Title

Singlet oxygen-mediated photochemical cross-linking of an engineered fluorescent flavoprotein iLOV

Journal

Journal of Biological Chemistry, 300(11)

ISSN

0021-9258

Authors

Jones, Benjamin J

Greene, Brandon L

Publication Date

2024-11-01

DOI

10.1016/j.jbc.2024.107845

JBC RESEARCH ARTICLE

Singlet oxygen-mediated photochemical cross-linking of an engineered fluorescent flavoprotein iLOV

Received for publication, July 29, 2024, and in revised form, August 30, 2024 Published, Papers in Press, September 30, 2024,
https://doi.org/10.1016/j.jbc.2024.107845

Benjamin J. Jones[1] and **Brandon L. Greene**[1,2,*]

From the [1]*Department of Chemistry and Biochemistry the University of California Santa Barbara, Santa Barbara, California, USA;* [2]*Biomolecular Science and Engineering Program, The University of California Santa Barbara, Santa Barbara, California, USA*

Reviewed by members of the JBC Editorial Board. Edited by Joan B. Broderick

Genetically encoded photoactive proteins are integral tools in modern biochemical and molecular biological research. Within this tool box, truncated variants of the phototropin two light-oxygen-voltage flavoprotein have been developed to photochemically generate singlet oxygen ($^1O_2$) *in vitro* and *in vivo*, yet the effect of $^1O_2$ on these genetically encoded photosensitizers remains underexplored. In this study, we demonstrate that the "improved" light-oxygen-voltage flavoprotein is capable of photochemical $^1O_2$ generation. Once generated, $^1O_2$ induces protein oligomerization *via* covalent cross-linking. The molecular targets of protein oligomerization by cross-linking are not endogenous tryptophans or tyrosines, but rather primarily histidines. Substitution of surface-exposed histidines for serine or glycine residues effectively eliminates protein cross-linking. When used in biochemical applications, such protein–protein cross-links may interfere with native biological responses to $^1O_2$, which can be ameliorated by substitution of the surface exposed histidines of improved" light-oxygen-voltage or other $^1O_2$-generating flavoproteins.

The reactivity of molecular oxygen (O_2) presents a challenge to both aerobic life and life at the aerobic-anaerobic interface due to the production of reactive oxygen species (ROS) from O_2 (1–3). These ROS, such as singlet oxygen ($^1O_2$), superoxide ($\bullet O_2^-$), hydrogen peroxide (H_2O_2), and hydroxyl radicals ($\bullet$OH), are capable of oxidatively damaging biomolecules, which disrupt downstream cellular functions, leading to oxidative stress, tumorigenesis, and cell death (4–6). On the other hand, ROS also play a crucial role in signal transduction and posttranslational enzyme processing (3, 7, 8). The interplay between detrimental and beneficial contributions of ROS to biological function remains unclear, and thus, tools that selectively generate ROS under defined conditions, which do not interfere with the biological response under study, would be of value.

Among the ROS family, $^1O_2$ is the only species that is not generated by electron transfer, but rather by energy transfer in a triplet–triplet annihilation reaction between the ground state triplet $^3O_2$ and another triplet species, or the oxidative decomposition of peroxides (9). In the biological context, energy transfer partner triplets are typically relaxation products of chromophore electronic excited singlet states. The ability to generate $^1O_2$ photochemically has motivated the development of genetically encodable photoactive flavoproteins based on the *Arabidopsis thaliana* light-oxygen-voltage (LOV)-sensing domain of phototropin 2 (10, 11). These proteins bind a flavin mononucleotide (FMN) cofactor and photochemically generate $^1O_2$, but the quantum yield of $^1O_2$ generation and the distribution of other ROS products (*e.g.* $\bullet O_2^-$) remain unsettled (10, 12, 13).

By genetically tagging $^1O_2$-generating proteins to specific tissue-localized genes, LOV-based proteins can, in principle, be used to examine the effect of $^1O_2$ on biomolecules and biomolecular pathways, but the reactivity of $^1O_2$ with the LOV photosensitizer itself has been unexplored. The reactivity of $^1O_2$ toward common biomolecules is distinct from other ROS, where it acts as an electrophile and a Diels–Alder reactant, targeting nucleoside bases, unsaturated lipids, and amino acid residues histidine (H), cysteine (C), methionine (M), tyrosine (Y), and tryptophan (W) (14). All currently known $^1O_2$-generating flavoproteins contain at least three of these reactive amino acids, and thus could directly react with photogenerated $^1O_2$.

In this study, we investigate the reactivity of photogenerated $^1O_2$ with the "improved" LOV (iLOV) protein, a highly homologous flavoprotein to two previously reported $^1O_2$-generating proteins mini $^1O_2$ generator (miniSOG) and $^1O_2$-producing protein (15). We find that photogenerated $^1O_2$ reacts with the protein to form iLOV–iLOV cross-links. The reactivity is mediated by surface exposed histidines. This observation is consistent with previous studies, and we propose a mechanism of cross-linking involving a histidine attack of a 2-oxo-histidine product of $^1O_2$ and another histidine Nε from a second iLOV (16). These findings suggest that such protein-based $^1O_2$-generating photosensitizers may, themselves, engage in *in vivo* reactivity, which could affect subsequent $^1O_2$ responses *in vivo*. Alternatively, the $^1O_2$ mediated proximity labeling could be leveraged for investigating protein–protein interactions.

Results

Following purification of iLOV, we consistently observe two species by SDS-PAGE analysis, despite extensive effort to

* For correspondence: Brandon L. Greene, greene@chem.ucsb.edu.

optimize purification, small ubiquitin-like modifier protein (SUMO) cleavage, FMN reconstitution, and iLOV isolation (Fig. 1*A*). The most intense band in the SDS-PAGE of the purified iLOV product exhibits an apparent molecular weight of 13 kDa, similar to the expected molecular weight of iLOV of 13,232 Da (isotopic average mass), while a fainter band of variable intensity exhibits an apparent molecular weight of 26 kDa. The minor species is similar in electrophoretic mobility to the noncleaved $(H)_6$-SUMO-iLOV (25,360 Da) iLOV parent protein, but is also the expected molecular weight of a purified iLOV dimer that cannot be chemically separated by denaturation and reduction with 2-mercaptoethanol.

We turned to mass spectrometry to determine the nature of this minor protein species. Using MALDI-ToF mass spectrometry, we identified four peaks with apparent *m/z* values of 13,230, 26,460, 39,760, and 53,140 with progressively lower intensity (Fig. 1*B*). Based on the similarity of the 13,230 *m/z* MALDI peak to the expected mass of iLOV (0.02% mass difference) and the observed 26,460 *m/z* of the second peak, this higher molecular weight species is most consistent with an iLOV dimer (26,462 Da, assuming a bond is formed between the two proteins *via* formerly R−H groups, 0.01% mass difference), as opposed to a potentially noncleaved $(H)_6$-SUMO-iLOV (1% mass difference). Similarly, the 39,760 *m/z* peak appears consistent with an iLOV trimer with an expected molecular weight of 39,692 Da (0.2% mass difference) and the peak at 53,140 *m/z* is similar in mass to an iLOV tetramer at 52,922 Da (0.4% mass difference).

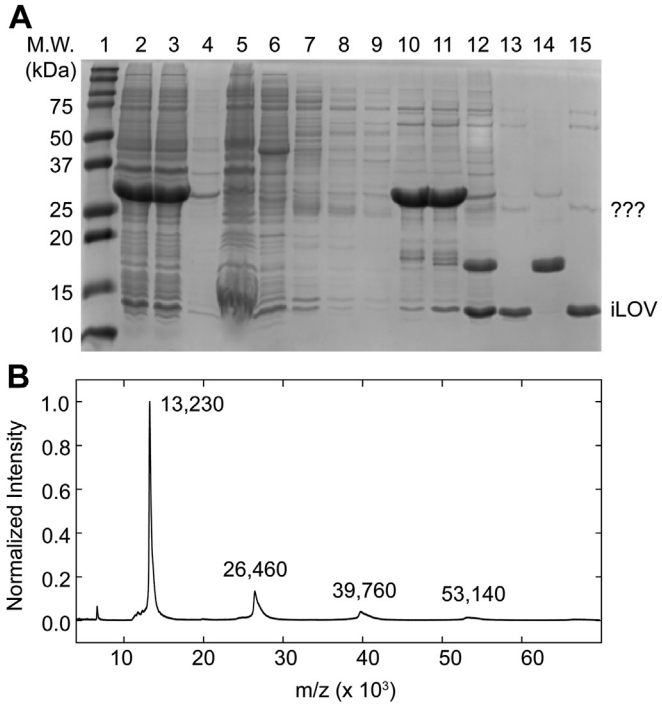

Figure 1. Recombinant expression, purification, and MALDI-ToF analysis of cross-linked iLOV. *A,* SDS-PAGE gel depicting the purification of iLOV. Lanes: one molecular weight marker, 2 *Escherichia coli* cell lysate, three soluble lysate, four cell debris, five immediate Ni-NTA flow-through, 6 to 9 Ni-NTA flow-through during the imidazole wash, 10 Ni-NTA eluted $(H)_6$-SUMO-iLOV, 11 desalted $(H)_6$-SUMO-iLOV, 12 ULP1-digestion product, 13 ULP1-digested Ni-NTA flow-through, 14 ULP1-digestion Ni-NTA elution, 15 desalted ULP1-digested Ni-NTA flow-through. *B,* MALDI-ToF mass spectrum of purified iLOV suggests oligomerization of WT iLOV. iLOV, improved light-oxygen-voltage; Ni-NTA, nickel nitrilotriacetic acid; ULP1, ubiquitin-like protease 1.

