Research Article

Rocco Zerlotti, Aba Losi, Eugenia Polverini*

Oxygen diffusion pathways in mutated forms of a LOV photoreceptor from *Methylobacterium radiotolerans*: A molecular dynamics study

https://doi.org/10.1515/bmc-2022-0013
received February 8, 2022; accepted March 8, 2022

Abstract: *Mr*4511 from *Methylobacterium radiotolerans* is a photoreceptor of the light, oxygen voltage (LOV) family, binding flavin mononucleotide (FMN) as a chromophore. It exhibits the prototypical LOV photocycle, with the reversible formation of an FMN-Cys71 adduct via fast decay of the FMN triplet state. *Mr*4511 has high potential as a photosensitiser for singlet oxygen (SO) upon mutation of C71. *Mr*4511-C71S shows a triplet lifetime (τ_T) of several hundreds of microseconds, ensuring efficient energy transfer to dioxygen to form SO. In this work, we have explored the potential diffusion pathways for dioxygen within *Mr*4511 using molecular dynamics (MD) simulations. The structural model of wild-type (wt) *Mr*4511 showed a dimeric structure stabilised by a strong leucine zipper at the two C-terminal helical ends. We then introduced *in silico* the C71S mutation and analysed transient and persistent oxygen channels. MD simulations indicate that the chromophore binding site is highly accessible to dioxygen. Mutations that might favour SO generation were designed based on their position with respect to FMN and the oxygen channels. In particular, the C71S-Y61T and C71S-Y61S variants showed an increased diffusion and persistence of oxygen molecules inside the binding cavity.

Keywords: photosensitisation, singlet oxygen, blue-light sensors, LOV domain, simulations

*** Corresponding author: Eugenia Polverini,** Department of Mathematical, Physical and Computer Sciences, University of Parma, Parco Area delle Scienze 7/A, 43124 Parma, Italy, e-mail: eugenia.polverini@unipr.it
Rocco Zerlotti: NeuroTrans ETN Network c/o Nanion Technologies GmbH, Ganghoferstraße 70/a, 80339 München, Bayern, Germany; Department of Mathematical, Physical and Computer Sciences, University of Parma, Parco Area delle Scienze 7/A, 43124 Parma, Italy
Aba Losi: Department of Mathematical, Physical and Computer Sciences, University of Parma, Parco Area delle Scienze 7/A, 43124 Parma, Italy

Abbreviations

BL	blue light
BsYtvA-LOV	*Bacillus subtilis* YtvA LOV domain
ET	energy transfer
FAD	flavin adenine dinucleotide
FMN	flavin mononucleotide
FP	fluorescent protein
HB	hydrogen bonds
LOV	light, oxygen voltage
miniSOG	mini Singlet Oxygen Generator
MD	molecular dynamics
PDB	protein data bank
PR	position restrained
RMSD	root mean square deviation
RMSF	root mean square fluctuation
ROS	reactive oxygen species
SO	singlet oxygen
SOPP	singlet oxygen photosensitising protein

Introduction

Light, oxygen, voltage (LOV) proteins are UVA/blue-light (BL) photoreceptors that bind riboflavin derivatives as chromophores, primarily flavin mononucleotide (FMN) or, in fungal proteins, flavin adenine dinucleotide (FAD) [1]. Flavin is embedded within a globular and compact α/β fold of ca. 110 amino acids, named LOV domain [2], in most proteins fused to a variety of effector and regulatory modules [3]. In the dark adapted state (referred to as LOV_{450} from the approximate absorption maximum in the visible range), the chromophore is noncovalently bound within the LOV domain and exhibits a typical green fluorescence [4]; BL illumination triggers a photocycle with the formation of a photoproduct (LOV_{390}) via the 2–4 μs decay of the flavin triplet excited state [5]. In LOV_{390}, FMN (or FAD) is covalently bound at position C4a to a nearby cysteine residue; FMN-N5 becomes

protonated and fluorescence is lost, while the biological activity is activated (Figure 1) [1]. In the dark, LOV_{390} thermally returns to the LOV_{450} state with recovery lifetime (τ_R) ranging from a few seconds to days [6].

The large number of LOV proteins - wide spread in the three superkingdoms - their modular architecture, and the increasing understanding of their activation and signal transduction mechanisms are offering many chances to engineer LOV-based actuators for optogenetics [7–9]. Engineered LOV domains are also exploited as fluorescent reporters, collectively named flavin-binding fluorescent proteins (FbFPs), which, differently from prototypical FPs of the green fluorescent protein type, are small, minimally perturbative, and oxygen independent [4]. A further and more recent application is their exploitation as genetically encoded photosensitisers for singlet oxygen (SO) and other reactive oxygen species (ROS) such as the superoxide anion and the hydroxyl radical [10]. The triplet state of free or LOV-bound FMN has, in fact, an energy content of ca. 200 kJ/mol [11], a perfect configuration to perform efficient energy transfer (ET) to O_2 (a triplet state) via the Dexter mechanism yielding the strong oxidant SO [$^1O_2$ ($a^1\Delta_g$)] that lies at 94 kJ/mol [12]. Photosensitised formation of SO is exploited for a variety of biophysical and medical

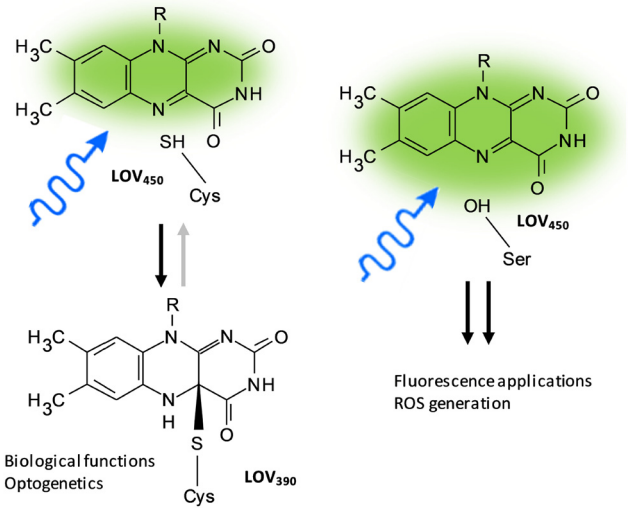

Figure 1: Left, simplified photocycle of a wt LOV protein. The dark-adapted state LOV_{450} is fluorescent with ca. 20% quantum efficiency. After BL illumination, the photoproduct LOV_{390} is formed with ca. 30–50% efficiency via the 2–4 ms decay of the FMN triplet state, fluorescence is lost and biological functions is activated, a feature also exploited in light control of cell functions (optogenetics). Right, when the reactive Cys is modified (here the example of serine) the photoproduct is not formed, fluorescence efficiency is higher (30–50%) and the triplet lifetime becomes longer: these features are suitable for fluorescence microscopy and photodynamic activity with generation of ROS [1].

