Targeted Activation in Localized Protein Environments via Deep Red Photoredox Catalysis

Nicholas Eng Soon Tay[1,4‡], Keun Ah Ryu[2‡], John L. Weber[1], Aleksandra K. Olow[3], David R. Reichman[1], Rob C. Oslund[2*], Olugbeminiyi O. Fadeyi[2*], Tomislav Rovis[1*]

[1]Department of Chemistry, Columbia University, 3000 Broadway, New York, NY, 10027, USA
[2]Exploratory Science Center, Merck & Co., Inc., Cambridge, MA, 02139, USA
[3]Genetics and Pharmacogenomics, Merck & Co., Inc., South San Francisco, CA, 94080, USA
[‡]These authors contributed equally
*Corresponding authors: rob.oslund@merck.com, olugbeminiyi.fadeyi@merck.com, tr2504@columbia.edu

Abstract

State-of-the art photoactivation strategies in chemical biology provide spatiotemporal control and visualization of biological processes. However, using high energy light (λ < 500 nm) for substrate or photocatalyst sensitization can lead to background activation of photoactive small molecule probes and reduce its efficacy in complex biological environments. Here we describe the development of targeted aryl azide activation via deep red light (λ = 660 nm) photoredox catalysis and its use in photocatalyzed proximity labeling. We demonstrate that aryl azides are converted to triplet nitrenes via a novel redox-centric mechanism and show that its spatially localized-formation requires both red light and a photocatalyst-targeting modality. This technology was applied in different colon cancer cell systems for targeted protein environment labeling of epithelial cell adhesion molecule (EpCAM). We identified a small subset of proteins with previously known and unknown association to EpCAM, including CDH3, a clinically relevant protein that shares high tumor selective expression with EpCAM.

Introduction

New targeted activation strategies are essential for uncovering fundamental biological processes that inspire clinical applications. From methods that enable cellular microenvironment mapping to technologies that hijack biological machinery, fulfilling this translational promise requires a deep understanding of the chemical principles that underlie these tools. A prime example is the use of aryl azides which has widespread use in many chemical biology-based applications within the context of photoaffinity labeling (PAL)[1,2], photo-uncaging[3,4], protein conjugation[5,6,7,8], protein cross-linking[9], and enzyme-mediated protein labeling[10]. Of these approaches, the ability to activate aryl azides via ultraviolet (UV) or visible light irradiation has surfaced as a highly attractive approach for labeling, uncaging, or conjugation (Figure 1A). However, limitations such as unselective/background labeling, a requirement for high energy light, and the generation of multiple reaction pathways from current photoactivation methods restrict downstream biological applications. As a consequence, direct photolysis presents challenges for activation of aryl azides in a spatially-controlled manner. This has important implications for profiling protein interactions within complex physiological environments. Thus, low-energy light driven technologies that enable spatio-temporal activation of aryl azides are highly desirable.

The identity of the aryl azide plays an important role in determining the photoactivation reaction pathways (Figure 1B). For example, singlet aryl nitrene (**1**) ($t_{1/2}\sim$ 1–10 ns)[11] is formed via UV photolysis of unsubstituted aryl azides. This leads to rapid two-step rearrangement to the benzazirine and subsequent ring-expansion to a long-lived 7-membered ketenimine (**2**) that is susceptible to hydrolysis or nucleophilic trapping ($t_{1/2}\sim$ 5 ms–1 s depending on the

concentration of intercepting species)[12,13,14]. To overcome this limitation, perfluorinated aryl azides (PFAA) have been adopted[15,16] to enable complementary reactivity through the singlet nitrene, leading to *N*-centered covalent adducts[17,18]. This intermediate can additionally relax via intersystem crossing (ISC) (k_{ISC} ~ 10^6–10^7 s^{-1})[19] to a triplet nitrene[20,21], which has divergent reactivity relative to its singlet form[17,21,22]. Given that multiple reactive pathways are generated through direct high energy light photoactivation, identifying methods that incorporate lower energy light systems and do not directly activate aryl azides (Figure 1B), would allow for a targeted activation platform in physiologically relevant environments.