To complement the MALDI analysis, we also gel-extracted both the predicted iLOV monomer and potential dimer bands from the SDS-PAGE gel, performed tryptic digestions and subjected them to ultraperformance liquid chromatography tandem ToF/quadrupole mass spectrometry (UPLC-MS/MS) analysis (Fig. S1). In both the lower and higher apparent molecular weight gel bands several unique ions were observed by both their primary mass spectrum (MS^1) *m/z* and their ion fragments in the secondary mass spectrum (MS^2) to be derived from iLOV, with 86% of the iLOV sequence identified. A complete list of the identified peptides, measured and calculated masses, and fragment ions identified for both the monomer and dimer bands are provided in Tables S1 and S2, respectively. Some of the expected $(H)_6$-SUMO tryptic peptides were observed in the UPLC-MS/MS analysis of the digested monomer band at low intensities, but none were observed in the dimer band. Collectively, the mass spectrometry and SDS-PAGE analysis support the assignment of the higher molecular weight species in purified iLOV samples as homomeric cross-links of iLOV.

We suspected the origin of iLOV cross-linking was due to photochemical reactions mediated by the protein-bound FMN cofactor. To investigate this hypothesis, we performed time-dependent illuminations of iLOV and analyzed the products by SDS-PAGE (Fig. 2*A*). Following illumination of iLOV we observe the accumulation of the 26 kDa band with a concomitant diminished intensity of the monomer 13 kDa band. This effect can be seen after as little as 30 s of illumination, while after 2 h of illumination the original monomer had nearly disappeared. In nonilluminated samples, little change in the ratio of dimer to monomer band intensities occurred over the course of the experiment. Analysis of each gel extracted, trypsin digest of the illuminated dimer and nonilluminated monomer by UPLC-MS/MS revealed a loss of three peptide fragments, residues −3 to 3, 80 to 90, and 94 to 110 (Fig. 2*B*, Tables S3, and S4). These peptides contain amino acids with known reactivity toward $^1O_2$, including histidine, methionine, and tryptophan, but only histidine is contained in all of the peptides. We also searched for potential post-translational modifications that may be expected following a reaction with $^1O_2$. We identified a +32 Da modification of the C-terminal peptide composed of residues 94 to 110 (Fig. S2). Based on the fragmentation pattern of this peptide in the MS^2 spectrum, the +32 Da modification is C-terminal to L104 within the amino acid sequence -DGSDHV-COOH.

Photochemical protein modifications are commonly associated with O_2 and ROS. Given that all the prior investigations were conducted aerobically, we sought to test whether O_2 and ROS were involved in iLOV cross-linking. We repeated the photochemical cross-linking experiment anaerobically and analyzed the photoproducts by SDS-PAGE (Fig. 3*A*). Over the course of the illumination, we observed little change in the ratio of the iLOV monomer and dimer, in contrast to the same experiment performed aerobically in otherwise identical

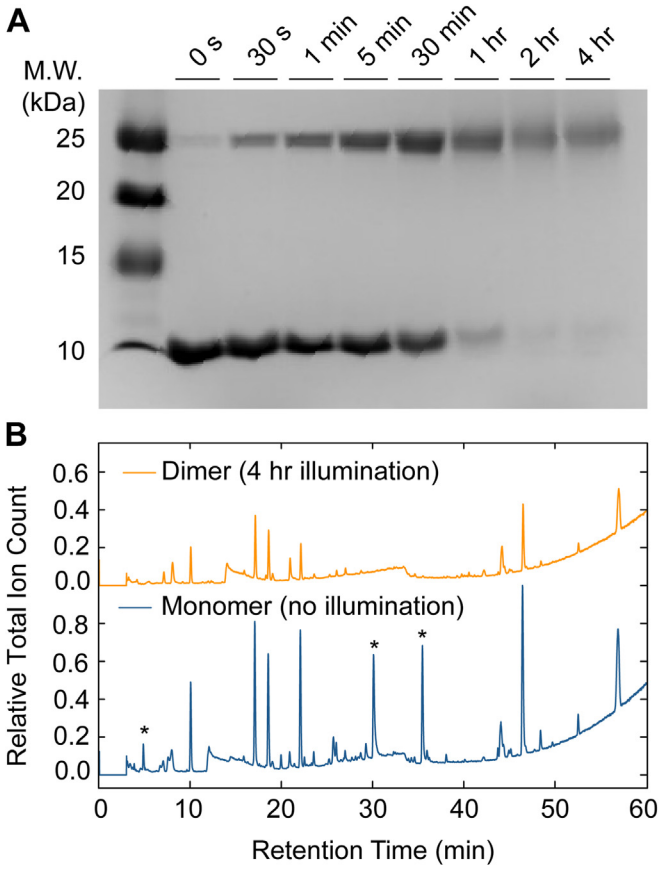

Figure 2. Photochemical dimerization of iLOV. *A,* SDS-PAGE analysis of illumination time-dependent dimerization of iLOV. *B,* UPLC-MS analysis of total ion counts of trypsin digested SDS-PAGE gel-excised monomer (*navy*) and dimer (*orange*) bands. Peptides lost upon illumination in the dimer band are indicated by *asterisks*. iLOV, improved light-oxygen-voltage.

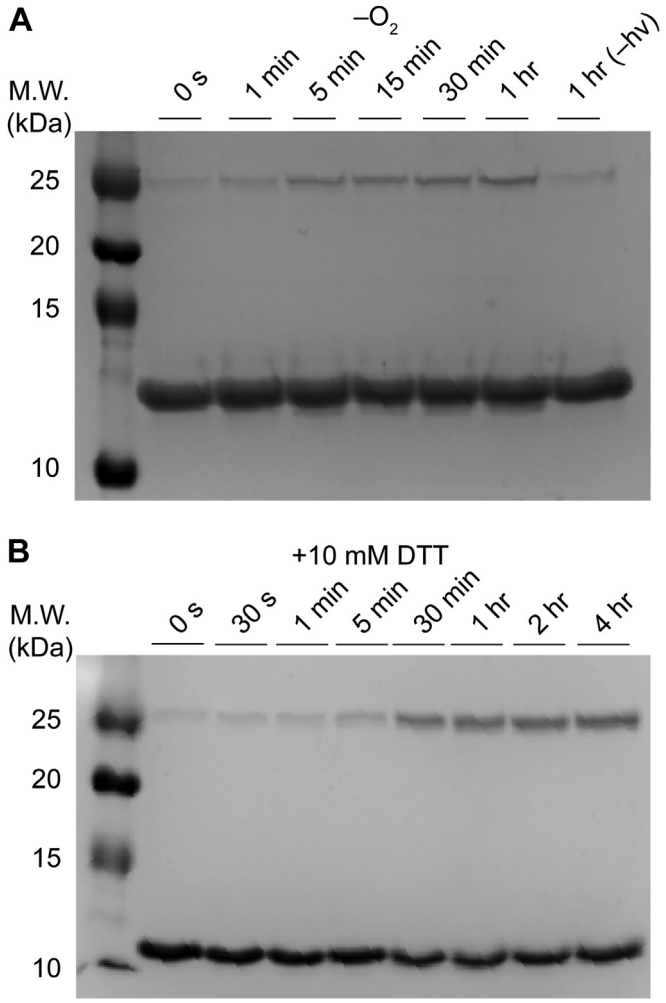

Figure 3. Cross-linking of iLOV in reducing or anaerobic conditions. *A,* illumination of anaerobic iLOV. *B,* illumination of an aerobic samples of iLOV in the presence of 10 mM DTT. iLOV, improved light-oxygen-voltage.

conditions. We then repeated the photochemical cross-linking experiments aerobically, supplemented with 10 mM of the sacrificial reductant DTT, which does not readily react with $^3O_2$, but does effectively scavenge the ROS $^1O_2$ and other oxidized biomolecules (14) (Fig. 3*B*). Again, we observed a substantial suppression of photochemical dimerization, compared to control experiments where DTT was omitted. The cumulative suppression of cross-linking in either O_2-free or DTT supplemented conditions supports a role of O_2-mediated generation of ROS that cross-link iLOV.

Given the known reactivity of LOV-based flavoproteins toward O_2, with both $^1O_2$ and $\bullet O_2^-$ being reported as photochemical products of reactions with $^3O_2$, we sought to identify the ROS responsible for iLOV cross-linking (11, 13, 17). The singlet oxygen sensor green (SOSG) fluorophore is a $^1O_2$ sensor which generates an increase in fluorescence following reaction with $^1O_2$ (18) and is nonreactive with other ROS such as H_2O_2 and $\bullet O_2^-$ (19). Illumination of iLOV in the presence of SOSG results in an increase in fluorescence of SOSG (Fig. 4). In the absence of $^3O_2$, no production of $^1O_2$ is observed by iLOV. The generation of $^1O_2$ was also dependent on iLOV; no fluorescent SOSG was observed in the absence of iLOV in the presence or absence of $^3O_2$. These results suggest that $^1O_2$ is a photoproduct of iLOV in the presence of $^3O_2$.

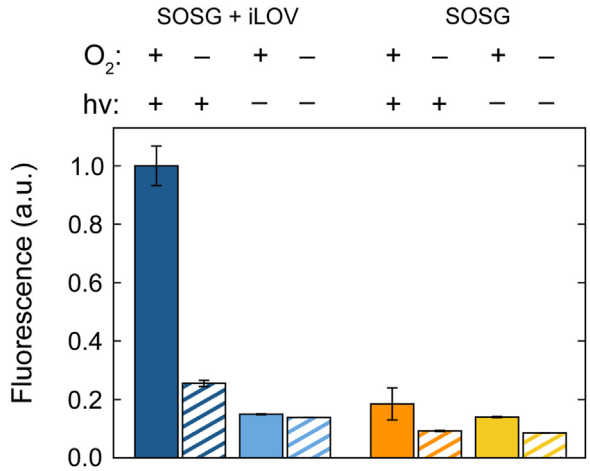

Figure 4. iLOV-mediated photochemical $^1O_2$ generation, observed through SOSG fluorescence in the presence (+) and absence (–) of O_2 and illumination (hν). The fluorescence response in all conditions is normalized to the SOSG signal of iLOV illumination aerobically (*navy*). Anaerobic SOSG responses shown in stripped bars. Photochemical response of SOSG alone is shown in *orange* and *yellow*. iLOV, improved light-oxygen-voltage; SOSG, singlet oxygen sensor *green*.