applications, e.g. photodynamic therapy of cancer and microbe inactivation, cell ablation, induction of cell stress responses in optogenetics, and oxidative tagging in electron microscopy [7,13]. Nevertheless, the Dexter ET process is diffusion limited, and its efficiency depends crucially on the triplet lifetime τ_T of the photosensitiser, i.e. in wild-type (wt) LOV domains, the reactive cysteine nearby FMN quenches the triplet state of the chromophore forming the photoproduct on the short microsecond time scale, thus impeding formation of SO. In LOV domains where this cysteine has been exchanged with serine or alanine, τ_T becomes accordingly longer, in the tens of microseconds range [5,14]. The first LOV domain engineered for the purpose of generating SO was a so-called mini singlet oxygen generator (miniSOG), where six mutations were introduced into the native protein [15], τ_T was 31 µs in aqueous, aerated solution [16], but the quantum yield of formation for SO (Φ_Δ) remained very modest (ca. 0.03) [17]; for comparison, free FMN has Φ_Δ = 0.51 in aqueous solution [18]. The enhanced miniSOG variants developed later, baptised singlet oxygen photosensitising proteins (SOPPs), notably have shown that tryptophan 81 and glutamine 103 (miniSOG numbering) represent major quenchers of the FMN triplet state, together with poorly identified tyrosine residues: beyond the additional (with respect to miniSOG) Q103V mutation, the best performing variant, SOPP3, bears the W81L change ensuring τ_T = 135 µs and Φ_Δ = 0.61 (aerated D_2O buffer) [19]. Interestingly, W81, localised at an edge-to-edge distance of ca. 11 Å, is conserved in about 75% of sequenced LOV domains at this position and in the absence of the reactive cysteine represents the major quencher of the FMN triplet state by electron transfer forming a transient radical pair [20,21]. Residual problems, also in the enhanced SOPP variants of miniSOG, are the instability and the bleaching of the chromophore both *in vitro* and *in vivo* under prolonged illumination [22,23]. Recently, the *Mr*4511 LOV protein from the plant symbiont *Methylobacterium radiotolerans* was characterised and shown to be a promising photosensitiser for SO with Φ_Δ ca. 0.2 upon the single mutation of the reactive Cys71 into Gly or Ser [24]. *Mr*4511 (164 aa) is built of a standalone LOV domain with short N- and C-terminal flanking regions [25]. The position corresponding to W81 of miniSOG is occupied by Q112, and this undoubtedly confers an unusually long τ_T to the C71S and C71G variants (240 and 340 µs, respectively) as demonstrated by inserting a Trp residue at position 112, which shortens τ_T by one order of magnitude, concomitantly reducing Φ_Δ to ca. 0.01 [24]. In addition, *Mr*4511 has unusual stability towards denaturation with urea, a promising feature for applications *in vivo* [24]. The prominent role of Trp residues and, with less efficiency, tyrosines in quenching the FMN triplet state in *Mr*4511 has

been highlighted in a recent work: Trp or Tyr residues in the proximity of the chromophore (naturally occurring or artificially inserted) shorten τ_T, whereas Phe residues are inert [26]. The results with different *Mr*4511-C71S variants also show that τ_T and Φ_Δ are positively correlated.

In order to understand at the molecular level the propensity of *Mr*4511 to be engineered as a photosensitiser for SO and suggest further improvements, in this work, we have built a stable dimeric 3D model for the *Mr*4511 structure, which till now was never been resolved. After creating its C71S variant *in silico*, we have investigated the presence of transient and persistent oxygen channels using molecular dynamics (MD) simulations and analysed the diffusion of molecular oxygen inside the protein. Mutations that might favour SO generation were then designed based on their position with respect to the FMN and the oxygen channels, taking into account the ability of certain amino acids to quench the FMN triplet state and SO. Such mutations were inserted in the C71S variant and simulated in the presence of oxygen to compare its diffusibility and persistence in the binding pocket, leading to a double variant that, enhancing oxygen diffusion to the FMN-binding site, promise to be a stronger and efficient photosensitiser for SO.

Methods

Modelling of *Mr*4511 LOV domain and its variants

The sequence of the *Mr*4511 LOV domain was obtained by Uniprot Knowledgebase (UniProtKB code: B1M516) [27]. The model was previously built in a monomeric form by the Swiss-Model server [28], using as a template the structure of the aureochrome 1a LOV domain (protein data bank [PDB] ID: 5a8b [29]) from *Phaeodactylum tricornutum* that showed the highest sequence identity (48%) and a total coverage of the core region. This template had the long C-terminal helix bent along the β-sheet side. However, the built monomeric structure for *Mr*4511 turned out to be unstable at MD simulations due to wide oscillations of the C-terminal helix. Therefore, considering the hypothesis of a dimeric structure and looking for a higher coverage of the C-terminal region as well as the core, a new template was found in the PpSB1-LOV BL photoreceptor protein (PDB ID: 5j3w [30]) of *Pseudomonas putida* that has a lesser sequence identity (34.7%) but a higher C-terminal coverage. This template is not only in a dimeric form, but the dimer has a pairing of the C-terminal helices held together by a leucine zipper. A similar leucine zipper, involving even more leucine residues (four pairs of residues) than *Pseudomonas*, is present also in *Mr*4511, supporting the template choice and the hypothesis that this motif could help to stabilise the structure. The raw model was built using the Swiss-Pdb Viewer software [31] to check the alignment and obtain a dimeric structure. It was then submitted to the the Swiss-Model server [31] to obtain a refined structure. The refined model included residues 19–164, lacking only the first 18 residues, therefore allowing the formation of at least a fragment of an N-terminal helix. The two FMN cofactors were inserted in the binding pocket, building the coordinates by superimposition with the 5j3w structure. The last residues of the C-terminal helix (155–164) were built using the Swiss-Pdb Viewer software [31] and the rotamers of key residues (Asn103, Asn113, and Gln134), forming the H-bond network that interacts with the isoalloxazine rings, were refined with the same software. The quality of the model was assessed both by the Swiss-Model server and by the Swiss-Pdb Viewer software (threading energy and structural analyses). Nevertheless, the model underwent an MD simulation to regularise and equilibrate the structure until a stable one was reached (see below).

