A QM/MM Study on the Initiation Reaction of Firefly Bioluminescence—Enzymatic Oxidation of Luciferin

Mohan Yu 1,2 and Yajun Liu 1,3,*

1 Center for Advanced Materials Research, Advanced Institute of Natural Sciences, Beijing Normal University at Zhuhai, Zhuhai 519087, China; tmsdtkl@foxmail.com
2 School of Chemistry and Chemical Engineering, Liaoning Normal University, Dalian 116029, China
3 Key Laboratory of Theoretical and Computational Photochemistry, Ministry of Education, College of Chemistry, Beijing Normal University, Beijing 100875, China
* Correspondence: yajun.liu@bnu.edu.cn; Tel.: +86-10-58801160

Abstract: Among all bioluminescent organisms, the firefly is the most famous, with a high luminescent efficiency of 41%, which is widely used in the fields of biotechnology, biomedicine and so on. The entire bioluminescence (BL) process involves a series of complicated in-vivo chemical reactions. The BL is initiated by the enzymatic oxidation of luciferin (LH2). However, the mechanism of the efficient spin-forbidden oxygenation is far from being totally understood. Via MD simulation and QM/MM calculations, this article describes the complete process of oxygenation in real protein. The oxygenation of luciferin is initiated by a single electron transfer from the trivalent anionic LH2 (L3−) to O2 to form [L•− ··· O2•−]; the entire reaction is carried out along the ground-state potential energy surface to produce the dioxetanone (FDO•−) via three transition states and two intermediates. The low energy barriers of the oxygenation reaction and biradical annihilation involved in the reaction explain this spin-forbidden reaction with high efficiency. This study is helpful for understanding the BL initiation of fireflies and the other oxygen-dependent bioluminescent organisms.

Keywords: firefly bioluminescence; luciferin oxidation; mechanism; single electron transfer; QM/MM

1. Introduction

The firefly is the most efficient bioluminescent system for converting chemical energy into light with the extremely high luminescence efficiency of 41% [1]. Its bioluminescence (BL) has been applied widely in biotechnology and biomedical fields [2,3]. The entire firefly BL process can be roughly divided into four stages (Figure 1) [4–7]: oxidation of luciferin (LH2) to a dioxetanone (FDO), decomposition of the dioxetanone to produce excited-state oxyluciferin (bioluminophore), fluorescence emission [8–10] and LH2 regeneration. The mechanism of the last three stages has been explained theoretically in detail [11–16], but there is no comprehensive and reliable theoretical study on the mechanism of the first stage (blue dotted box in Figure 1). Oxygenation is the initial reaction for not only firefly BL but also all oxygen-dependent BL systems. A thorough and reliable investigation of this process is of great significance for understanding the mechanism of all oxygen-dependent BL [17–20].

In 2015, Branchini et al. detected the presence of superoxide anion (O2•−) in the chemical model reaction of firefly BL and suggested that firefly BL is induced by single electron transfer (SET) from LH− to oxygen [21]. There are some corresponding theoretical studies on the oxygenation of luciferin. In 2018, an umbrella sampling molecular dynamics simulation and QM/MM study pointed out the approach of the oxygen moving inside the protein and defined the formation of FDO−, but did not provide details along the reaction path [22]. Our group has investigated the oxygenation process in DMSO, and described a complete process with potential energy curves (PECs) of both ground state (S0) and triplet state (T1) to confirm the SET mechanism [23]. However, this calculation
was not performed in the real protein, and the conclusion could not reflect the essence of the enzymatic reaction. Although there is experimental evidence and corresponding theoretical calculations, the below questions have not yet been thoroughly answered. What is the entire reaction process from $L^2^-\cdot$-AMP + $O_2^*$ to FDO$^-$ in protein? What is the difference between the oxygenation pathway in a solvent and in protein? How does the spin-forbidden reaction of firefly BL occur so efficiently? To answer these three questions is the purpose of this article.

Figure 1. The four stages of a cycle of firefly bioluminescence. * indicates the excited state.

