) and CLZ(7H) structures (see Tables 3, 4).

The comparison of the geometries indicates that the CLZ(1H) structure is in closest agreement with the experimental data. All of the calculation methods show that the disagreement for this structure is smaller than that for the 7H structure.

6. The structure of the CLM emitter of the fluorescence reaction has been analyzed. It has been shown that the inclusion of electron correlations makes it possible to obtain a structure close to the experimentally observed structure; this is important for the correct estimate of the luminescence properties. The calculation without electron correlations in such structures gives an incorrect geometry of the system and, as a result, leads to mistakes in the understanding of the nature of photophysical processes.

Coelenterazine can exist in two isomeric forms, CLZ(1H) and CLZ(7H). Since these structures are geometrically very similar, it is impossible to experimentally determine which tautomeric form exists in solvents. The CLZ(1H) and CLZ(7H) structures are also energetically very similar. The inclusion of correlations reduces the energy difference between these structures and this difference almost vanishes when the solvent is taken into account. The comparison of the atomic structures indicates that the initial structure of the fluorescence process is the CLZ(1H) structure rather than the CLZ(7H) structure, as was accepted previously.

This work was supported by the Russian Foundation for Basic Research (project no. 07-04-00930-a), by the Presidium of the Russian Academy of Sciences (program "Molecular and Cellular Biology"), and by the Siberian Branch, Russian Academy of Sciences (project no. 2).

REFERENCES

1. G. A. Stepanyuk, Z.-J. Liu, S. S. Markova, et al., Photochem. Photobiol. Sci. **7**, 442 (2008).
2. Z.-J. Liu, E. S. Vysotski, C.-J. Chen, et al., Protein Sci. **9**, 2085 (2000).
3. J. F. Head, S. Inouye, K. Teranishi, et al., Nature **405**, 372 (2000).
4. F. G. Predergast, Nature **405**, 291 (2000).
5. K. Mori, S. Maki, H. Niwa, et al., Tetrahedron **62**, 6272 (2006).
6. K. Hori, H. Charbonneau, R. C. Hart, et al., Proc. Natl. Acad. Sct. USA **74**, 4285 (1997).
7. P. Hohenberg and W. Kohn, Phys. Rev. Lett. **136**, B864 (1964).
8. W. Kohn and L. Sham, Phys. Rev. Lett. **140**, A1133 (1965).
9. J. P. Perdew and Y. Wang, Phys. Rev. B. **45**, 13244 (1992).
10. S. J. Vosko, L. Wilk, and M. Nusair, Can. J. Phys. **58**, 1200 (1980).
11. C. Moller and M. S. Plesset, Phys. Rev. **46**, 618 (1934).
12. A. D. Becke, J. Chem. Phys. **109**, 2093 (1998).
13. S. Miertus, E. Scrocco, and J. Tomasi, Chem. Phys. **55**, 117 (1981).
14. S. Miertus and J. Tomasi, Chem. Phys. **65**, 239 (1982).
15. M. Cossi, V. Barone, R. Cammi, et al., Chem. Phys. Lett. **255**, 327 (1996).
16. V. Barone, M. Cossi, B. Mennucci, et al., J. Chem. Phys. **107**, 3210 (1997).

Translated by R. Tyapaev