The C71S mutation was introduced using the Swiss-Pdb Viewer software on the mean structure obtained as described below, whereas Y61T and Y61S mutations were introduced on the mean structure obtained in the same manner for the C71S variant. Finally, three models were built: with C71S mutation alone and with C71S-Y61T or C71S-Y61S double mutations.

MD simulations

MD simulations were performed with the GROMACS 2019.4 software package [32], using the Charmm27 force field [33] with modified parameters for the FMN cofactor [34,35]. The system was simulated at neutral pH; histidines were kept in the neutral form, while FMN had a net charge of –2, at the phosphate group level. Protein was embedded in a rectangular box, built considering a layer of 1 nm around the solute. The box was filled by water molecules with the Solvate plugin of GROMACS, using the TIP3P water model, which is optimized for Charmm force fields, and molecules randomly inserted inside the binding pocket were manually deleted [35]. Na^+ and Cl^- ions were added to reach a physiological salt concentration of 0.1 M and to preserve the system neutrality. After an energy minimisation of the whole system, the solvent was allowed to relax around protein during 100 ps

long position-restrained (PR) simulations. Subsequently, full-motion simulations were carried out for 200–300 ns, with a timestep of 2 fs. Each simulation was replicated at least two times, with the exception of the ones of C71S-Y61T and C71S-Y61S variants in water because their stability was rapidly assessed. A table with the performed simulations and their duration is shown in the Supplementary Material (Table S1). Both PR and full MD simulations were performed in the PTN ensemble at T = 300 K and P = 1 bar using a velocity rescaling thermostat and a Berendsen barostat for temperature and pressure coupling. Periodic boundary conditions were applied to the system.

The convergence to a stable conformation was assessed by root mean square deviation (RMSD) profiles and matrices calculated on Cα atoms using GROMACS subroutines. On the more stable simulation of the wild-type protein (namely, wt_r3; Table S1 and Figure S1C), a mean structure was calculated on the frames of the convergent part of the trajectory in the following way: after the calculation of the arithmetic mean of the coordinates, the structure showing the minimum value of RMSD with respect to the arithmetic mean was selected among those used for the calculation and used as a "real" mean structure, belonging to the structures of the trajectory. This mean structure was subsequently used for the building of the C71S variant and for the analyses of its cavities and tunnels. On the more stable simulation of the C71S variant (namely, C71S_r2; Table S1 and Figure S3B), a mean structure was calculated on the frames of the convergent part of the trajectory in the same way described above.

For MD simulations in the presence of oxygen, the pdb file for the oxygen molecule was extracted from the pdb file of oxymyoglobin (PDB ID: 1A6M). Parameters for molecular oxygen were already embedded in the Charmm27 force field. Fifty oxygen molecules were randomly inserted in the box of solvent surrounding the protein to ensure a relevant probability to observe the entrance of oxygen inside the protein cavities in a reasonable simulation time.

The structure and trajectory analyses were performed with VMD [36] and Swiss-Pdb Viewer software and using GROMACS subroutines. Volume maps for water and oxygen occupancy were calculated by VMD software, skipping the first 5 ns of each trajectory to ensure that the system has reached equilibrium. A neighbourhood of 8 Å around the O4 atom of FMN was considered to comprehend all the binding pocket.

Calculation of protein cavities and tunnels

To investigate protein cavities and tunnels forming during the MD trajectories, the software Caver Analyst 2.0 [37] was used. To calculate cavities in the *Mr*4511-C71S structure, a probe radius of 1.4 Å was used, including the cofactor FMN in the structure. To investigate the tunnel formations along a trajectory, the cavities found in the protein core at no more than 5 Å from the alloxazine rings were chosen. A neighbourhood of 3 Å of the centre of the cavity was used as a starting point for tunnel search. The mean radius of the tunnel was set to 1.4 Å, with a bottleneck (the minimum radius) of 0.9 Å. A clustering of the results was made, with a threshold of 3.5 Å for the mean distance among tunnels in the same cluster.

Results

Model of *Mr*4511 LOV domain

The stability of the dimeric model of the *Mr*4511 LOV domain was assessed by three 300 ns long MD simulation replicas, performed under standard conditions (T = 300 K and P = 1 bar, Table S1). By means of RMSD plots and matrices, a structural convergence was verified (Figures S1 and S2), which allowed us to build a mean structure that could be the starting point for subsequent simulations (Section 2). Root mean square fluctuation (RMSF) plots show as well peaks of flexibility in the predicted regions.

The typical structure of LOV domains consists of a core formed by five β-strands and four helices, named Aβ–Bβ–Cα–Dα–Eα–Fα–Gβ–Hβ–Iβ [3] (Figure 2). The regions flanking this core at the N- and C-terminals exhibit variable α-helical structures, named A′α and Jα, respectively. The pocket hosting the FMN chromophore is lined by two core helices, one of which is a 3_{10}-helix. Instead of two arginine residues protruding from these two helices as in the other LOV photoreceptors, the structural model of *Mr*4511 shows three arginines (Arg72, Arg88, and Arg92) that keep in place the ligand through salt bridges formed with its phosphate group (Figure 3a). The network of H bonds, interacting with the polar part of the isoalloxazine ring of FMN and involving a conserved triad made by two asparagine residues (Asn103 and Asn113) and the so-called "flipping" glutamine (Gln134) [35,38], is also present and contributes to the stability and