Visible-light photocatalysis is an emerging strategy for probing complex biological systems as the selective delivery of photons enables catalytic, chromoselective substrate activation via the controlled formation of high-energy species that covalently bind to desired biomolecule targets.[23] To fully realize the potential of visible light, a paradigm shift in photoactivation must turn focus from substrate to catalyst, and low energy wavelengths of light—ideally in the deep red (DR) or near-infrared (NIR) range ($\lambda > 600$ nm)— used to maximize the physiological relevance of photocatalytic applications. To achieve this goal, the discovery and implementation of new systems involving localized, catalytic activation of latent chemical probes will be key. In particular, recent endeavors of photoredox catalysis via light-induced electron transfer (ET) or energy transfer (EnT) into chemical biology have been particularly fruitful, leading to early-stage applications for cancer phototherapy[24] and proximity labeling[25]. In the latter example, carbene intermediates are used, as they offer a tight labeling radius ($t_{1/2}$~ 4 ns), but suffer from background quenching with water and oxygen in their singlet and triplet forms, respectively.[26] Replacing these reactive species with triplet nitrenes, which are less sensitive to reactions with

oxygen[17,27] and protic solvents[18], will improve reactivity with biomolecules in complex physiological environments relative to carbenes. Additionally, versatility of aryl azides beyond proximity labeling, especially within the context of targeted photo-uncaging via azide-to-aniline reduction, should encourage exploratory efforts into new mechanisms for aryl azide activation and localized drug delivery.[4]

Although photocatalytic approaches utilizing high-energy blue light (450 nm) to form the nitrene intermediate via ET[28] or EnT[29,30] mechanisms have been reported, the potential for direct photoactivation of aryl azides through these methods hinder their utility for probing biological environments in a targeted fashion. In view of this limitation, we sought to develop a new low energy light-based catalytic platform for biomolecule labeling via triplet nitrene generation (Figure 1C). Recently, we disclosed a method for DR and NIR photoredox catalysis where osmium photocatalysts harvest low-energy light to facilitate synthetic reactions via ET and EnT[31]. We therefore envisioned a DR light-based photocatalytic labeling approach in which an osmium complex is tethered to a targeting modality for localized generation of reactive triplet nitrene species within the desired region of a complex biological system (Figure 1D). Given the inherent value of understanding protein communities within disease states such a technology would unlock new strategies for selective therapeutic development. With this in mind, here we report the development of a DR light-mediated photoredox technology and use it to profile the local protein environment of EpCAM, a tumor selective antigen, within different colon cancer monoculture, spheroid, and dissociated tumor tissue cell systems.