Regardless of the identity of the ROS, the molecular target of ROS cross-linking reactivity in iLOV must be identified if this chemistry is to be controlled. Several amino acids are susceptible to reactivity with diverse ROS, primarily W, Y, C, M, and H. The iLOV protein has four redox-active aromatic amino acid residues, one tryptophan and three tyrosines residues. In model studies, both tryptophan and tyrosine have been shown to covalently cross-link following exposure to ROS (14, 20). The sole tryptophan residue of iLOV has been proposed to be an electron donor in photochemical electron transfer, generating a neutral W• and a neutral FMN semiquinone. This photochemical generation of an neutral FMN semiquinone in iLOV has only been demonstrated in an iLOV mutant, Q103D (17), but the possibility of this reactivity in authentic iLOV could function as a route to FMN reduction and $•O_2^-$ production. Alternatively, tyrosine and tryptophan also react with $^1O_2$ to form reactive intermediates that may chemically cross-link (14).

To examine the role of tyrosine and tryptophan amino acids in protein cross-linking, we mutated each of these residues for the redox-inactive amino acid phenylalanine (F), denoted hereafter as iLOV$^{\Delta YW}$. Initially, we examined the excited state dynamics of the iLOV FMN *via* fluorescence lifetime measurements of the reconstituted protein. The FMN singlet excited state (^{1}FMN*) lifetime was unchanged by the substitution of any one redox-active amino acid or the concerted substitution of all four of them, with iLOV ^{1}FMN* fluorescence decay lifetimes of 4.43 (0.01) ns for iLOV and 4.42 (0.01) ns in iLOV$^{\Delta YW}$ (Fig. S3, *A* and *B* and Table S5). The lack of difference in the ^{1}FMN* lifetime between iLOV and iLOV$^{\Delta YW}$ does not rule out a role of the ^{3}FMN* state in photochemical cross-linking *via* tyrosine or tryptophan photooxidation. To address the role of ^{3}FMN*-mediated aromatic amino acid oxidation and iLOV cross-linking, we measured the cross-linking propensity of iLOV$^{\Delta YW}$. The iLOV$^{\Delta YW}$ protein exhibited similar or increased levels of photochemical cross-linking, determined by SDS-PAGE, as the parent iLOV (Fig. S3*C*). Thus, another amino acid residue enables photochemical cross-linking of iLOV.

To further test the potential role of $•O_2^-$ in photochemical cross-linking, we compared the $^1O_2$ generation efficiency of iLOV and iLOV$^{\Delta YW}$, as the two ROS are generated from the same triplet flavin photoproduct. Using SOSG as a fluorescent reporter for $^1O_2$ (Fig. 5), we observed that the iLOV$^{\Delta YW}$ exhibited a larger SOSG response than the WT protein (p value of 0.00015), suggesting $•O_2^-$ formation is suppressed in the iLOV$^{\Delta YW}$ mutant, resulting in a greater quantum yield of $^1O_2$ production from the flavin triplet. Illumination of samples of iLOV supplemented by large quantities of superoxide dismutase (SOD) did not affect the propensity for cross-linking and even resulted in a distribution of cross-linked products between iLOV and SOD as evidenced by SDS-PAGE (Fig. S4*A*). The reaction of xanthine oxidase (XO) with $^3O_2$ and xanthine is known to produce $•O_2^-$ and H_2O_2, but not $^1O_2$, in a light-independent reaction, allowing the interrogation of $•O_2^-$-induced protein cross-linking (21, 22). Incubation of iLOV with XO and xanthine in the dark did not result in detectable cross-linking (Fig. S4*B*), again supporting the unique role of $^1O_2$.

The evidence presented herein is consistent with $^1O_2$-mediated photochemical cross-linking of iLOV *via* an amino acid other that tyrosine or tryptophan. The known reactivity of $^1O_2$ with histidine, and the potential of these reactive products toward covalent cross-linking is well documented (23–25). There are three histidine residues in our iLOV protein (Fig. 5, *A* and *B*), two associated with the source phototropin 2 LOV protein, and one introduced during engineering. We mutated each H85 and H109 to serine (S) and H-2 to glycine (G), individually and collectively (iLOV$^{\Delta H}$), to determine their role in photochemical cross-linking. None of the single histidine

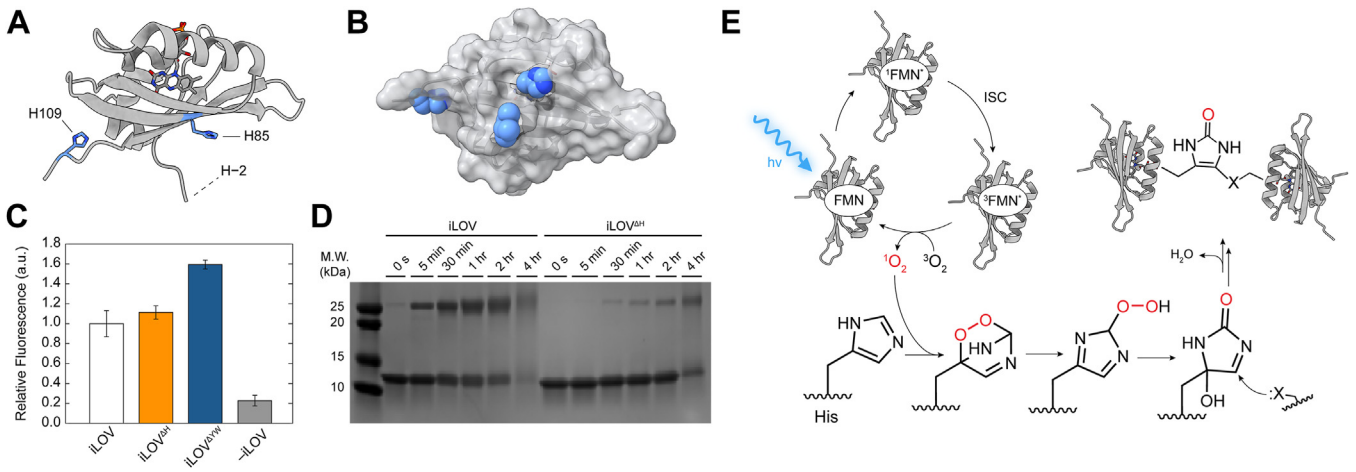

Figure 5. The role of histidines in iLOV cross-linking. *A,* an X-ray crystal structure of iLOV (PDB: 4EES) depicting the locations of each histidine residue. H – 2 is not present in the crystal structure as it was introduced only in this work. *B,* the same crystal structure depicting the van der Waals protein surface. The native histidine residues and the N-terminal I1 (as a proxy for the N-terminal H – 2) are depicted in *blue spheres*. *C,* SOSG response following 30 min illumination of iLOV (*white*), iLOV$^{\Delta H}$ (*orange*, *p*-value of 0.21), iLOV$^{\Delta YW}$ (*navy*, *p* value of 0.00015), and a –iLOV control (*gray*). *D,* effect of histidine residues on photochemical cross-linking measured by illumination time-dependent SDS-PAGE. *E,* the proposed mechanism of $^1O_2$/histidine-mediated protein cross-linking. iLOV, improved light-oxygen-voltage; SOSG, singlet oxygen sensor *green*.

iLOV mutants exhibited significant attenuation of photochemical cross-linking by SDS-PAGE analysis (Fig. S5). On the other hand, significantly less cross-linking was observed for the iLOV$^{\Delta H}$ with statistical change in the ^{1}O2 generation efficiency relative to the WT protein (Fig. 5, *C* and *D*). To examine potential posttranslational modifications in the iLOV$^{\Delta H}$ protein following illumination, we performed tryptic digest of gel-extracted iLOV$^{\Delta H}$ preillumination and postillumination. After 30 min of illumination, two new peptides were identified with MS^1 and MS^2 masses consistent with the oxidation of W81 (+32 Da and +4 Da) with similar, but distinct retention times to the parent nonoxidized W peptide (Figs. S6, S7, and Table S6). Neither of these W oxidation photoproducts were observed in the parent iLOV after illumination, suggesting this is a minor modification when histidines are present. We also observed a loss of a mass associated with the N-terminal −3-3 peptide, despite the H−2G mutation. This peptide also contains two methionines, which we suspect are potential sites of posttranslational modification in the absence of histidines.

Understanding the nature of the photochemical ROS product(s) as well as the mechanism of generation is essential for rational engineering of iLOV photochemical properties. If $^1O_2$ generation by iLOV mediates protein cross-linking *via* surface exposed histidine residues, limiting $^1O_2$ generation *via* triplet–triplet energy transfer should suppress cross-linking activity. Energy transfer is a donor-acceptor distance-dependent process, and thus, the proximity of ^{3}FMN* and $^3O_2$ may dictate $^1O_2$ generation efficiency *via* a "type II" process (13, 26). To address the potential for O_2 to approach the FMN cofactor we performed five independent molecular dynamics (MD) trajectory simulations of iLOV in the presence of O_2 over a collective 1.4 μs. Each simulation exhibited average protein backbone atom RMSD < 2 Å, with only transient excursions >2 Å (Fig. S8). In each of five independent MD trajectory simulations O_2 penetrated into the protein structure, occupying at least one of three sites, denoted "X," "Y," and "Z," within 10 Å of the FMN cofactor (Fig. 6*A*).

The O_2-binding site X, identified in previous MD simulations, is lined with the alanine (A) residue A40, which is a glycine (G) in the $^1O_2$-generating LOV protein "miniSOG" (10, 27, 28). To determine whether steric access to this site was impactful on $^1O_2$ generation, we mutated the A40 residue in iLOV to either phenylalanine, isoleucine (I), valine (V), and glycine, and quantitated their respective $^1O_2$ generation efficiency by reaction with SOSG *via* illumination-dependent fluorescence response. While A40F exhibited weak binding of the FMN cofactor after reconstitution, A40I, A40V, and A40G all exhibited varying degrees of diminished $^1O_2$ yields after illumination, as determined by the SOSG fluorescent $^1O_2$ reporter (Fig. 6*B*).