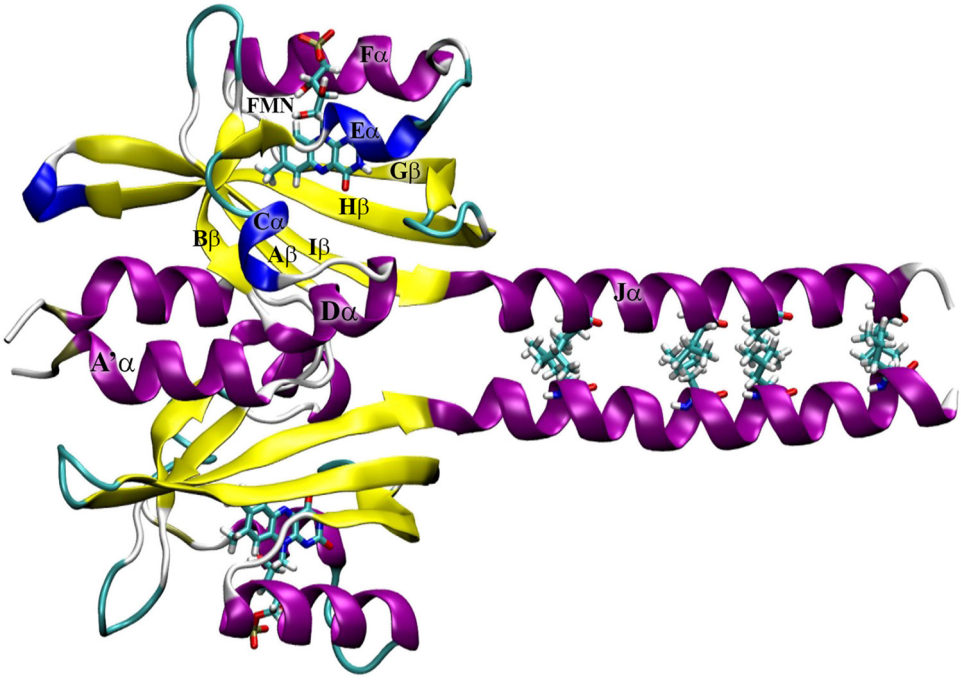

Figure 2: *Mr*4511 dimeric model, coloured by secondary structure. The four leucine residues forming the leucine zipper at the long C-terminal helices are highlighted in stick and coloured by atom type. The secondary structure elements are labelled.

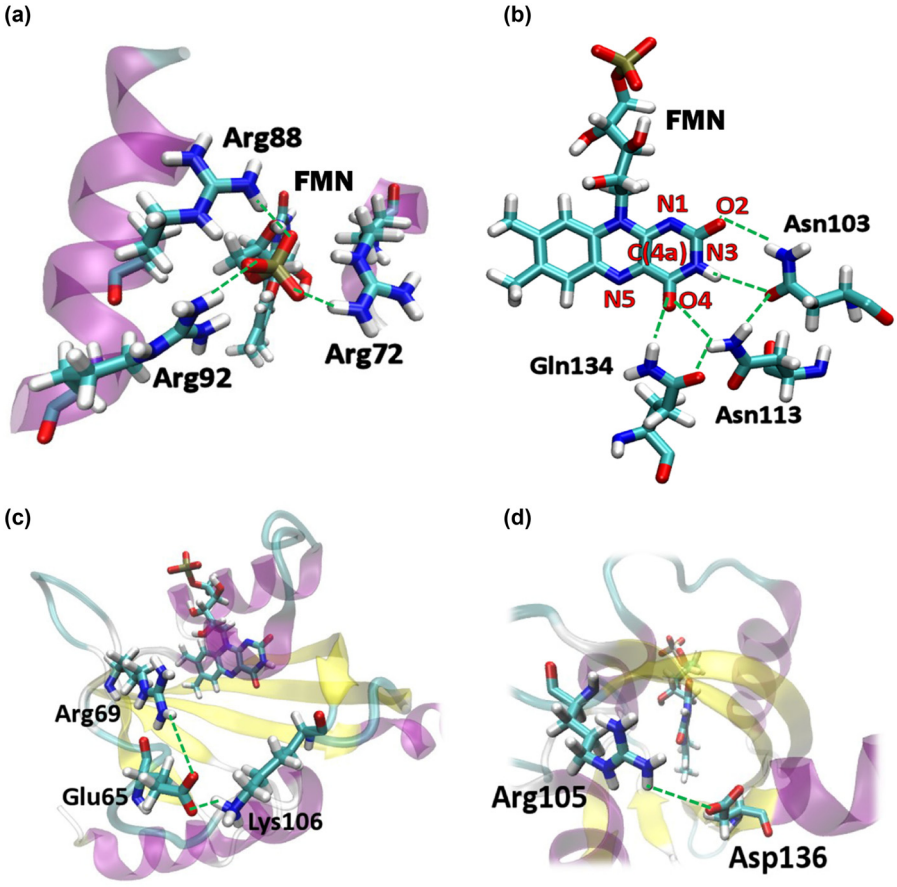

Figure 3: Key structural features of LOV domain in *Mr*4511. (a) The three arginine residues interacting with FMN phosphate group. (b) The residues triad establishing an H-bond network with FMN. (c) and (d) The two conserved salt bridges lining the doorways to the binding pocket.

functionality of the ligand (Figure 3b). This last residue, although not fully conserved in the LOV series, participates in signal transmission by reorienting its lateral chain after the formation of the photoproduct [39]. We note that it was recently shown that even when this residue is naturally absent or artificially mutated, LOV domains undergo light-induced conformational changes and remain partially functional, possibly due to the transient ingress of water molecules in the photoproduct, as suggested by MD simulation results [40].

Interestingly, as mentioned in Section 2, in the C-terminal helix a leucine pattern was observed involving a high number of leucine residues (four residues: 143, 150, 154, and 161). The presence of this pattern – the so-called ZIP domain – supported the hypothesis of a dimeric model for *Mr*4511 in which the long terminal helix is extended to allow the formation of a leucine zipper (Figure 2).

In the MD trajectories, two conserved salt bridges, namely, Lys106–Glu65 and Arg105–Asp136, are observed to play an important role in the permeability to water (and – as we will discuss below – to oxygen molecules) of the binding pocket (Figure 3c and d). Such salt bridges line what were already recognised as doorways for water in a previous study on the YtvA LOV domain from *Bacillus subtilis* (*Bs*YtvA-LOV) [35]. Differently from *Bs*YtvA-LOV, however, where the binding cavity was rather impermeable to water molecules and where their penetration strongly perturbed the conserved FMN–Gln–Asn–Asn H-bond network, in *Mr*4511, water can more easily enter and exit from the binding pocket, and the perturbation of hydrogen bonds (HB) is reversible. The Arg105–Asp136 salt bridge has a pivotal role for water ingress (and oxygen – as we will see below) being a gate for one of the main cavities from which originates one of the most persistent tunnels connecting the solvent with the binding pocket. Regarding the Lys106–Glu65 salt bridge, a particular behaviour is detected. In *Bs*YtvA-LOV, the basic residue Lys97 – analogous to *Mr*4511 Lys106 – was involved in an alternative salt bridge with two acidic residues, both on the opposite Cα helix, maintaining in this way a limited entrance to the binding cavity. On the contrary, in *Mr*4511 it is the acidic residue Glu65 that switches between Lys106 and Arg69, in the same Cα helix of Glu65 (Figure 3c). Finally, the presence of a very stable water molecule forming a bridge between the NH backbone atom of Cys71 (the cysteine forming the photoproduct in the wt protein) and the O2′ atom of the ribityl chain of FMN is notable. A water molecule in this position is conserved in all the crystal structures of LOV domains [41], supporting the reliability of our structural model.