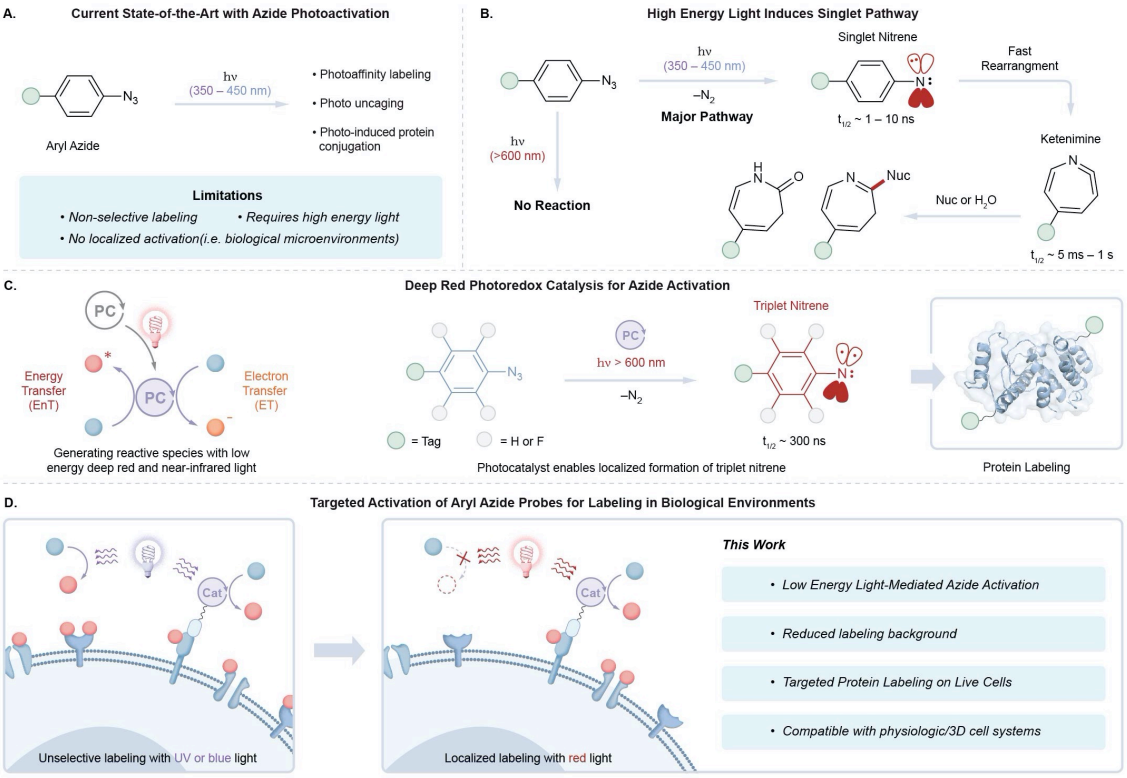

Figure 1. (A) Contemporary uses for azide photolysis in chemical biology. (B) Mechanism of azide photolysis and relevant intermediates. (C) Leveraging electron- or energy transfer mechanisms associated with deep red and near-infrared photoredox catalysis for azide activation. (D) Targeted azide activation for biological labeling with red light.

Results

To examine the impact of blue light irradiation on background azide activation (Figure 2A), we performed a trapping experiment with nonfluorinated azide **3**, which is known to form a ketenimine intermediate upon photolysis. Using 468 nm LEDs, we observed 85% yield of azepine **4** alongside a 5% yield of aniline **5**. This result was surprising as it suggests that irradiating the absorption region tail of azide **3** is sufficient to induce azide photolysis (Figure 2B). To understand the role of a photocatalyst in the activation mechanism, we next irradiated azide **3** in the

presence of catalytic iridium photocatalyst (**6**) — a well-documented triplet sensitizer with an excited state oxidation potential ($E(Ir^{IV/III*})$ = –0.89 V vs SCE) insufficient for ET-mediated azide activation. Contrary to expected azepine product, we instead isolated aniline **5** in 75% yield, with only 3% of **4** observed. This result suggests that aryl azide activation via energy transfer results in triplet nitrene formation that does not produce the ketenimine intermediate as recently proposed.[30] Noting that one-electron reduction pathways can lead to aniline formation, we used an osmium photocatalyst (**7**) with a greater redox match ($E(Os^{III/II*})$ = –1.05 V vs SCE) with the azide reduction potential ($E_{p/2}$ = –1.25 V vs SCE), and obtained **5** in 85% yield when irradiated with 468 nm LEDs. Given the broadband absorption of photocatalyst **7** and its defined triplet metal-ligand charge transfer (^{3}MLCT) band in the 530–700 nm range (Figure 2B), we were interested in the effects of decoupling azide light- and redox-dependent activation pathways. When the azide is irradiated with 660 nm LEDs in the absence of photocatalyst, full recovery of the azide is observed (see supplemental information) whereas when **7** is added, 50% conversion of **3** to **5** is observed. These results reveal that (1) no background activation is observed under red light irradiation, (2) light and redox decoupling is unviable when using blue light, (3) an exergonic redox match between the photocatalyst and azide may be necessary for efficient activation. To investigate this latter point, we examined the conversion of perfluorinated azide **8** (Ep/2 = –0.79 V vs. SCE) to aniline **9**, and found that quantitative formation of 9 was observed in an hour, with 2-hydroxytetrahydrofuran as the only byproduct (Figure 2C). The reaction is dependent on the reducing ability of the photocatalyst rather than triplet energy as photocatalysts **10** and **11**, which have similar triplet energies but lower $E(Os^{III/II*})$, are less efficient at converting the azide to aniline.

Importantly, we further found this aniline reduction method displays broad substrate compatibility within the context of PFAA featuring several commonly-used bio-orthogonal probes (Figure 2D). Simple functional groups such as benzamide (**12**), benzyl alcohol (**13**), nitrile (**14**), biotin (**15**), sugar (**16**–**19**), and tryptamine (**20**) substrates are unaffected, with the aniline obtained in near quantitative yields. Bioorthogonal groups such as alkyl azide (**21**) and alkyne (**22**) are also unreactive, with the former example notable as the conditions developed by Liu and coworkers activate both alkyl and aryl azides[28]. Alkyl diazirine (**23**) is also compatible with the reduction, highlighting the potential for achieving tandem photoactivation using diazirine and PFAA as complementary photoactivatable groups. Finally, we found that this photocatalytic reduction strategy could also be used for prodrug release, as exemplified by catalytic uncaging of PFAA-bound hydroxycoumarin, sugars, and doxorubicin (Figure S1).