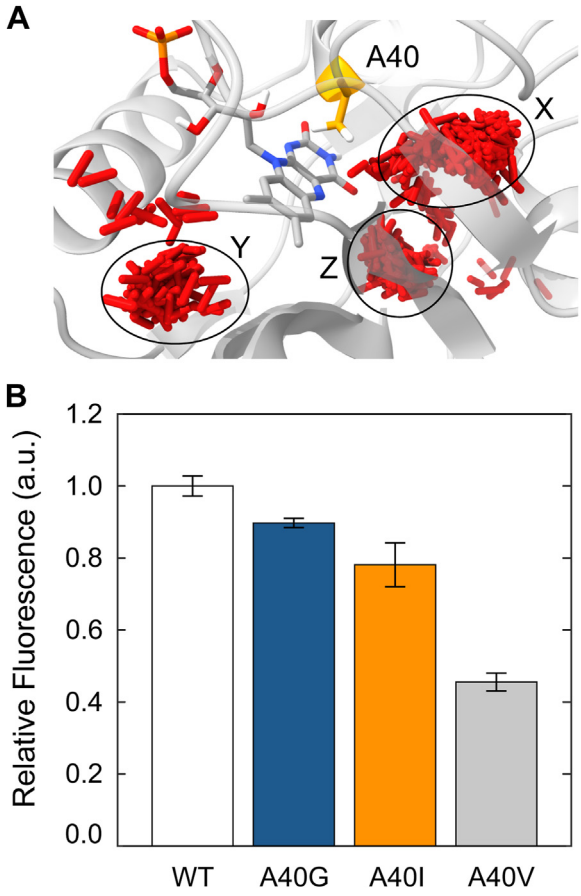

Figure 6. Predicted O_2 diffusion into iLOV and role of A40 in access to FMN cofactor. *A,* molecular dynamics simulations predict three sites occupied by O_2 within iLOV labeled "X," "Y," and "Z" near A40. *B,* $^1O_2$ generation by iLOV and various A40 mutants determined by SOSG fluorescence after illumination. *p* values, calculated using a *t* test against iLOV^{WT} were 0.0043, 0.000095, and 0.0048 for A40G, A40V, and A40I mutants, respectively. FMN, flavin mononucleotide; iLOV, improved light-oxygen-voltage; SOSG, singlet oxygen sensor *green*.

Discussion

The redox environment of cells in aerobic conditions is dictated, in part, through ROS. The effects of ROS are complex, involving both the control of essential cellular functions as well as detrimental damage to biomolecular species such as proteins, nucleic acids, and lipids (14, 29). Genetically encodable flavoproteins based on the LOV domain have been employed to spatio temporally perturb the redox environment *via* photochemical generation of $^1O_2$ or other ROS in cellular environments (10, 30–34). In these studies, the LOV proteins are assumed to be noninteracting with photo-generated $^1O_2$ or other ROS sources, such that downstream cellular responses can be investigated with fidelity to the ROS, yet the reactivity of LOV-based proteins with $^1O_2$ has not been thoroughly investigated.

In this study, we find that iLOV, a member of the LOV proteins that photochemically generate $^1O_2$, reacts directly with $^1O_2$ to form homo-oligomeric protein covalent cross-links, evidenced by SDS-PAGE and mass spectrometry. The $^1O_2$-mediated protein cross-linking is light- and $^3O_2$-dependent, and can be mitigated by thiol-based $^1O_2$ scavengers. Oligomerization proceeds through sequential steps of addition (*i.e.,* monomer to dimer to trimer, *etc.*), shown by the MALDI intensity profile and SDS-PAGE analysis, but whether a previously oligomerized iLOV can serve as a photochemical $^1O_2$

generator remains to be determined. Using site-directed mutagenesis (SDM) targeting the redox-active aromatic residues tryptophan and tyrosine in iLOV, we see an enhanced capacity for $^1O_2$ generation and photochemical cross-linking. We ascribe this phenomenon to the suppression of ^{3}FMN*-W/Y electron transfer, eliminating a competing pathway for ^{3}FMN* quenching, and therefore increasing the yield of $^1O_2$. This implies that $\bullet O_2^-$ is a photoproduct of iLOV, in addition to $^1O_2$. Even concentrations of SOD as high as 100 μM did not suppress photochemical cross-linking, supporting the assertion that $\bullet O_2^-$ is not responsible for the observed cross-linking, but we acknowledge that $^1O_2$ is a known inhibitor of SOD which could complicate the interpretation of these results (35, 36). The role of $\bullet O_2^-$ is more directly demonstrated in the absence of observable cross-linking in the presence of xanthine and XO, a system known to produce $\bullet O_2^-$ but not $^1O_2$. Together with our estimation of $^1O_2$ by SOSG for iLOV and iLOV$^{\Delta YW}$, these results suggest that $\bullet O_2^-$ generation is a competing pathway for photochemical reactivity with $^3O_2$, but $\bullet O_2^-$ generation does not result in protein cross-linking. Conversely, mutation of surface exposed histidines to serine or glycine resulted in the ablation of photochemical cross-linking, while $^1O_2$ generation was maintained (Fig. S4*A*). Whether as a product of photooxidation or ground-state reactivity with $^1O_2$, W81 appears to be a secondary degradation target of iLOV, with +32 Da and +4 Da modifications which we tentatively assign to dioxyindolalanine or N-formylkynurenine and kynurenine, respectively. A role for H_2O_2 must also be considered. While H_2O_2 is not a known photoproduct of either miniSOG or iLOV, the $\bullet O_2^-$ produced by iLOV would be disproportionated to O_2 and H_2O_2 by SOD. We see no evidence for increased cross-linking of iLOV in the presence of SOD, again suggesting $^1O_2$ initiates iLOV cross-linking.

Based on the evidence presented herein and literature precedent, we propose a mechanism for photochemical cross-linking described in Figure 5 (16, 37). The protein structure and dynamics facilitate $^3O_2$ penetration into iLOV's core where it resides for extended periods of time in one of three pockets, identified by MD simulations. These sites poise $^3O_2$ for triplet–triplet annihilation with the ^{3}FMN* photoproduct of FMN excitation, following ^{1}FMN* intersystem crossing. While distinct O_2 binding sites are supported by our MD simulations, a series of mutant proteins designed to test the role of steric access to one of these binding sites revealed no correlation between steric bulk and $^1O_2$ yield (Fig. 6), suggesting the exact nature of the encounter complex, and the potentially pleotropic role of these mutations remain unresolved, particularly in relation to flavin binding, as evidenced by the significantly reduced FMN binding in the A40F mutant (Table S9). Once generated, $^1O_2$ diffuses out of the protein and reacts with surface-exposed histidine residues in iLOV, resulting in an endoperoxide Diels–Alder adduct intermediate. This assertion is buttressed by the MS^2 identification of a +32 Da modification of the C-terminal peptide fragment -DGSDHV-COOH of illuminated dimeric iLOV. This unstable species quickly rearranges to form a hydroperoxide intermediate, and later a reactive urea Schiff base species. This species is then vulnerable to nucleophilic attack by several possible nucleophilic residues on the protein surface. Based on the loss of histidine-containing peptides in the tryptic digest mass spectrometry analysis of dimerized proteins, we propose histidine–histidine cross-links as the most common. The similarities between all phototropin 2–based LOV proteins suggests that $^1O_2$-mediated cross-linking is a common phenomenon.

Our results demonstrate that, at least in the *in vitro* conditions examined in this study, iLOV photochemically cross-links to form oligomers *via* histidine residues. The implications of iLOV photochemical cross-linking on *in vivo* studies involving the use of genetically encoded $^1O_2$ generators remain to be determined. Several differences between the conditions investigated herein and those that may be expected *in vivo* will likely diminish iLOV cross-linking including decreased $^1O_2$ quantum yield due to light scattering and inner filter effects of competing chromophores, competing reactions such as $^1O_2$ reduction by cellular small molecules, enzymes or nucleic acids, or limited O_2 availability in microaerobic or anaerobic conditions. Nevertheless, given the success of prior studies in generating $^1O_2$ *in vivo*, the potential for iLOV heter- or homo-oligomeric cross-linking deserves consideration (10, 14). Given the potential for spatio and temporally resolved $^1O_2$ generation by this genetically encoded flavoprotein, $^1O_2$ generation and cross-linking properties of iLOV could be utilized for *in vivo* affinity labeling and pull down of proteins reactive to ROS in a spatio-temporal manner *via* iLOV-binding antibodies or other similar technologies. For example, genetically fusing iLOV to one of two proteins suspected of forming a protein–protein complex would place the site of $^1O_2$ generation in close proximity to the protein–protein complex of interest. Upon illumination, iLOV-generated $^1O_2$ could attack histidines on the surface of one of the target proteins, forming a reactive electrophile, and subsequently generating cross-links with the proximal protein–protein complex partner. The iLOV fluorescence or affinity to antibodies could then be used to identify and enrich the protein–protein complex for identification. Alternatively, $^1O_2$-mediated protein cross-linking, particularly between iLOV and other proteins involved in the ROS response pathway, could complicate interpretation of experimental results. This can be controlled for by ablation of the iLOV histidines, which we show successfully inhibits iLOV cross-linking, limiting $^1O_2$ chemistry to cross-linking other proteins in the biological milieu or alternative pathways of $^1O_2$ quenching detailed above.