2. Computational Details

For the LH$_2$, the H atom on the hydroxyl in benzothiazole moiety of luciferin is easy to lose in the luciferase environment (Scheme 1). Besides, it has been proven that the dioxetanone decomposition is caused by FDO in its anionic form (FDO$^-$). In addition, the H atom on C$_4$ site of LH-AMP is removed by the adjacent residue in luciferase; this process has already been verified by theoretical calculation [22]. Therefore, the complex $L^2^-\cdot$-AMP (for convenience, this trivalent complex is named $L^3^-$) is the actual reactant in this study. $L^2^-\cdot$ and $O_2^*$ are produced from SET by $L^3^-$ to $O_2$, which induces the subsequent superoxide anion addition reaction. The North American firefly Photinus pyralis luciferase (PDB ID 4G37) [24] was chosen for its structure, which is suitable for providing a starting point for simulating the oxygenation reactions. The 2.5 ns molecular dynamic (MD) simulation was performed to consider protein fluctuation. The initial structures for the QM/MM calculations were started at the snapshot of 1800 ps from the MD trajectory (Figure S1). The chosen QM region contains $L^3^-$ and $O_2$ with a total of 59 atoms, as shown in Figure 2. All QM/MM calculations were performed by a two-layered ONIOM method encoded in the Gaussian16 program [25]. The UM06-2X [26]/6-311G (d, p) [27,28] method with broken-symmetry technology was adopted for the QM region, and the remainder of the system (MM region) used the Amber force field (parm96). The QM/MM calculations and the MD simulation were based on the Gaussian 16 package [25] and AMBER 16 [29], respectively. Computational details are given in the Supporting Information.
Scheme 1. Two half reactions (adenylation and oxygenation) in the initiation of firefly bioluminescence.

Figure 2. QM/MM computational model. The gray ribbon in the background represents the protein environment (left). The atomic labels for key atoms (right). For details see Figure S2.

3. Results and Discussion

The firefly oxygen addition of LH$_3^-$ is initiated by a SET process from LH$_3^-$ to 3O$_2$ to produce two free doublet radicals, LH$_2^-$ and O$_2^*^-$; this reactant complex (RC) has been experimentally [21] confirmed. Since the RC is a biradical ionic pair formed by LH$_2^-$ and O$_2^*^-$ radicals, $[^1LH^2^- ... O_2^*-]$ and $[^1LH^2^- ... O_2^*^-]$ are both possible initial states. For the spin density on the atom for $[^1LH^2^- ... O_2^*^-]$ and $[^1LH^2^- ... O_2^*^-]$, see Table S2. Regarding $[^1LH^2^- ... O_2^*-]$ as the initial state, from RC to the final product FDO$^-$ and AMP (for convenience, we defined FDO$^- +$ AMP as P), three transitions states (TSs) and two intermediates (Ints) were located. The relative energy profiles of SET oxygenation on LH$_3^-$ are shown in Figure 3. The key geometric parameters and Mulliken charges population for all stationary points on LH$_3^- ... O_2^*^-$ moieties are summarized in Table 1. As shown in Figure 3, the RC $[^1LH^2^- ... O_2^*-]$ (RC) is formed by electrostatic force and van der Waals interaction. The O$_4$-O$_6$ bond distance is 1.311 Å in 1RC, which is longer than it is in 3O$_2$ (1.205 Å). This implies that O$_2^*^-$, rather than 3O$_2$, attacks the LH$_2^-$. The charge distribution of O$_2^*^-$ is $-1.00 |e|$ when it is just generated via the SET process. However, partial negative charge transferred from O$_2^*^-$ to LH$_2^-$ in the formation of 1RC, and the Mulliken charge on O$_2^*^-$ is $-0.88 |e|$ and $-2.12 |e|$ on LH$_2^-$, respectively. The expectation value of the $S^2$ operator ($<S^2>$) is 1.00, which indicates that the 1RC has obvious biradical characteristics. With the process of nucleophilic addition, the C$_4$-O$_6$ bond becomes shorter and the C$_8$-C$_4$-O$_6$-O$_2$ dihedral angle gradually twists to the closure of four-membered cyclic peroxide. 1RC forms Int1 through TS1 via a biradical annihilation. This process is accompanied by a small amount of back negative charge transfer (CT) from LH$_2^-$ to O$_2^*^-$ (see Table 1).
The bond length of C₄-O₆ is 1.401 Å in Int1, which does not change much until TS2. Int1 forms Int2 with a four-membered cyclic structure through TS2.