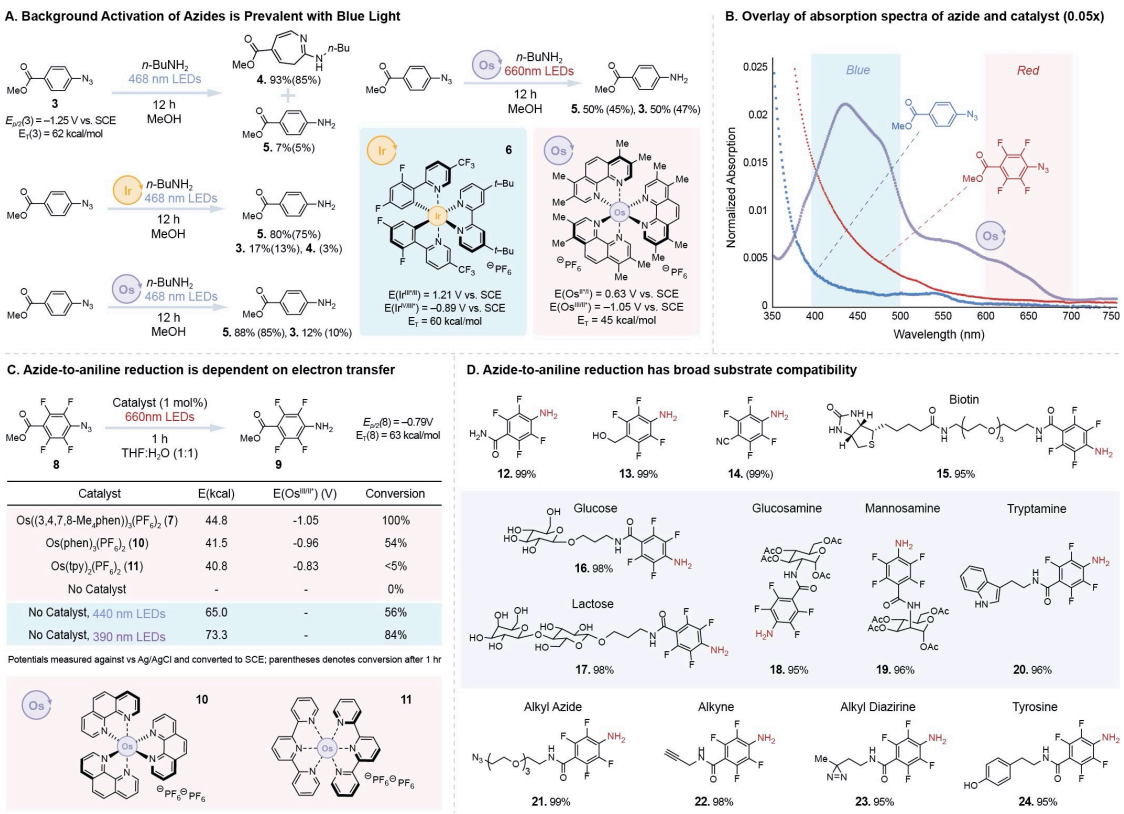

Figure 2. (A) Mechanistic investigations into the photolysis and catalytic activation of azides, isolated yields in parentheses. (B) Overlaid absorption spectrums of unsubstituted azide **3**, perfluorinated azide **8**, and osmium photocatalyst **7**. (C) Azide-to-aniline reduction studies suggest an electron transfer pathway is necessary. (D) Broad substrate compatibility is observed for azide-to-aniline reduction (See supplemental information for experimental conditions, ^{19}F NMR yields in parentheses).