Experimental procedures

Materials

The iLOV gene was obtained as a gracious gift from the Mukherjee lab, inside a pQE80L vector backbone (38). The 6× polyhistidine-tagged ubiquitin-like protease 1 (ULP1) was expressed and purified as previously described (39, 40, 41). All DNA oligomers used as primers for PCR were ordered from Integrated DNA Technologies. DpnI, Phusion HF polymerase,

chemically competent DH5-α and BL21 (DE3) *Escherichia coli,* the Qiagen QIAprep Spin Miniprep and Gel Extraction Kits, and the NEBuilder HiFi DNA Assembly kit were purchased from New England Biolabs (NEB). IPTG was purchased from Merck. Sephadex G-25 fine resin, LB Miller broth, NaH_2PO_4, iodoacetamide, TFA, DTT, formic acid, urea, PMSF, bovine SOD, bovine XO, xanthine, and Amicon Ultra centrifugal filter units were purchased from Millipore Sigma. FMN sodium salt, HisPur nickel nitrilotriacetic acid (Ni-NTA) resin, Pierce 660 nm protein assay kit, SOSG, and Imperial protein stain were obtained from Thermo Fisher Scientific. Bradford reagent and N,N,N′,N′- tetramethylethylenediamine were purchased from VWR. Mini-PROTEAN TGX-Precast protein gels containing a polyacrylamide gradient from 4 to 20%, 2x Laemmli sample buffer, and Precision Plus Protein Standards were purchased from Bio-Rad. The single-color, cold, visible, mounted, light-emitting diode (LED) with a wavelength of 450 nm and a bandwidth of 18 nm was obtained from ThorLabs. The MALDI steel target plates, Sinapinic acid matrix, and MALDI MS calibration standards (Protein Standards I and II) were obtained from Bruker. Bisacrylamide and kanamycin sulfate were obtained from Genesee Scientific. Millipure water was obtained by flowing deionized water through a 4-module Barnstead E-Pure ultrapure water system until a resistivity of >17 MΩ/cm was reached.

Construction of pSUMO-iLOV plasmid vector

The iLOV gene was cloned into the pET backbone of pHYRSF53 (AddGene #64696) *via* Gibson assembly. The resulting inducible expression gene codes for iLOV N terminally tagged with a 6 × polyhistidine tag and the SUMO with flexible serine and glycine residues in between fragments, denoted $(H)_6$-SUMO-iLOV. Fragments for this Gibson assembly were generated by PCR using primers with sequences shown in the Table S7. The fragments were gel-purified using a Qiagen Gel Extraction kit and assembled using an NEBuilder HiFi DNA Assembly kit. The assembled product was used to transform chemically competent DH5-α *E. coli* which were then plated onto agar plates containing LB (Miller, pH 7) and 50 μg/ml Kanamycin as the selection antibiotic. After growth overnight, single colonies were picked and used to inoculate a small culture in 5 ml LB media with 50 μg/ml Kanamycin in 15 ml sterile culture tubes. This culture was allowed to grow for 18 to 20 h at 37 °C shaking at 220 rpm. The plasmid was purified from these cells using a Qiagen QIAprep Spin Miniprep Kit. The purified plasmids were sent to the UC Berkeley DNA Sequencing Facility to undergo full plasmid nanopore sequencing. After validation of the correct sequences, the plasmids were used as templates for subsequent rounds of mutagenesis described in the following section.

SDM of iLOV mutants

SDM was carried out using PCR primers containing nucleotide mismatches consisting of a codon for the appropriate amino acid replacement (Table S8). The pSUMO-iLOV assembled in the previous section was used as a plasmid template for PCR using Phusion HF polymerase. The resulting PCR product was digested with DpnI to remove methylated template DNA and transformed into DH5-α cells. Plasmid DNA was purified and sequence verified as previously described. After validation of the correct mutation(s), the product plasmids were used as templates for subsequent rounds of mutagenesis for double, triple, and quadruple mutants.

Growth and purification of iLOV

The pSUMO-iLOV plasmid was used to transform BL21 (DE3) *E. coli* which were then plated on LB agar plates with Kanamycin and grown overnight. A single colony was picked and grown in 5 ml LB with 50 μg/ml Kanamycin overnight. From this overnight seed culture, 750 μl were added to each of two sterile 250 ml Erlenmeyer flasks containing 75 ml LB with Kanamycin. These flasks were shaken at 200 rpm at 37 °C until the cells reached an A_{600} of 0.6. At this point small aliquots of cells were collected, pelleted, flash frozen, and stored in a −80 °C freezer for future growths. Simultaneously, to one of the remaining 75-ml cultures, a final concentration of 20 μM of IPTG was added to induce protein expression. The two flasks were left to grow for another 5 h before samples of each were removed for SDS-PAGE to confirm expression of protein at the appropriate molecular weight.

After confirmation of protein expression, large-scale cultures were prepared. About 15 ml of cultured cells from the uninduced flasks were used to inoculate 1.5 L of LB, supplemented with 50 μg/ml Kanamycin. These cultures were allowed to grow and expression was induced as previously described. After growth, the cells were pelleted at 3000*g* for 10 min at 4 °C. The yellow pellet was then flash-frozen in liquid nitrogen and stored at −80 °C for purification at a later date.

To purify $(H)_6$-SUMO-iLOV, the cell pellet was placed on ice and allowed to thaw slowly. This pellet was resuspended in 200 mM NaCl, 50 mM $H_2PO_4^-$/HPO_4^{2-}, and 1.5 mM DTT at pH 7.50 (buffer A) supplemented with 1 mM PMSF to a volume of 5 ml per gram of cell paste and homogenized with a glass homogenizer. The cells were then lysed by passing the mixture 5 to 6 times through a French press homogenizer operating at 14,500 psi. The resulting lysate was then placed into 50 ml centrifuge tubes and pelleted at 14,000*g* for 10 min at 4 °C to remove insoluble debris. The supernatant was placed in a beaker with a stir bar on ice, while a 1:5 (v/v) ratio of 6% (w/v) streptomycin sulfate was added dropwise. The solution was allowed to stir on ice for 10 min, and then the insoluble material was pelleted at 29,000*g* for 30 min. The resulting supernatant was combined in a 1:1 ratio with buffer A supplemented with 15 mM imidazole and applied to a HisPur Ni-NTA resin, equilibrated with the same buffer. The column was then washed with 30 column volumes of buffer A supplemented with 50 mM imidazole. The protein was eluted with buffer A supplemented with 250 mM imidazole. Fractions were pooled based on either yellow color or positive reactions with Bradford reagent. The eluted $(H)_6$-SUMO-iLOV was then

dialyzed in the dark overnight at 4 °C against 1000 to 2000 × (v/v) buffer A. The pure $(H)_6$-SUMO-iLOV was either used immediately or supplemented with 10% (v/v) glycerol, flash-frozen in liquid nitrogen, and stored at −80 °C.

ULP1 digestion and loading of FMN

The recombinantly expressed and purified $(H)_6$-SUMO-iLOV was first digested with ULP1 to cleave the N-terminal $(H)_6$-SUMO fragment then reconstituted with FMN before separating the $(H)_6$-SUMO fragment from iLOV *via* Ni-NTA chromatography. The resulting cleaved protein contains the additional (−3)S-H-M-M(0) N-terminal amino acids, relative to the original iLOV (15). In all steps, the protein quantitation of iLOV variants were measured *via* the bicinchoninic acid–based Pierce 660 assay kit (42). The FMN content was determined by the absorbance at 448 nm with an estimated extinction coefficient of 14,800 M^{-1} cm^{-1} (43) for the protein-bound cofactor.

To cleave the $(H)_6$-SUMO fragment, 30 to 40 mg of $(H)_6$-SUMO-iLOV, or variants, were mixed in a 1:500 M ratio of ULP1 to $(His)_6$-SUMO-iLOV and incubated at room temperature for <2 h. To reconstitute iLOV, free FMN was then added to this mixture at 50 × molar excess of iLOV from a saturated solution of FMN in 200 mM NaCl, 50 mM $H_2PO_4^-$/HPO_4^{2-} at pH 8.0. The actual concentration of FMN was determined by measuring the absorbance of FMN at 450 nm using a molar extinction coefficient of 12,500 M^{-1} cm^{-1} for the free cofactor (44). The mixture was then heated to 60 °C for 15 min, 5 °C below the unfolding temperature of the apo protein, and then cooled back to room temperature.

The reconstituted iLOV in the FMN, $(H)_6$-SUMO, and ULP1 mixture was then passed through a second Ni-NTA column equilibrated with buffer A supplemented with 25 mM imidazole to remove the cleaved $(H)_6$-SUMO fragment, and the flow through was collected. The resin was further washed with buffer A supplemented with 25 mM imidazole at pH 7.50 until the characteristic yellow color of iLOV was no longer visible on the column. The eluted iLOV and FMN solution was then concentrated in a 10 kDa molecular weight cut-off Amicon Ultra centrifugal filter. This concentrated solution was buffer-exchanged using a Sephadex G-25 desalting column into 200 mM NaCl, 50 mM $H_2PO_4^-$/HPO_4^{2-} at pH 7.50 (buffer B), which clearly resolved free FMN and reconstituted iLOV. Initial and postreconstitution FMN/protein content is provided in Table S9, and the corresponding FMN UV-visible spectra are reported in Fig. S9. Pure, reconstituted iLOV was then flash-frozen in liquid nitrogen, and stored at −80 °C for later use. The final yield of iLOV after purification, SUMO cleavage, reconstitution, and buffer exchange, routinely ranged from 10 to 20 mg iLOV/g wet cell paste.

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis

For visualization of protein samples by SDS-PAGE, purified iLOV samples were diluted to 33 μM (0.43 mg/ml) into phosphate buffer, while impure protein samples were diluted to about 0.7 mg/ml of total protein determined by the aforementioned Pierce 660 assay. Equal volumes of diluted protein and 2 × Laemmli sample buffer containing 5% v/v 2-mercaptoethanol were mixed and heated to 98 ℃ for 5 to 10 min. Denatured samples were then either loaded into homemade 16% bisacrylamide polyacrylamide gels containing a 9.3% stacking layer or into premanufactured gels containing a polyacrylamide gradient from 4 to 20%. Additionally pre-stained protein standards were added alongside samples to estimate molecular weights. A voltage of 90 V was applied for 15 min before switching to 180 V for approximately 45 min or until the prestained ladders migrated across the gel. The gels were then washed in water for 5 min before being bathed in Imperial protein stain for several hours or overnight. The gels were then destained in water for several hours or overnight before either imaging or gel extraction.