Figure 3. The S₀ and T₁ relative energy profile of the stationary points for oxygenation of L³⁻ in protein by ONIOM (UM06-2X/MM) method. (The unit of energies is kcal mol⁻¹) Here and in later figures, carbon atoms are shown in green, oxygen in red, nitrogen in blue, phosphorus in orange and hydrogen in white.
The key geometric parameters (in Å and degrees) and the Mulliken charges population on L\(^{2--}\) and O\(_2^\bullet\)\(^-\) of the stationary points of L\(^{3--}\) oxygenation by ONIOM method (UM06-2X/MM). (See atomic labels in Figure 2).

|        | O\(_2^\bullet\) | C\(_4\)-C\(_6\) | C\(_6\)-O\(_7\) | C\(_7\)-O\(_8\) | C\(_4\)-O\(_2\) | C\(_9\)-O\(_3\) | C\(_4\)-C\(_4\)-O\(_3\)-O\(_b\) | \(\rho(O_2^\bullet^-)\) | \(\rho(L^{2--})\) |
|--------|----------------|----------------|----------------|----------------|----------------|----------------|---------------------------------|----------------|----------------|
| 1RC    | 1.311          | 1.462          | 1.206          | 1.412          | 2.539          | 3.358          | 119.8                          | −0.88           | −2.12           |
| TS1    | 1.376          | 1.493          | 1.201          | 1.390          | 1.890          | 3.139          | 118.2                          | −0.98           | −2.02           |
| Int1   | 1.436          | 1.568          | 1.200          | 1.370          | 1.401          | 2.682          | 63.0                           | −1.01           | −1.99           |
| TS2    | 1.440          | 1.576          | 1.198          | 1.386          | 1.392          | 2.321          | 49.8                           | −0.92           | −2.07           |
| Int2   | 1.446          | 1.598          | 1.232          | 1.493          | 1.450          | 1.554          | 17.9                           | −0.69           | −2.31           |
| TS3    | 1.447          | 1.552          | 1.185          | 2.056          | 1.476          | 1.416          | 1.3                            | −0.56           | −2.44           |
| P      | 1.445          | 1.530          | 1.178          | 2.617          | 1.494          | 1.375          | −5.1                           | −0.52           | −2.48           |

The main structural changes from Int1 to Int2 are the shortening of the C\(_6\)-O\(_b\) bond and the torsion of the C\(_6\)-C\(_4\)-O\(_3\)-O\(_b\) dihedral angle, which are accompanied by an obvious negative CT from O\(_2^\bullet\)\(^-\) to L\(^{2--}\) (see Table 1). The bond length of C\(_6\)-O\(_3\) does not change much until Int2. After Int2, the C\(_6\)-O\(_3\) bond began to break and leads to P with the departure of the AMP group. This process is accompanied by −0.17 CT from O\(_2^\bullet\)\(^-\) to L\(^{2--}\) and adjustment of the C\(_6\)-C\(_4\)-O\(_3\)-O\(_b\) dihedral angle. (See Table 1).

For the case of \(3[L^{2--} \ldots O_2^\bullet\] as the initial state, the energies of all stationary points at the T\(_1\) state were evaluated at the corresponding S\(_0\) geometries, except \(3\)RC was optimized (Figure S3). The optimized \(3[L^{2--} \ldots O_2^\bullet\] is 0.7 kcal mol\(^{-1}\) higher than \(1[L^{2--} \ldots O_2^\bullet\]. The \(3[L^{2--} \ldots O_2^\bullet\] at the \(1[L^{2--} \ldots O_2^\bullet\] geometry is 1.6 kcal mol\(^{-1}\) higher than \(1[L^{2--} \ldots O_2^\bullet\]. Except for \(3[L^{2--} \ldots O_2^\bullet\], the energy of each stationary point on the T\(_1\) PES is much higher than the corresponding one on the S\(_0\) PES. Obviously, the LH\(_2\) oxygenation reaction occurs on the S\(_0\) potential energy surface (PES). It is worth mentioning that this is quite different from our previous calculation in DMSO [23]. In a solvent, the reaction first along the T\(_1\) PES; after an intersystem crossing (ISC), the reaction takes place on the S\(_0\) PES. Meanwhile, the biradical annihilation occurs along with the ISC process and finally a four-membered cyclic structure is formed (Figure S4).