We next undertook density functional theory (DFT) calculations to elucidate the potential mechanism for obtaining triplet nitrene. We ruled out Dexter EnT as a viable pathway as the vertical and adiabatic singlet-to-triplet energies (ΔG = 88.3 and 63.3 kcal/mol respectively) necessary for obtaining the triplet azide **25** are significantly greater than the triplet energy of the osmium photocatalyst **7** (ΔG = 44.8 kcal/mol).[32] By contrast, we found an ET transfer mechanism to be energetically favorable as the reduced azide (**26**) could be significantly stabilized through

both solvation and fluorination (Figure 3A). Additionally, we found that the barrier to N_2 dissociation upon reduction is minimal (Figure 3B) and, as expected, N_2 loss is highly exothermic (Figure 3A). Subsequent exothermic re-oxidation of the reduced nitrene (**27**) by $\mathbf{7}^{+}$ then reveals the triplet nitrene (**28**). Thus, we propose a photoredox-catalyzed, stepwise reduction-dissociation-oxidation pathway occurs (Figure 3C), taking advantage of the stability of charged intermediates to form triplet nitrene **28**. This novel mechanism is further supported by qualitative trends with solvation and fluorination; using a solvent with a lower dielectric constant or replacing fluorine with hydrogen raises the energy of the anionic azide (whereas the triplet energy remains approximately identical), resulting in a lower reaction completion rate (Figure S2, and Table S1).

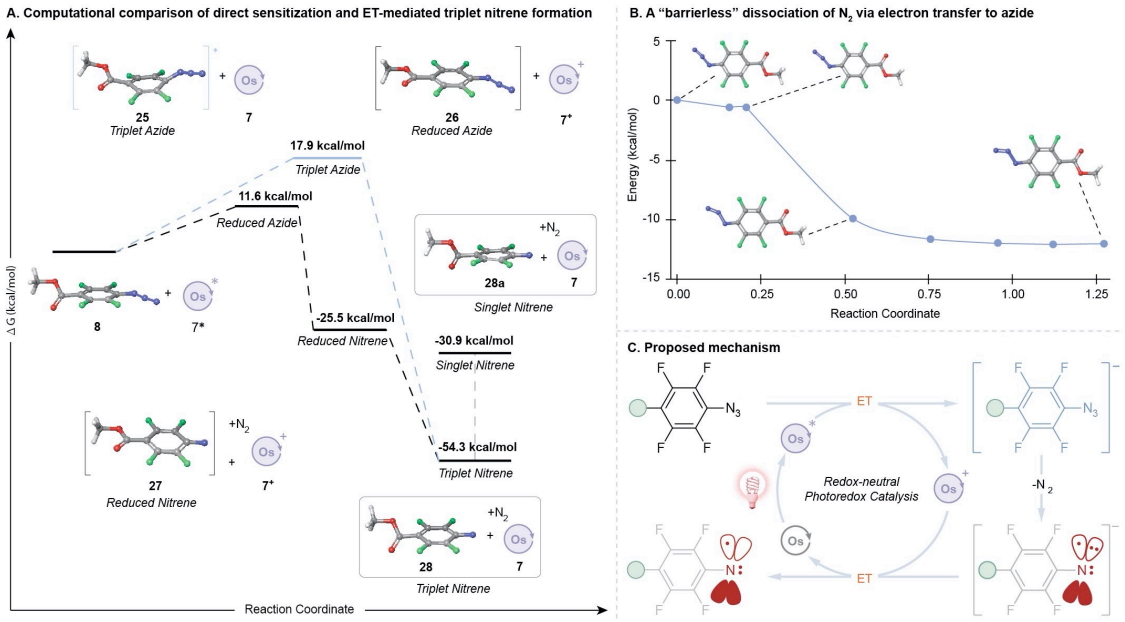

Figure 3. (A) Density functional theory analyses of azide-to-triplet nitrene pathways via electron transfer or triplet Dexter energy transfer. (B) Computational model for electron-transfer-mediated N_2 dissociation from the azide. (C) Putative mechanism for accessing triplet nitrenes via a redox-neutral photoredox cycle.