Analysis of cavities and tunnels and simulations of dioxygen diffusion

*Mr*4511-C71S variant

To prevent the formation of the photoproduct and allow ET from the FMN triplet state to dioxygen, generating SO, the reactive Cys71 was mutated *in silico* in a serine residue. This variant has been previously experimentally investigated [24]. The structure was checked for stability by means of MD simulation replica 200 ns long. Again, RMSD plots and matrices assessed the reaching of convergence with no modifications in the structure with respect to the wild-type protein, and RMSF plots reported very low oscillations in the expected flexible regions (loops and termini) (Figure S3).

The *Mr*4511-C71S structure was first investigated for the presence of cavities. Three cavities were found from which O_2 could diffuse in the vicinity of FMN, thus minimising quenching of the triplet state by amino acid side-chains (in particular, His, Met, Trp, and Tyr [20,26,42]) and enhancing local oxygen concentration. The first cavity is lined by the Arg105–Asp136 salt bridge and the residues involved in the H-bond network; the second lies at one side of FMN, behind Eα and Fα helices; and the third lies

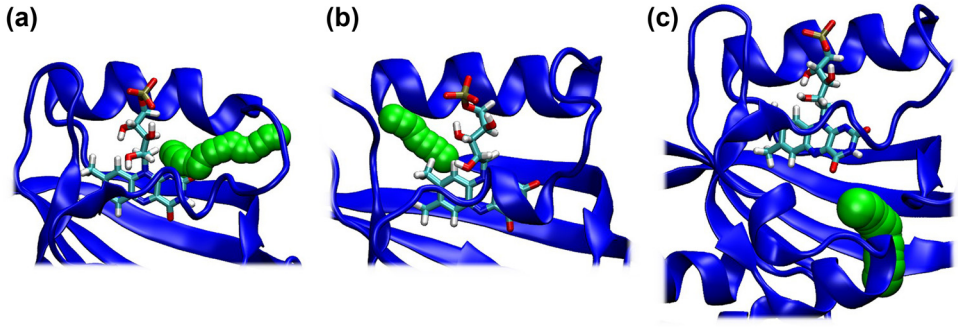

Figure 4: More persistent tunnels leading to the three binding cavities and connecting outside to the FMN cofactor in the C71S variant.

at the other side of FMN, near the methylic substituents of alloxazine.

Starting from these cavities, the formation of tunnels suitable for oxygen diffusion into the binding pocket was explored along the MD trajectories, and the similar ones (i.e. below a threshold of 3.5 Å as a maximum centre-to-centre distance, see Section 2) were clustered. Representative tunnels of the more persistent clusters were further analysed. Three tunnels were identified, one for each cavity, which connect the outer surface of the protein to the FMN-binding pocket, leading in the vicinity of the cofactor (Figure 4). It is noticeable that tunnels form dynamically within the protein, changing their radius with the thermal motions of the sidechains of the residues. Subsequently, the ability of oxygen to diffuse inside the binding pocket was investigated by means of MD simulations performed after the insertion of 50 O_2 molecules in the box of water surrounding the protein. In the trajectories, multiple events of ingress of oxygen molecules in the binding cavity were observed, all passing through the three identified tunnels. In addition, the permanence of oxygen inside the binding cavity is very long, reaching more than 58 ns (Table 1), and concerns, in particular, the neighbourhood of the FMN cofactor.

This is especially true for the molecules diffusing through the tunnel created thanks to the breaking of the Arg105–Asp136 salt bridge and penetrating into the first cavity mentioned above, lined by the residues of the H-bond network. Looking at the volume maps of the occupancy of oxygen inside the binding pocket (Figure 5), this region has the highest values. It is worth noting that the highest mobility of oxygen in the whole binding pocket leads to lower values of occupancy in a limited region with respect to water. Water is, in fact, less movable because of its ability to form interactions due to its polarity. It is just for its different polarity that the high presence of O_2 near the HB network does not perturb it as water does. In fact, in the analogous volume maps calculated for water (Figure 5), the highest occupancy is again in the HB network region, where, on the contrary, water perturbs the site and is trapped by H-bond interactions both with the sidechain of the conserved triad and with FMN itself.

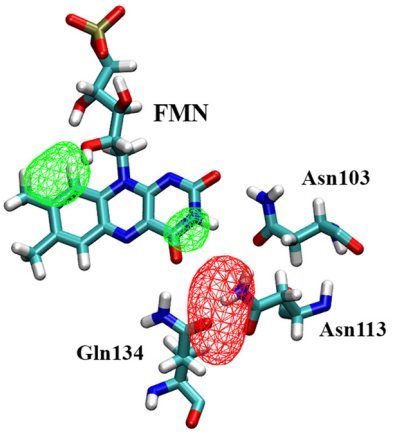

Figure 5: Occupancy volume maps of oxygen (red, isosurface value 5%) and water (green, isosurface value 25%) in the binding pocket of C71S variant (replica C71S_r1 is shown as an example). The volume maps were averaged on the whole trajectory after skipping the first 5 ns. A neighbourhood of 8 Å around atom O4 of FMN cofactor includes the whole binding pocket. The FMN and the HB networking triad are shown as a reference.

Interestingly, it is worth noting the presence of the conserved water molecule near residue 71, as observed for the wt protein, in this case forming a bridge between the hydroxyl group of mutated Ser71 in place of cysteine, and the O2′ atom of the ribityl chain of FMN.