Photochemical cross-linking of iLOV mutants

To initiate photochemistry, we used a 450 nm LED, mounted onto a home-made illumination apparatus delivering 3 W of power. The apparatus positions the light 1.6 cm above the center of a ring of six sample holders made to fit 1.5 ml plastic microcentrifuge tubes (*ca*. 1 mW/mm^2 irradiance). A solution of 100 μl of purified and reconstituted iLOV was prepared to a final concentration of 100 μM and placed into open opaque amber 1.5 ml microcentrifuge tubes. The tubes were then illuminated for various times. At specified time points, 10 μl was removed and combined with 20 μl water and 30 μl 2 × Laemmli buffer with 5% v/v 2-mercaptoethanol and flash-frozen in liquid nitrogen until all time points were collected. Aliquots were also collected for UV-visible absorption analysis at 4 h of illumination, diluted 10-fold into buffer B and measured on a Cary 60 Agilent UV-visible absorption spectrometer. The samples were then thawed, heated to 98 °C for 5 min, then resolved by SDS-PAGE. For experiments performed in the presence of reductant, and DTT was added to concentrations of 10 mM. For experiments performed with SOD, lyophilized SOD was dissolved in water and added to the illumination mixture in various concentrations.

For experiments performed in the dark with XO, xanthine was dissolved to a concentration of 10 mM in 0.1 M NaOH, while lyophilized XO was dissolved in water. Xanthine and XO were then added to the mixture to final concentrations of 400 μM and 20 U/l, respectively. Samples with XO were kept in the dark, and no LED illumination was performed.

For anaerobic illumination, protein samples were degassed on a Schlenk line prior to illumination. Next, the sealed, anaerobic samples would be removed from the Schlenk line to be placed into a Vacuum Atmospheres glovebox (<1 ppm O_2). Illumination of the degassed samples was performed inside the glovebox *via* the apparatus used aerobically described above with identical settings.

Matrix-assisted laser desorption ionization-time-of-flight mass spectrometry

Samples of iLOV that were exposed to ambient light over a 2-day purification process in the absence of a reductant were

dialyzed overnight against 2000 to 3000 × (v/v) of 5 mM ammonium bicarbonate to remove salts. The dialysis was repeated in fresh buffer for a few more hours. The samples were then diluted to 10 to 20 μM of protein in 5 mM ammonium bicarbonate. A saturated solution of sinapinic acid matrix in a 30:70 (v/v) mixture of acetonitrile and 0.1% TFA in water served as the matrix. Equal volumes of protein and matrix were mixed and 1 μl was applied to a MALDI steel target plate and allowed to dry at room temperature. MALDI-ToF mass spectra were collected on a Bruker Microflex LRF MALDI ToF (ONR N00014-12-1-0792) in linear mode configuration at *ca*. 70% laser power. For mass calibration, protein standards I and II, mixed together, were prepared in the same way as iLOV samples. Peaks were picked automatically by the Bruker PolyTools software to generate the polynomial calibration and verified manually.

Detection of $^1O_2$

SOSG was used to quantify $^1O_2$ generation. Standard solutions of 100 μg SOSG in 33 μl of methanol were prepared immediately before use. Samples of 50 μl of 116 μM iLOV were mixed with 1.5 μl of the SOSG stock solution in open opaque amber colored plastic microcentrifuge tubes and illuminated as described previously for 1 h or nonilluminated in control experiments. For negative controls, SOSG in buffer B was used in the absence of iLOV. After illumination, 1450 μl of buffer B was added to each sample to dilute the fluorescent SOSG and minimize inner filter effects, and the sample was loaded into a fluorescence quartz cuvette. The emission spectra of SOSG between 520 and 700 nm was collected in a Duetta Spectrometer (Horiba Scientific). The excitation wavelength was 505 nm with a bandpass of 3 nm. Samples were temperature-controlled at 25 °C during data acquisition. For the construction of bar charts, triplicate samples were illuminated and prepared. Triplicate emission spectra were integrated in MATLAB and normalized to the intensity of the triplicate average of iLOV WT samples. Error bars were constructed from the SD of each triplicate. Statistical calculation of p values demonstrating a statistical difference between $iLOV^{WT}$ and mutant SOSG responses to $^1O_2$ yield was performed using a t test. This was calculated using the "ttest2" function as part of the Statistics and Machine Learning toolbox add on to MATLAB.

Polyacrylamide gel extraction

The procedure used to gel extract peptides for Fig. S1 and Tables S1 and S2 is described in the Supplementary Information. After identifying the heavy molecular weight band as dimeric iLOV, the gel extraction procedure for iLOV samples was modified to increase peptide yields by excluding the reaction with iodoacetamide (iLOV lacks cysteine residues, so this omission decreases the number of washing steps required). Gel extraction of iLOV monomer and dimer bands reported in Figure 2 was performed as follows: illuminated samples of iLOV were resolved by electrophoresis on a 16% bisacrylamide gel. The monomer and dimer bands of interest were excised using a razor blade and minced to ∼1 mm^3 pieces. The minced gel sample was mixed 3 × with 150 μl 50 mM ammonium bicarbonate and then was pelleted *via* centrifugation at 3000*g*, and supernatant was discarded. Following this initial wash, the gel pieces were incubated in 500 μl acetonitrile for 10 min at room temperature. This acetonitrile was removed and the pieces were allowed to air-dry for 15 min at room temperature. At this point, 30 μl of 50 μg/ml trypsin in 50 mM ammonium bicarbonate was added to the gel pieces to reswell them. The samples were incubated on ice for 20 min. Once, the gel pieces had resumed their former size, enough 50 mM ammonium bicarbonate was added to just cover the gel pieces. The samples were incubated at 37 °C overnight with shaking. The digestion mixture was then moved to a new clean test tube. The same volume of 1% formic acid was added to the gel pieces and incubated for 15 min at room temperature. This extraction solution was removed and added to the digestion mixture. To extract hydrophobic peptides, 1% formic acid in 75% acetonitrile were added to the gel pieces following incubation for 15 min. This was added to the digestion/extraction mixture. The digestion/extraction mixture was placed on a SpeedVac to dry completely. At this point, about 30 μl of 1% formic acid in 30% acetonitrile were added to resolubilize digested peptides. These samples were then flash-frozen in liquid nitrogen and stored at −80 °C until analysis by HPLC-MS/MS as described below on a later date.

Peptide mapping by UPLC-MS/MS

Initially, tryptic peptides of gel-extracted iLOV samples exposed only to ambient light used for Fig. S1 and Tables S1 and S2 were resolved and analyzed as described in the Supplementary information (45–47). Tryptic peptides of gel-extracted iLOV samples intentionally illuminated with blue light (Figs. 2*B*, S2, S6, S7, and Tables S3, S4, and S6) were injected into a Shimadzu 9030 UPLC-MS/MS equipped with a C18 column with dimensions of 150 mm × 2.0 mm i.d. (2.2 μm particle size) equilibrated with aqueous 0.1% (v/v) formic acid (solution A) and resolved in the UPLC against an isocratic gradient from 0 to 60% (v/v) of acetonitrile with 0.1% (v/v) formic acid (solution B) over 90 min at a flow rate of 0.2 ml/min. For samples with $iLOV^{\Delta H}$, the gradient was changed to 0 to 55% over 25 min. The eluent was injected into the qTOF mass spectrometer calibrated with sodium iodide. Mass spectra were collected in data-dependent acquisition mode with MS^1 spectra collected over a range of 200 to 2500 m/z (or 200–1500 m/z for $iLOV^{\Delta H}$,samples). Five dependent events were chosen for fragmentation to create MS^2 spectra of precursor peptides with m/z values of 400 to 1500. An event time of 0.2 s, a base peak chromatogram intensity threshold ion count of 200, and a collision energy of 35 V was used. To identify potential W oxidation products in $iLOV^{\Delta H}$ the range in which precursor peptides were selected to fragment was narrowed to 685 to 730 corresponding to the mass range of the single peptide containing W81 (residues 80–90). In order to observe the effect this illumination had on another peptide of

iLOV$^{\Delta H}$, also formerly containing histidine (residues 94–110), this run was repeated with a fragmentation range of 910 to 960. However, no significant differences between the illuminated and nonilluminated spectra were detected for this run. The MS^1 spectra of these two otherwise identical replicates showed consistent results in terms of both the number of peaks in the chromatograms and the observed MS^1 *m/z* values. MS^2 spectra covered a mass range of 200 to 2500 *m/z* (or 200–1500 *m/z* for iLOV$^{\Delta H}$, samples).

Experimental design and statistical rationale

Data used for the construction of Figure 2*B*, S2, S6, S7, and Tables S3, S4, and S6 were analyzed using FragPipe v.22.0 (https://fragpipe.nesvilab.org/) (48–51). The data used for this analysis consisted of tryptic peptides originating from SDS page gel-purified bands intentionally illuminated with blue light. Positive controls included iLOV bands that were not illuminated by light. The data for Figure 2*B* originated from a single illumination injected in triplicate under identical MS parameters. The data in Figs. S6, S7, and Table S6 are the results of a single illumination, injected in duplicate, where each replicate was run with different MS^2 parameters targeting specific precursor peptides as described above. The data contained in each figure or table are a representative result from one of the replicates under the conditions previously specified.