To further confirm the triplet nitrene pathway, we compared the product distributions for our method against UV- or blue-light photolysis for toluene (Figure 4, top) and dimethyl sulfoxide (Figure 4, bottom). UV photolysis of **8** in the presence of toluene yields a mixture of aryl and benzylic amination products (**29** and **30** respectively), whereas our DR photoredox method yields only the benzylic C–H insertion product from toluene, as well as azobenzene **31**. Observing the latter product is important as it is indicative of triplet nitrene formation[11,21]. Blue-light photolysis of **8** in dimethyl sulfoxide leads to quantitative formation of sulfoximine **32** whereas our DR photoredox method yields a 1:1 mixture of **31** and **32**. Furthermore, investigation of additional reaction partners reveals general reactivity with C(sp^3)–H bonds, alkenes, and sulfides

with azide **8** under DR photoredox conditions (Figure S3), suggesting the utility of exploiting triplet perfluoroaryl nitrenes to label within biological environments where many of these functional groups can be found. Taken together, these results show that our photoredox method enables high-energy nitrene formation with low-energy light via a redox mechanism, and the resulting triplet perfluorinated aryl nitrenes have similar reactivity to their singlet counterparts, although stepwise mechanisms are likely involved.

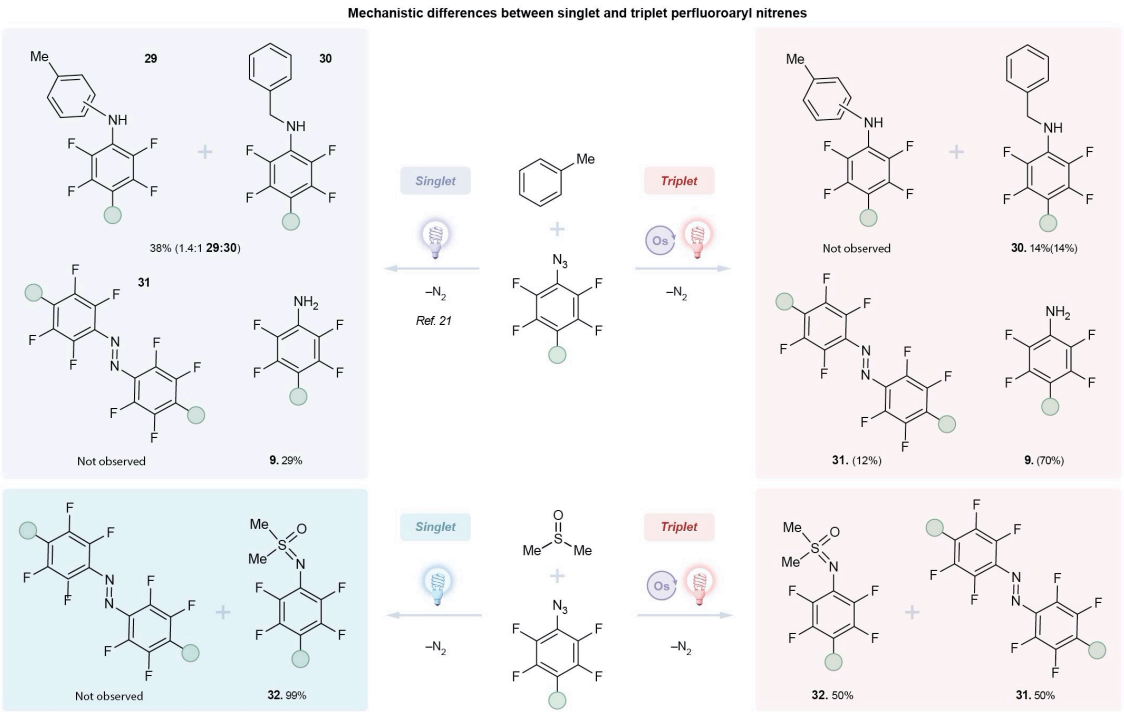

Figure 4. Distinguishing between singlet and triplet nitrene mechanisms for perfluoroaryl azides for C–H amination of toluene (top) and imidation of dimethyl sulfoxide (bottom) (See supplemental information for experimental conditions, ^{19}F NMR yields are listed in parentheses).

Having identified DR photoredox conditions that induce activation of various PFAAs, we next investigated the application of this method to achieve covalent protein labeling in aqueous buffer conditions. In anticipation of utilizing this photocatalyst system in a targeted fashion within biological environments, we designed an osmium photocatalyst containing an alkyne linker for easy attachment to targeting modalities (Figure S4). As a proof-of-concept experiment, we combined carbonic anhydrase (CA) with PFAA-biotin (**15**) and free osmium-alkyne photocatalyst followed by irradiation with red light (660 nm) using a recently developed biophotoreactor[33] (Figure 5A). Under these conditions, we observed biotinylation of CA in the presence of both the photocatalyst and red light whereas the absence of light or the use of red

light without photocatalyst does not induce biotinylation (Figure 5B). Comparable results were achieved using osmium photocatalysts without the alkyne linker (Figure S5). Also, similar to our chemical transformation results in Figure 2, we found that the direct exposure of PFAA to blue light led to biotinylation with CA (Figure 5B), further highlighting that background labeling can occur from direct activation of aryl azides with higher energy wavelengths of light. Finally, we investigated protein labeling as a function of red-light exposure time and observed increasing biotinylation levels as a function of light duration confirming the light dependent nature of our labeling system (Figure 5C).