*Mr*4511-C71S-Y61T and *Mr*4511-C71S-Y61S variants

In the MD simulations previously described, it was observed that oxygen molecules explored the surface region around the Glu65–Ly106 salt bridge without being capable of penetrating inside due to the presence of the bulky sidechain of Tyr61. Therefore, to design a protein variant even more efficient in the production of SO favouring oxygen entrance,

Table 1: Presence and persistence of O_2 molecules in the binding pocket

Variant	Presence of O_2 in the binding pocket[a] (% of frames)	Highest persistence time[b] (ns)	Number of different molecules entering[c]
C71S	35	58.3	18
C71S-Y61T	40	31.4	39
C71S-Y61S	58	72.2	22

[a]The Binding pocket is calculated considering a neighbourhood of 8 Å around atom O4 of FMN. Values are averaged on the two MD replicas.
[b]The highest persistence time between the two replicas is reported.
[c]Each number corresponds to the replica with the persistence time reported in the previous column.

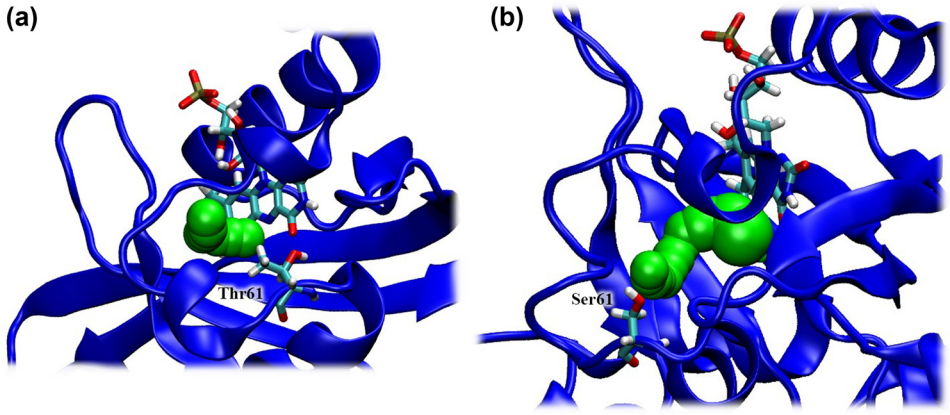

Figure 6: New tunnel connecting outside to the FMN cofactor in the C71S-Y61T (a) and C71S-Y61S (b) variant.

two mutants of Tyr61 were proposed: a threonine, which with lower steric hindrance still maintains both polarity and hydrophobicity; and a serine, to further reduce the size but preserve polarity features. The choice of mutating a tyrosine is supported by the fact that this kind of residue is efficient to quench the triplet state of FMN [43]. Both C71S-Y61T and C71S-Y61S variants' stability was confirmed again by MD simulations (Figure S4), and their cavity and tunnels were explored. As expected, in both variants, a further cavity was detected between the mutated residue and the isoalloxazine rings of FMN, and new tunnels from this cavity to the outside were formed during the trajectories (Figure 6). MD simulations in the presence of oxygen showed in both variants high diffusion of O_2 through these new tunnels, as well as through the other ones previously detected, increasing the percentage of frames in which oxygen is present in the FMN-binding pocket, higher for C71S-Y61S than for C71S-Y61T (Table 1). Comparing, among the three variants, the number of different O_2 molecules entering in the binding cavity with the highest persistence time of the same molecule in it (Table 1), it seems that for C71S-Y61T, more distinct oxygen molecules enter but less persistently; *vice versa*, for the C71S-Y61S variant, the highest persistence time of a single molecule corresponds to a lower number of distinct oxygen molecules able to enter in the pocket. It seems that the channel in C71S-Y61T could be more permeable, with the O_2 molecules entering and exiting more easily than in the C71S-Y61S variant, which, on the contrary, tends to keep oxygen inside. The methyl group of mutated residue Thr61 is, in fact, surrounded by a hydrophobic cluster that maintains the sidechain in an orientation that keeps the channel open most of the time. The occupancy volume

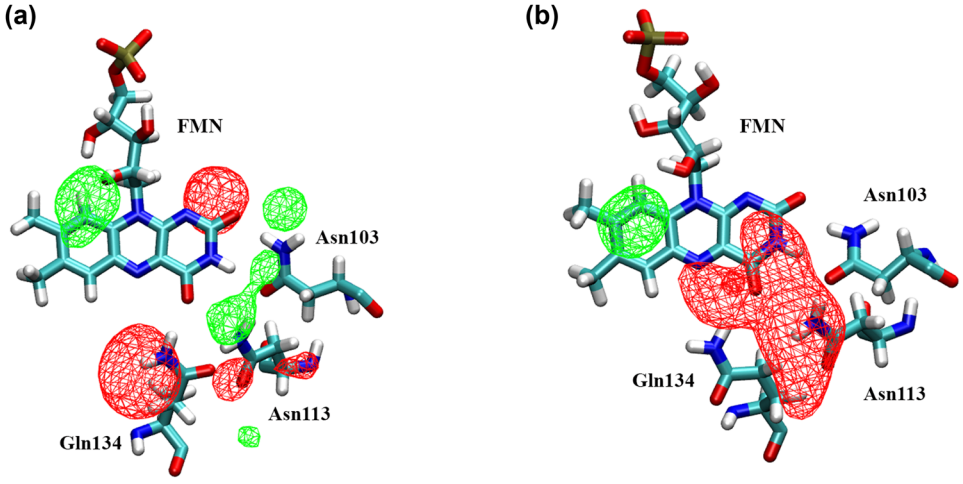

Figure 7: Occupancy volume maps of oxygen (red, isosurface value 5%) and water (green, isosurface value 25%) in the binding pocket of C71S-Y61T (a) and of C71S-Y61S (b) variants (replica C71S-Y61T_oxy_r1 and C71S-Y61S_oxy_r1 are shown as an example). The volume maps were averaged on the whole trajectory after skipping the first 5 ns. A neighbourhood of 8 Å around atom O4 of FMN cofactor includes the whole binding pocket. The FMN and the HB networking triad are shown as a reference.

maps confirm this behaviour, showing that oxygen tends to be spread around the binding pocket in the threonine variant, and assembled around the HB network in the serine one, which for both cases agrees with the higher presence in space and time of oxygen in the binding site with respect to the C71S variant (Figure 7).

Interestingly, water is not able to pass through the new tunnels to diffuse inside the binding pocket because it remains trapped in water bridges formed among residues Glu65, Ly106, and Arg69 of the broken salt bridge. Therefore, the cavity and tunnels created by the mutations have proved to be effective in oxygen selection with respect to water, preventing the filling of the binding cavity by water molecules and the consequent perturbation of the key interactions. Also in these double variants, the conserved water molecule bridging the hydroxyl group of mutated Ser71 and the O2′ atom of the ribityl chain of FMN is still present.