Peak lists for data analyzed in FragPipe were exported directly from Shimadzu's LabSolutions Postrun software (https://www.shimadzu.com/an/products/software-informatics/labsolutions-series/index.html) in the form of .mzML files with centroid peaks. Analysis of data presented in Figure 2*B* was performed with a sequence database composed of the UniProt database accessed on 12/6/2022 in addition to the FASTA sequences for iLOV and cleaved $(H)_6$-SUMO added manually for a total of 569,340 entries. Decoy peptides were generated by FragPipe with reverse sequences of each entry as a negative control for the analysis. For data presented in Figs. S6, S7 and Table S6 the goal was to observe any modifications on specific precursors rather than for the identification of the protein species. As such, the sequence "database" only included the sequence for iLOV$^{\Delta H}$ along with its associated reverse sequence decoy added by the FragPipe software. Peptides from the database were generated by "digesting" the sequences with trypsin (cleaves C terminally of lysine/arginine except preceded by proline) with no allowed missed cleavages. No fixed modifications were used (no iodoacetamide was used to prepare these samples as described in the Methods section titled "Polyacrylamide gel extraction"), but variable modifications included oxidation on methionine (+15.9949 Da), N-terminal acetylation (+42.0106), oxidation on histidine (+31.9988 Da and +15.9994 Da), and oxidation on tryptophan (+31.9988 Da and +3.9994 Da). The mass tolerance for precursor peptides was set to a range of 40 ppm, while a mass tolerance for fragment ions was set to 2 Da. Statistical tests are employed by FragPipe's MSFragger module to generate PeptideProphet Probabilities as previously described (48). Within MSFragger MS^2 analysis, unfragmented precursor peaks were ignored (within an *m/z* range of −1.5–1.5 Da) and spectral processing identified the top 150 peaks from each spectrum then removed peaks with relative intensities of ≤0.01. In the analysis, no decoy peptides were detected.

Fluorescence lifetime measurements

Fluorescence lifetime measurements were performed using time-correlated single photon counting (TCSPC) technique (52). An approximately 100 femtosecond (fs) excitation pulses with wavelength of 400 nm was generated by doubling the fundamental frequency of a Ti:Sapphire laser (SpectraPhysics Tsunami) in a commercial optical harmonics generator (Inrad/Coherent). The laser repetition rate was reduced by a home-made acousto-optical pulse picker in order to avoid saturation of the chromophore. The TCSPC detection system is equipped with a single-photon counting avalanche photodiode module (Photon Micro Devices) and electronics board (Becker & Hickl SPC-630) and has an instrument response time of less than 30 ps. The triggering signal for the TCSPC board was generated by sending a small fraction of the laser beam onto a fast (400 MHz bandwidth) Si photodiode (Thorlabs Inc). Fluorescence signal was collected in 90 degrees geometry and dispersed in Acton Research SPC-500 monochromator after passing through a pump blocking, long wavelength-pass, autofluorescence-free, interference filter (Omega Filters, ALP series). The monochromator was equipped with a CCD camera (Roper Scientific PIXIS-400) allowing for monitoring of the fluorescence spectrum. Fluorescence transients were not deconvolved with the instrument response function since their characteristic decay time constants were much longer than the width of the system response to the excitation pulse.

MD simulations

MD modeling of O_2 diffusion into iLOV was performed similarly as previously described with minor modifications using the visual MD software (https://www.ks.uiuc.edu/Research/vmd/) (27, 53). The crystal structure PDB ID 4EES was used as the starting coordinates for heavy atoms of FMN and protein (15). Ten O_2 molecules were placed close to the protein surface randomly. Hydrogen atoms were added assuming the conventional protonation states of ionizable groups at pH 7.00, while histidine residues were modeled as neutral. Water molecules resolved in the crystal structure were not removed, and a solvation box was generated with an edge width of 96.497 Å. Sodium and chloride ions were added to a final concentration of 0.15 M. The additive force field CHARMM36 was employed to model iLOV, and waters were modeled with the TIP3P intermolecular potential function (54, 55). Force field parameters for oxidized FMN were obtained from a prior study (56) as well as those used for O_2 (57).

Classical MD simulations were performed using the NAMD 3.0 software package (https://www.ks.uiuc.edu/Research/namd/) (58). An isothermal-isobaric (NPT) ensemble was maintained at 300 K and 1.01325 atm with Nosé–Hoover Langevin piston pressure control and Langevin dynamics. Periodic boundary conditions were employed and a Particle Mesh Ewald algorithm was used to calculate long-range

interactions. A cut-off for nonbonded interactions was set to 12 Å. The SHAKE algorithm is used to constrain bonds involving hydrogen atoms. An integration time step of 2 fs was used and a total of 1425.15 ns were simulated in five independent simulations. RMSD plots of backbone atoms in iLOV over the course of each ∼285 ns simulation are shown in Fig. S8.

Conclusion

In this study, we have demonstrated that the iLOV protein photochemically generates $^1O_2$, which causes protein cross-linking *in vitro*. Based on our results from mutagenesis studies, the photochemical cross-linking this chemistry can be rationally modulated to either bias cross-linking or ameliorate it for specified applications *in vitro* and *in vivo*. We hope these results will inform biotechnological development of genetically encoded fluorescent tools and provide insight into the role of $^1O_2$ reactivity in biological systems with potential therapeutic applications.

Data availability

The MS proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) *via* the MassIVE partner repository with the data set identifier MSV000095393.

Supporting information—This article contains supporting information (41, 45, 46).

Acknowledgments—The authors acknowledge Prof. Arnab Mukherjee for the gracious gift of the iLOV plasmid and for fruitful discussion. The authors also acknowledge Juan Carlos Cáceres for helpful suggestions and guidance. The Department of Chemistry and Biochemistry Mass Spectrometry Facility instrumentation was supported by the Department of Defense DURIP grant number N00014-23-1-2197. The Optical Characterization Facility instrumentation was supported by DoD ARO DURIP grant number 66886LSRIP. The research reported here used shared facilities of the UC Santa Barbara Materials Research Science and Engineering Center (MRSEC, NSF DMR-1720256), a member of the Materials Research Facilities Network (http://www.mrfn.org). Additionally, the authors gratefully acknowledge the contributions of undergraduate researchers Brett Haynes, Ian O'Connor, and Collin Origer.

Author contributions—B. J. J. and B. L. G. investigation; B. J. J. and B. L. G. methodology; B. J. J. and B. L. G. formal analysis; B. J. J. and B. L. G. writing–review and editing; B. L. G. supervision; B. J. J. and B. L. G. conceptualization; B. J. J. data curation; B. J. J. and B. L. G. writing–original draft; B. J. J. and B. L. G. visualization; B. J. J. software; B. L. G. resources; B. L. G. project administration; B. L. G. funding acquisition.

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: FMN, flavin mononucleotide; H_2O_2, hydrogen peroxide; iLOV, improved light-oxygen-voltage; LED, light-emitting diode; LOV, light-oxygen-voltage; MD, molecular dynamics; moniSOG, mini $^1O_2$ generator; MS, mass spectrum; Ni-NTA, nickel nitrilotriacetic acid; ROS, reactive oxygen species; SDM, site-directed mutagenesis; SOD, superoxide dismutase; SOSG, singlet oxygen sensor green; SUMO, small ubiquitin-like modifier protein; TCSPC, time-correlated single photon counting; ULP1, ubiquitin-like protease 1; UPLC-MS/MS, ultraperformance liquid chromatography tandem ToF/quadrupole mass spectrometry; XO, xanthine oxidase.

References

1. Fridovich, I. (1999) Fundamental aspects of reactive oxygen species, or what's the matter with oxygen? *Ann. N. Y. Acad. Sci.* **893**, 13–18
2. Khademian, M., and Imlay, J. A. (2020) Do reactive oxygen species or does oxygen itself confer obligate anaerobiosis? The case of *Bacteroides thetaiotaomicron*. *Mol. Microbiol.* **114**, 333–347
3. Zuo, L., Zhou, T., Pannell, B. K., Ziegler, A. C., and Best, T. M. (2015) Biological and physiological role of reactive oxygen species – the good, the bad and the ugly. *Acta Physiol.* **214**, 329–348
4. Nathan, C., and Cunningham-Bussel, A. (2013) Beyond oxidative stress: an immunologist's guide to reactive oxygen species. *Nat. Rev. Immunol.* **13**, 349–361
5. Feig, D. I., Reid, T. M., and Loeb, L. A. (1994) Reactive oxygen species in tumorigenesis. *Cancer Res.* **54**, 1890s–1894s
6. Dixon, S. J., and Stockwell, B. R. (2014) The role of iron and reactive oxygen species in cell death. *Nat. Chem. Biol.* **10**, 9–17
7. Schreck, R., Rieber, P., and Baeuerle, P. A. (1991) Reactive oxygen intermediates as apparently widely used messengers in the activation of the NF-kappa B transcription factor and HIV-1. *EMBO J.* **10**, 2247–2258
8. Greene, B. L., Kang, G., Cui, C., Bennati, M., Nocera, D. G., Drennan, C. L., *et al.* (2020) Ribonucleotide reductases: structure, chemistry, and metabolism suggest new therapeutic targets. *Annu. Rev. Biochem.* **89**, 45–75
9. Foote, C. S. (1968) Photosensitized oxygenations and the role of singlet oxygen. *Acc. Chem. Res.* **1**, 104–110
10. Shu, X., Lev-Ram, V., Deerinck, T. J., Qi, Y., Ramko, E. B., Davidson, M. W., *et al.* (2011) A genetically encoded tag for correlated light and electron microscopy of intact cells, tissues, and organisms. *PLoS Biol.* **9**, e1001041
11. Westberg, M., Holmegaard, L., Pimenta, F. M., Etzerodt, M., and Ogilby, P. R. (2015) Rational design of an efficient, genetically encodable, protein-encased singlet oxygen photosensitizer. *J. Am. Chem. Soc.* **137**, 1632–1642
12. Ruiz-González, R., Cortajarena, A. L., Mejias, S. H., Agut, M., Nonell, S., and Flors, C. (2013) Singlet oxygen generation by the genetically encoded tag miniSOG. *J. Am. Chem. Soc.* **135**, 9564–9567
13. Pimenta, F. M., Jensen, R. L., Breitenbach, T., Etzerodt, M., and Ogilby, P. R. (2013) Oxygen-dependent photochemistry and photophysics of "MiniSOG," a protein-encased flavin. *Photochem. Photobiol.* **89**, 1116–1126
14. Di Mascio, P., Martinez, G. R., Miyamoto, S., Ronsein, G. E., Medeiros, M. H. G., and Cadet, J. (2019) Singlet molecular oxygen reactions with nucleic acids, lipids, and proteins. *Chem. Rev.* **119**, 2043–2086
15. Christie, J. M., Hitomi, K., Arvai, A. S., Hartfield, K. A., Mettlen, M., Pratt, A. J., *et al.* (2012) Structural tuning of the fluorescent protein iLOV for improved photostability. *J. Biol. Chem.* **287**, 22295–22304
16. Liu, M., Zhang, Z., Cheetham, J., Ren, D., and Zhou, Z. S. (2014) Discovery and characterization of a photo-oxidative histidine-histidine cross-link in IgG1 antibody utilizing 18O-labeling and mass spectrometry. *Anal. Chem.* **86**, 4940–4948
17. Kopka, B., Magerl, K., Savitsky, A., Davari, M. D., Röllen, K., Bocola, M., *et al.* (2017) Electron transfer pathways in a light, oxygen, voltage (LOV) protein devoid of the photoactive cysteine. *Sci. Rep.* **7**, 13346
18. Tanaka, K., Miura, T., Umezawa, N., Urano, Y., Kikuchi, K., Higuchi, T., *et al.* (2001) Rational design of fluorescein-based fluorescence probes. Mechanism-based design of a maximum fluorescence probe for singlet oxygen. *J. Am. Chem. Soc.* **123**, 2530–2536