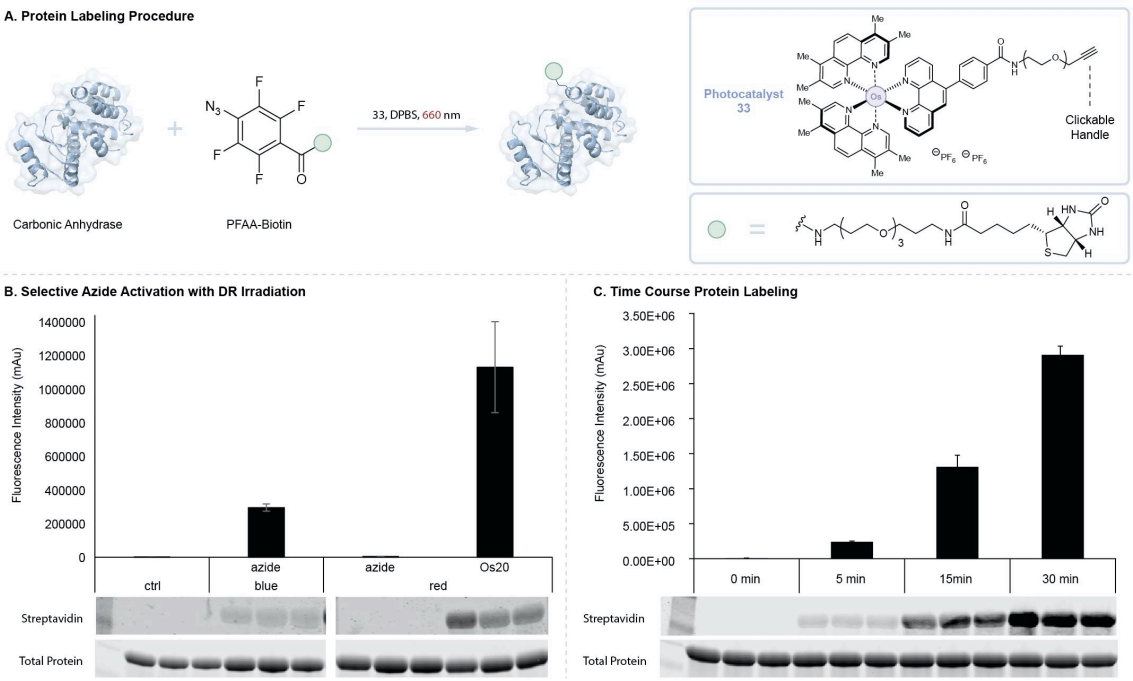

Figure 5. DR light-mediated protein labeling. A) Schematic depicting labeling of carbonic anhydrase in the presence of osmium-alkyne photocatalyst, PFAA-biotin, and red light. B) Western blot image of carbonic anhydrase labeled with PFAA-biotin in the presence or absence of blue light or presence or absence of red light and photocatalyst. Protein biotinylation is observed with red light and photocatalyst as well as blue light alone. Bar plots of replicate analysis of protein biotinylation measured by western blot. Error bars represent standard deviation of n = 3 experiments. C) Time course of protein biotinylation. Western blot analysis shows increased protein labeling over a 30-minute time course. Bar plots of replicate analysis of protein biotinylation measured by western blot. Error bars represent standard deviation of n = 3 experiments.