The volume map of water occupancy in the *Mr*4511-C71S-Y61T variant (Figure 7a) shows higher presence of water in the region of the H-bond network with respect to the C71S-Y61S variant due to the lower stability of the Arg105–Asp136 salt bridge (Figure 7b).

Conclusion

A structural model for *Mr*4511 was built and, in the folded sequence, all the structural features of a LOV domain were detected. The dimeric structure was supported by the presence of a leucine zipper to which another leucine must be added from the terminal residues of the Iβ strand. This hydrophobic interfacial sidechain packing could contribute to the high stability of this protein that was revealed under denaturing conditions as well [24,25]. More locally, the presence of a third arginine residue blocking the FMN in the binding site by interaction with its phosphate group likely assures the stability of the cofactor and the maintenance of its function also under denaturing conditions. The robustness of the *Mr*4511 structure can also be seen as resilience to perturbations: the permeability of the binding site to water, which is particularly high also with respect to other studied LOV domains such as *Bs*YtvA [35], frequently leads to a perturbation of the critical H-bond network, that is able to recover its stability when water detaches.

As expected, the same features are present in the three investigated variants, with some other remarkable advantages for photosensitisation. The *Mr*4511-C71S variant, together with the presence of glutamine instead of tryptophan in a place near FMN, has led to a long τ_T of several hundreds of microseconds [24]. The permeability of the binding pocket to dioxygen, allowed through channels, and its persistence inside the cavity ensure the quite high Φ_Δ experimentally observed [24]. After that, we designed new mutations to improve these already advantageous conditions enhancing SO production. The designed mutations share the dual effect of removing a quenching tyrosine and creating the space for new oxygen channels. To reach these purposes, the chosen residue was Tyr61, which was mutated *in silico* into threonine and serine, keeping as far as possible its physico-chemical features. The identified mutation permits the formation of a new channel that allows the oxygen entrance in a selective manner. Its doorway is, in fact, lined by a particular salt bridge that, switching between two different cations, traps water preventing its entrance. The two new channels seem to have different characteristics, permitting a more rapid exchange of oxygen in the C71S-Y61T variant and keeping the O_2 molecules inside the binding pocket in the C71S-Y61S variant. However, the two designed variants enhance O_2 diffusion to the binding cavity, allowing the access of a higher number of dioxygen molecules with a higher permanence, so enhancing the probability that the Dexter-type ET occurs. Overall, they promise to be good candidates as efficient, LOV-based genetically encoded photosensitisers for SO. Complementary approaches to this timely topic include the minimisation of internal quenching by aromatic amino acids [19,26] and the introduction of mutations that enhance binding preference towards riboflavin, leading to a better yield of SO production [44]. The increasing amount of experimental data, together with MD simulations as presented here, and molecular characterisation of novel LOV domains are thus building up a palette of genetically encoded photosensitisers with improved robustness and efficiency for *in vivo* applications.

Acknowledgements: This research benefits from the High-Performance Computing (HPC) facility of the University of Parma, Italy. We thank Dr. Sofia Secci for her contribution in the model stability.

Conflict of interest: Authors state no conflict of interest.

Data availability statement: The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

References

[1] Losi A, Gärtner W. Solving blue light riddles: new lessons from flavin-binding LOV photoreceptors. Photochem Photobiol. 2017;93(1):141–58.

[2] Stuffle EC, Johnson MS, Watts KJ. PAS domains in bacterial signal transduction. Curr Opin Microbiol. 2021;61(JUN):8–15.

[3] Glantz ST, Carpenter EJ, Melkonian M, Gardner KH, Boyden ES, Wong GK-S, et al. Functional and topological diversity of LOV domain photoreceptors. Proc Natl Acad Sci U S A. 2016;113(11):E1442–51.

[4] Buckley AM, Petersen J, Roe AJ, Douce GR, Christie JM. LOV-based reporters for fluorescence imaging. Curr Opin Chem Biol. 2015;27(C):39–45.

[5] Guo H, Kottke T, Hegemann P, Dick B. The Phot LOV2 domain and its interaction with LOV1. Biophys J. 2005;89(1):402–12.

[6] Arinkin V, Granzin J, Krauss U, Jaeger K, Willbold D, Batra-Safferling R. Structural determinants underlying the adduct lifetime in the LOV proteins of *Pseudomonas putida*. FEBS J. 2021;288(16):4955–72.

[7] Losi A, Gardner KH, Möglich A. Blue-light receptors for optogenetics. Chem Rev. 2018;118(21):10659–709.

[8] Seong J, Lin MZ. Optobiochemistry: genetically encoded control of protein activity by light. Annu Rev Biochem. 2021 Jun 20;90(1):475–501.

[9] Terazima M. Time-resolved detection of association/dissociation reactions and conformation changes in photosensor proteins for application in optogenetics. Biophys Rev. 2021;13(6):1053–9.

[10] Trewin AJ, Berry BJ, Wei AY, Bahr LL, Foster TH, Wojtovich AP. Light-induced oxidant production by fluorescent proteins. Free Radic Biol Med. 2018;128(Nov):157–64.

[11] Gauden M, Crosson S, van Stokkum IHM, van Gondelle R, Moffat K, Kennis JTM. Low-temperature and time-resolved spectroscopic characterization of the LOV2 domain of Avena sativa phototropin 1. In: Avrilleir S, Tualle JM, editors. Femtosecond Laser Applications in Biology. SPIE Bellingham, WA5463; 2004. 97–104.

[12] Bregnhøj M, Blázquez-Castro A, Westberg M, Breitenbach T, Ogilby PR. Direct 765 nm optical excitation of molecular oxygen in solution and in single mammalian cells. J Phys Chem B. 2015 Apr 30;119(17):5422–9.

[13] Souslova EA, Mironova KE, Deyev SM. Applications of genetically encoded photosensitizer miniSOG: from correlative light electron microscopy to immunophotosensitizing. J Biophotonics. 2017;10(3):338–52.

[14] Kottke T, Heberle J, Hehn D, Dick B, Hegemannt P. Phot-LOV1: Photocycle of a blue-light receptor domain from the green alga *Chlamydomonas reinhardtii*. Biophys J. 2003;84(2 I):1192–201.

[15] Shu X, Lev-Ram V, Deerinck TJ, Qi Y, Ramko EB, Davidson MW, et al. A genetically encoded tag for correlated light and electron microscopy of intact cells, tissues, and organisms. PLoS Biol. 2011;9(Apr):e1001041.