However, in luciferase, the reaction always occurs on the S\(_1\) PES. The main structural changes from S\(_0\) to S\(_1\) geometries, except \(3\)RC, TS1, and Int1. F-shaped π-π stacking forming between Phe-247 and L\(^{2--}\) is important for stabilizing the aromatic part of the benzothiazole moiety of L\(^{2--}\); this interaction exists all throughout the oxygenation. Besides, His-245 and Lys-433 had a strong H-bond interaction with the O atom on the AMP moiety of L\(^{3--}\). Lys-433 and O\(_7\) of L\(^{3--}\) also formed an H-bond. These H-bond interactions between positively charged residues and substrates lead to the negative charge of O\(_2^\bullet\)\(^-\) and L\(^{2--}\) moiety decrease. The H-bond interaction with 2.453 Å between Gly-246 and O\(_2^\bullet\)\(^-\) affects the relative position of O\(_2^\bullet\)\(^-\) and L\(^{2--}\). With the C\(_6\)-C\(_4\)-O\(_3\)-O\(_b\) dihedral angle gradual torsion, the H-bond between Gly-246 and O\(_2^\bullet\)\(^-\) becomes slightly strong, and then changes to 2.647 Å as Int1 is formed. Meanwhile, His-245 with O\(_b\) forms a weak H-bond. In short, these results demonstrated that the hydrogen-bond interactions between Gly-246 and O\(_2^\bullet\)\(^-\) are essential for the process of oxygen addition. The hydrogen-bonding interactions between Lys-433 and His-245 with AMP moiety of L\(^{3--}\) majorly stabilize the substrate and the negative charge on the O atom of AMP.
Figure 4. The interactions between RC, TS1, Int1 and residues in an active site by ONIOM method (UM06-2X/MM). (The protein environment is not shown).

4. Conclusions

Firefly BL is initiated by the reaction of LH$_2$ + $^3$O$_2$. This is a spin-forbidden reaction with usually a low efficiency, which contradicts the fact that the firefly is the most efficient bioluminescent system for converting chemical energy into light. In this letter, we addressed this issue via MD and QM/MM studies. In luciferase, the entire reaction starts with a SET process and occurs all along the S$_0$ PES. This is quite different for the reaction in a solvent, where the reaction first occurs along the T$_1$ PES and then along the S$_0$ PES after an intersystem crossing (ISC). Moreover, the rate-determined step obviously has a lower barrier in luciferase than in a solvent. The effect of enzymatic catalysis was analyzed. The present theoretical study provides strong theoretical evidence for the SET mechanism of LH$_2$ oxygenation of firefly BL, and is helpful for understanding the BL initiation of the other oxygen-dependent bioluminescent organisms.

Supplementary Materials: The following are available online. Computational details; the protonation states of Histidine; figure of molecular dynamics simulation; figure of QM/MM computational model; figure of optimized structure of RC; figures of computational model in DMSO; figure of S$_0$ and T$_1$ PECs of oxygenation in DMSO; Cartesian coordinates. Table S1: The protonation states of Histidine residue are computed by H++ program at pH = 7.8. Figure S1: RMS deviation of firefly luciferase backbone during 2.5 ns molecular dynamics simulation, Figure S2: QM/MM computational model, Figure S3: The optimized structure of RC, Figure S4: (a) Computational model in DMSO and labels of key atoms. (b) S0 and T1 PECs of oxygenation of A3- in DMSO at UM06-2X/6-31G (d, p) level, Table S2: The spin densities of RC.

Author Contributions: Manuscript conception, Y.L.; writing—original draft preparation, M.Y.; writing—review and editing, Y.L.; calculation, M.Y.; methodology, Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (Grant Nos. 21973005).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available from the authors.

Conflicts of Interest: The authors declare no conflict of interest.

Sample Availability: Samples of the compounds are available from the authors.

References
1. Ando, Y.; Niwa, K.; Yamada, N.; Enomot, T.; Irie, T.; Kubota, H.; Ohmiya, Y.; Akiyama, H. Firefly bioluminescence quantum yield and colour change by pH-sensitive green emission. Nat. Photonics. 2008, 2, 44–47. [CrossRef]
2. Kuchimaru, T.; Iwano, S.; Kiyama, M.; Mitsumata, S.; Kadonosono, T.; Niwa, H.; Maki, S.; Kizaka-Kondoh, S. A luciferin analogue generating near-infrared bioluminescence achieves highly sensitive deep-tissue imaging. Nat. Commun. 2016, 7, 11856. [CrossRef] [PubMed]