19. *Singlet Oxygen Sensor Green Reagent*. (2004). Invitrogen, Eugene, OR
20. Ludvíková, L., Štacko, P., Sperry, J., and Klán, P. (2018) Photosensitized cross-linking of tryptophan and tyrosine derivatives by rose bengal in aqueous solutions. *J. Org. Chem.* **83**, 10835–10844
21. De Pascali, F., Hemann, C., Samons, K., Chen, C.-A., and Zweier, J. L. (2014) Hypoxia and reoxygenation induce endothelial nitric oxide synthase uncoupling in endothelial cells through tetrahydrobiopterin depletion and S-Glutathionylation. *Biochemistry* **53**, 3679–3688
22. Lee, M.-C., Velayutham, M., Komatsu, T., Hille, R., and Zweier, J. L. (2014) Measurement and characterization of superoxide generation from xanthine dehydrogenase: a redox-regulated pathway of radical generation in ischemic tissues. *Biochemistry* **53**, 6615–6623
23. Marques, E. F., Medeiros, M. H. G., and Di Mascio, P. (2019) Singlet oxygen-induced protein aggregation: lysozyme crosslink formation and nLC-MS/MS characterization. *J. Mass Spectrom.* **54**, 894–905
24. Davies, M. J. (2003) Singlet oxygen-mediated damage to proteins and its consequences. *Biochem. Biophys. Res. Commun.* **305**, 761–770
25. Davies, M. J. (2004) Reactive species formed on proteins exposed to singlet oxygen. *Photochem. Photobiol. Sci.* **3**, 17–25
26. Schweitzer, C., and Schmidt, R. (2003) Physical mechanisms of generation and deactivation of singlet oxygen. *Chem. Rev.* **103**, 1685–1758
27. Polyakov, I., Kulakova, A., and Nemukhin, A. (2023) Computational modeling of the interaction of molecular oxygen with the miniSOG protein—a light induced source of singlet oxygen. *Biophysica* **3**, 252–262
28. Pietra, F. (2014) Molecular dynamics simulation of dioxygen pathways through mini singlet oxygen generator (miniSOG), a genetically encoded marker and killer protein. *Chem. Biodivers.* **11**, 1883–1891
29. Sies, H., Berndt, C., and Jones, D. P. (2017) Oxidative stress. *Annu. Rev. Biochem.* **86**, 715–748
30. Xu, S., and Chisholm, A. D. (2016) Highly efficient optogenetic cell ablation in C. elegans using membrane-targeted miniSOG. *Sci. Rep.* **6**, 21271
31. Westberg, M., Bregnhøj, M., Etzerodt, M., and Ogilby, P. R. (2017) No photon wasted: an efficient and selective singlet oxygen photosensitizing protein. *J. Phys. Chem. B.* **121**, 9366–9371
32. Makhijani, K., To, T.-L., Ruiz-González, R., Lafaye, C., Royant, A., and Shu, X. (2017) Precision optogenetic tool for selective single- and multiple-cell ablation in a live animal model system. *Cell Chem. Biol.* **24**, 110–119
33. Westberg, M., Etzerodt, M., and Ogilby, P. R. (2019) Rational design of genetically encoded singlet oxygen photosensitizing proteins. *Curr. Opin. Struct. Biol.* **57**, 56–62
34. Trewin, A. J., Berry, B. J., Wei, A. Y., Bahr, L. L., Foster, T. H., and Wojtovich, A. P. (2018) Light-induced oxidant production by fluorescent proteins. *Free Radic. Biol. Med.* **128**, 157–164
35. Kim, S. Y., Kwon, O. J., and Park, J.-W. (2001) Inactivation of catalase and superoxide dismutase by singlet oxygen derived from photoactivated dye. *Biochimie* **83**, 437–444
36. Escobar, J. A., Rubio, M. A., and Lissi, E. A. (1996) SOD and catalase inactivation by singlet oxygen and peroxyl radicals. *Free Radic. Biol. Med.* **20**, 285–290
37. Méndez-Hurtado, J., López, R., Suárez, D., and Menéndez, M. I. (2012) Theoretical study of the oxidation of histidine by singlet oxygen. *Chem. – Eur. J.* **18**, 8437–8447
38. Anderson, N. T., Weyant, K. B., and Mukherjee, A. (2020) Characterization of flavin binding in oxygen-independent fluorescent reporters. *Aiche J.* **66**, e17083
39. Guerrero, F., Ciragan, A., and Iwaï, H. (2015) Tandem SUMO fusion vectors for improving soluble protein expression and purification. *Protein Expr. Purif.* **116**, 42–49
40. Cáceres, J. C., Bailey, C. A., Yokoyama, K., and Greene, B. L. (2022) Selenocysteine substitutions in thiyl radical enzymes. *Methods Enzymol.* **662**, 119–141
41. Cáceres, J. C., Dolmatch, A., and Greene, B. L. (2023) The mechanism of inhibition of pyruvate formate lyase by methacrylate. *J. Am. Chem. Soc.* **145**, 22504–22515
42. Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., *et al.* (1985) Measurement of protein using bicinchoninic acid. *Anal. Biochem.* **150**, 76–85
43. Davari, M. D., Kopka, B., Wingen, M., Bocola, M., Drepper, T., Jaeger, K.-E., *et al.* (2016) Photophysics of the LOV-based fluorescent protein variant iLOV-Q489K determined by simulation and experiment. *J. Phys. Chem. B.* **120**, 3344–3352
44. van den Berg, P. A. W., Widengren, J., Hink, M. A., Rigler, R., and Visser, A. J. W. G. (2001) Fluorescence correlation spectroscopy of flavins and flavoenzymes: photochemical and photophysical aspects. *Spectrochim. Acta A. Mol. Biomol. Spectrosc.* **57**, 2135–2144
45. *ProteinLynx GLobal SERVER Version 2.2.5, Revision A*. (2006). Waters corporation, USA
46. Nesvizhskii, A. I., Keller, A., Kolker, E., and Aebersold, R. (2003) A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* **75**, 4646–4658
47. Li, G., Vissers, J. P. C., Silva, J. C., Golick, D., Gorenstein, M. V., and Geromanos, S. J. (2009) Database searching and accounting of multiplexed precursor and product ion spectra from the data independent analysis of simple and complex peptide mixtures. *Proteomics* **9**, 1696–1719
48. Kong, A. T., Leprevost, F. V., Avtonomov, D. M., Mellacheruvu, D., and Nesvizhskii, A. I. (2017) MSFragger: ultrafast and comprehensive peptide identification in shotgun proteomics. *Nat. Methods* **14**, 513–520
49. Yu, F., Teo, G. C., Kong, A. T., Haynes, S. E., Avtonomov, D. M., Geiszler, D. J., *et al.* (2020) Identification of modified peptides using localization-aware open search. *Nat. Commun.* **11**, 1–9
50. Yu, F., Haynes, S. E., Teo, G. C., Avtonomov, D. M., Polasky, D. A., and Nesvizhskii, A. I. (2020) Fast quantitative analysis of timsTOF PASEF data with MSFragger and IonQuant. *Mol. Cell Proteomics* **10**, 1575–1585
51. Teo, G. C., Polasky, D. A., Yu, F., and Nesvizhskii, A. I. (2020) A fast deisotoping algorithm and its implementation in the MSFragger search engine. *J. Proteome Res.* **20**, 498–505
52. Becker, W. (2005) *Advanced Time-Correlated Single Photon Counting Techniques*, Springer Science & Business Media, Cham, Switzerland
53. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38
54. Best, R. B., Zhu, X., Shim, J., Lopes, P. E. M., Mittal, J., Feig, M., *et al.* (2012) Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone Φ, ψ and side-chain χ1 and χ2 dihedral angles. *J. Chem. Theor. Comput.* **8**, 3257–3273
55. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935
56. Aleksandrov, A. (2019) A molecular mechanics model for flavins. *J. Comput. Chem.* **40**, 2834–2842
57. Wang, S., Hou, K., and Heinz, H. (2021) Accurate and compatible force fields for molecular oxygen, nitrogen, and hydrogen to simulate gases, electrolytes, and heterogeneous interfaces. *J. Chem. Theor. Comput.* **17**, 5198–5213
58. Phillips, J. C., Hardy, D. J., Maia, J. D. C., Stone, J. E., Ribeiro, J. V., Bernardi, R. C., *et al.* (2020) Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J. Chem. Phys.* **153**, 044130