[16] Westberg M, Bregnhøj M, Etzerodt M, Ogilby PR. Temperature sensitive singlet oxygen photosensitization by LOV-derived fluorescent flavoproteins. J Phys Chem B. 2017;121(12):2561–74.

[17] Ruiz-González R, Cortajarena AL, Mejias SH, Agut M, Nonell S, Flors C. Singlet oxygen generation by the genetically encoded tag minisog. J Am Chem Soc. 2013;135(26):9564–7.

[18] Baier J, Maisch T, Maier M, Engel E, Landthaler M, Bäumler W. Singlet oxygen generation by UVA light exposure of endogenous photosensitizers. Biophys J. 2006;91(4):1452–9.

[19] Westberg M, Bregnhøj M, Etzerodt M, Ogilby PR. No photon wasted: an efficient and selective singlet oxygen photosensitizing protein. J Phys Chem B. 2017;121(40):9366–71.

[20] Westberg M, Etzerodt M, Ogilby PR. Rational design of genetically encoded singlet oxygen photosensitizing proteins. Curr Opin Struct Biol. 2019;57(Aug):56–62.

[21] Ding Y, Kiryutin AS, Zhao Z, Xu QZ, Zhao KH, Kurle P, et al. Tailored flavoproteins acting as light-driven spin machines pump nuclear hyperpolarization. Sci Rep. 2020;10(1):18658.

[22] Torra J, Lafaye C, Signor L, Aumonier S, Flors C, Shu X, et al. Tailing miniSOG: structural bases of the complex photophysics of a flavin-binding singlet oxygen photosensitizing protein. Sci Rep. 2019;9(1):2428.

[23] Mogensen DJ, Westberg M, Breitenbach T, Etzerodt M, Ogilby PR. Stable transfection of the singlet oxygen photosensitizing protein SOPP3: examining aspects of intracellular behavior†. Photochem Photobiol. 2021;97(6):1417–30.

[24] Consiglieri E, Xu Q, Bregnhøj M, Westberg M, Ogilby PR, Losi A. Single mutation in a novel bacterial LOV protein yields a singlet oxygen generator. Photochem Photobiol Sci. 2019;18(11):2657–60.

[25] Consiglieri E, Xu QZ, Zhao KH, Gärtner W, Losi A. The first molecular characterisation of blue- and red-light photoreceptors from *Methylobacterium radiotolerans*. Phys Chem Chem Phys. 2020;22(22):12434–46.

[26] Ding Y, Zhao Z, Matysik J, Gärtner W, Losi A. Mapping the role of aromatic amino acids within a blue-light sensing LOV domain. Phys Chem Chem Phys. 2021;23(31):16767–75.

[27] Bateman A, Martin MJ, Orchard S, Magrane M, Agivetova R, Ahmad S, et al. UniProt: The universal protein knowledgebase in 2021. Nucleic Acids Res. 2021;49(D1):D480–9.

[28] Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 2018 Jul 2;46(W1):W296–303.

[29] Banerjee A, Herman E, Kottke T, Essen LO. Structure of a native-like aureochrome 1a LOV domain dimer from Phaeodactylum tricornutum. Structure. 2016;24(1):171–8.

[30] Röllen K, Granzin J, Panwalkar V, Arinkin V, Rani R, Hartmann R, et al. Signaling states of a short blue-light photoreceptor protein PpSB1-LOV revealed from crystal structures and solution NMR spectroscopy. J Mol Biol. 2016;428(19):3721–36.

[31] Guex N, Peitsch MC. SWISS-MODEL and the swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis. 1997;18(15):2714–23.

[32] van der Spoel D, Lindhal E, Hess B, Groenhof G, Mark AE, Berendsen HJC. GROMACS: fast, flexible and free. J Comp Chem. 2005;26:1701–18.

[33] Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM: a program for macromolecular energy minimization and dynamic calculations. J Comput Chem. 1983;4:187–217.

[34] Freddolino PL, Gardner KH, Schulten K. Signaling mechanisms of LOV domains: new insights from molecular dynamics studies. Photochem Photobiol Sci. 2013;12(7):1158–70.
[35] Polverini E, Schackert FK, Losi A. Interplay among the "flipping" glutamine, a conserved phenylalanine, water and hydrogen bonds within a blue-light sensing LOV domain. Photochem Photobiol Sci. 2020;19(7):892–904.
[36] Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. J Mol Graph. 1996;14(1):33–8.
[37] Jurcik A, Bednar D, Byska J, Marques SM, Furmanova K, Daniel L, et al. CAVER analyst 2.0: analysis and visualization of channels and tunnels in protein structures and molecular dynamics trajectories. Bioinformatics. 2018;34(20):3586–8.
[38] Ganguly A, Thiel W, Crane BR. Glutamine amide flip elicits long distance allosteric responses in the LOV protein vivid. J Am Chem Soc. 2017;139(8):2972–80.
[39] Iuliano JN, Collado JT, Gil AA, Ravindran PT, Lukacs A, Shin S, et al. Unraveling the mechanism of a LOV domain optogenetic sensor: a glutamine lever induces unfolding of the Jα helix. ACS Chem Biol. 2020;15(10):2752–65.
[40] Dietler J, Gelfer R, Kaiser J, Borin V, Renzl C, Pilsl S, et al. Signal transduction in light-oxygen-voltage receptors lacking the active-site glutamine; 2022. Preprint https//www.researchsquare.com/article/rs-956213/v1.
[41] Möglich A, Moffat K. Structural basis for light-dependent signaling in the dimeric LOV domain of the photosensor YtvA. J Mol Biol. 2007;373(1):112–26.
[42] Heelis PF, Parsons BJ, Phillips GO, McKellar JF. A laser flash photolysis study of the nature of flavin mononucleotide triplet states and the reactions of the neutral form with amino acids. Photochem Photobiol. 1978;28(2):169–73.
[43] Heelis PF, Parsons BJ, Phillips GO, McKellar JF. A laser flash photolysis study of the nature of flavin mononucleotide triplet states and the reactions of the neutral form with amino acids. Photochem Photobiol. 1978 Aug 1;28(2):169–73.
[44] Lafaye C, Aumonier S, Torra J, Signor L, von Stetten D, Noirclerc-Savoye M, et al. Riboflavin-binding proteins for singlet oxygen production. Photochem Photobiol Sci. 2022. doi: 10.1007/s43630-021-00156–1