Light dependent protein labeling via our DR photoredox technology opens the possibility to achieve targeted labeling of protein environments within cell-based systems. The high expression of EpCAM on many different tumor types has led to intense focus of EpCAM as a tumor associated antigen and biomarker[34,35]. Thus, enhancing our understanding of EpCAM cancer cell surface microenvironments can provide additional avenues for development of

tumor-selective therapeutics and biomarker-based diagnostics. Accordingly, we targeted EpCAM on the surface of colorectal cancer cells in different HCT116 cell systems with our osmium photocatalyst followed by irradiation with DR light to induce cell surface biotinylation for downstream protein enrichment, mass spectrometry-based proteomics, and bioinformatic analysis (Figure 6A). Targeted labeling of EpCAM was performed using a two-antibody system that consists of a primary α-EpCAM antibody and secondary antibody labeled with an osmium-alkyne photocatalyst (Figure S4) that binds to the primary antibody. HCT116 cells were labeled with the primary and secondary antibodies followed by addition of PFAA-biotin and DR-light irradiation. Similar to our protein labeling experiments, western blot analysis showed increased protein biotinylation that correlates with irradiation time (Figure 6B). As a negative control, targeted labeling with a non-binding isotype control primary antibody and secondary antibody photocatalyst resulted in background levels of protein biotinylation (Figure 6B). To investigate the potential of our DR photoredox system to generate reactive oxygen species (ROS) that can damage cellular environments, we measured singlet oxygen and hydrogen peroxide generation and observed minimal to no ROS increase compared to no light controls (Figures S6 and S7). This result was also consistent with the observation that cell viability is not affected by DR-light irradiation of cells labeled with the osmium-conjugated antibody system (Figure S8). Collectively, these results highlight compatibility of our osmium photocatalyst system for profiling within live cell environments.

With this photocatalytic targeting system in hand, we investigated EpCAM protein microenvironments on HCT116 cells within three distinct cell systems. This includes HCT116 cells grown in monoculture and detached in single cell suspension, grown in 3-D spheroid culture, or

dissociated cells from mice tumor xenograft tissue (Figure 6A). In each of these cell systems, EpCAM targeted biotinylation was performed with our antibody targeting approach followed by cell lysis and streptavidin bead-based enrichment of labeled proteins. Streptavidin-enriched proteins were then identified and quantified using TMT-based LC-MS/MS proteomic analysis. Volcano plot analysis was then performed to identify statistically significant enriched proteins that resulted from EpCAM targeting. EpCAM was detected as highly enriched in all three cell systems, confirming the ability of our technology to label and enrich the targeted protein (Figure S9-S11).

We next focused our attention on other commonly enriched membrane proteins, in addition to EpCAM, identified by our DR photocatalytic labeling technology. The surface proteins CD109, CD44, CDCP1, CDH3, DCBLD2, DDR1, DSC2, DSG2, ITGA2, ITGA3, NT5E, PTGFRN were enriched in at least two of the tested cell systems with EpCAM, CD44, CDH3, ITGA2, ITGA3 being enriched in all three (Figure 6C). The search tool for retrieval of interacting genes (STRING)[36] was used to obtain protein–protein interaction (PPI) network information for these proteins revealing both known and inferred protein interactions (Figure 6D). Further bioinformatic analysis of GO terms associated with proteins enriched in two or more cell systems resulted in the identification of biological processes associated with EpCAM function that include cellular adhesion, junction organization, and migration (Figure S12).

Additionally, we explored the gene expression profile for these enriched proteins across tumor vs normal tissue using data obtained from The Cancer Genome Atlas (TCGA) and Genotype-Tissue Expression (GTEx), respectively. Along with EpCAM, several genes (CDH3, DSG2, CDCP1, DSC2, and ITGA2) showed increased gene expression levels in tumor over normal tissue

(Figure 6E). Given our HCT116 cell system is derived from colorectal cancer, we performed a breakdown of gene expression for each of the enriched proteins by tumor indication (Figure 6F) and observed increased tumor-associated gene expression across different tumor types including increased gene expression in colorectal tumors for EpCAM, CDH3, and DSG2. We further analyzed gene co-expression analysis of EpCAM and each of the other enriched genes across individual primary tumor colon adenocarcinoma samples from TCGA and healthy colon and large intestine samples from GTEX and observed high co-expression with EpCAM in most of the identified hits, with the exception of NUP155, CD109 and DCBLD2 (Figure 6G). Of note, co-expression of EPCAM with CDH3 appears to be highly specific to primary colon adenocarcinoma tumors over normal tissue, highlighting the potential utility of co-targeting these two genes for increased tumor specificity in colon adenocarcinomas.

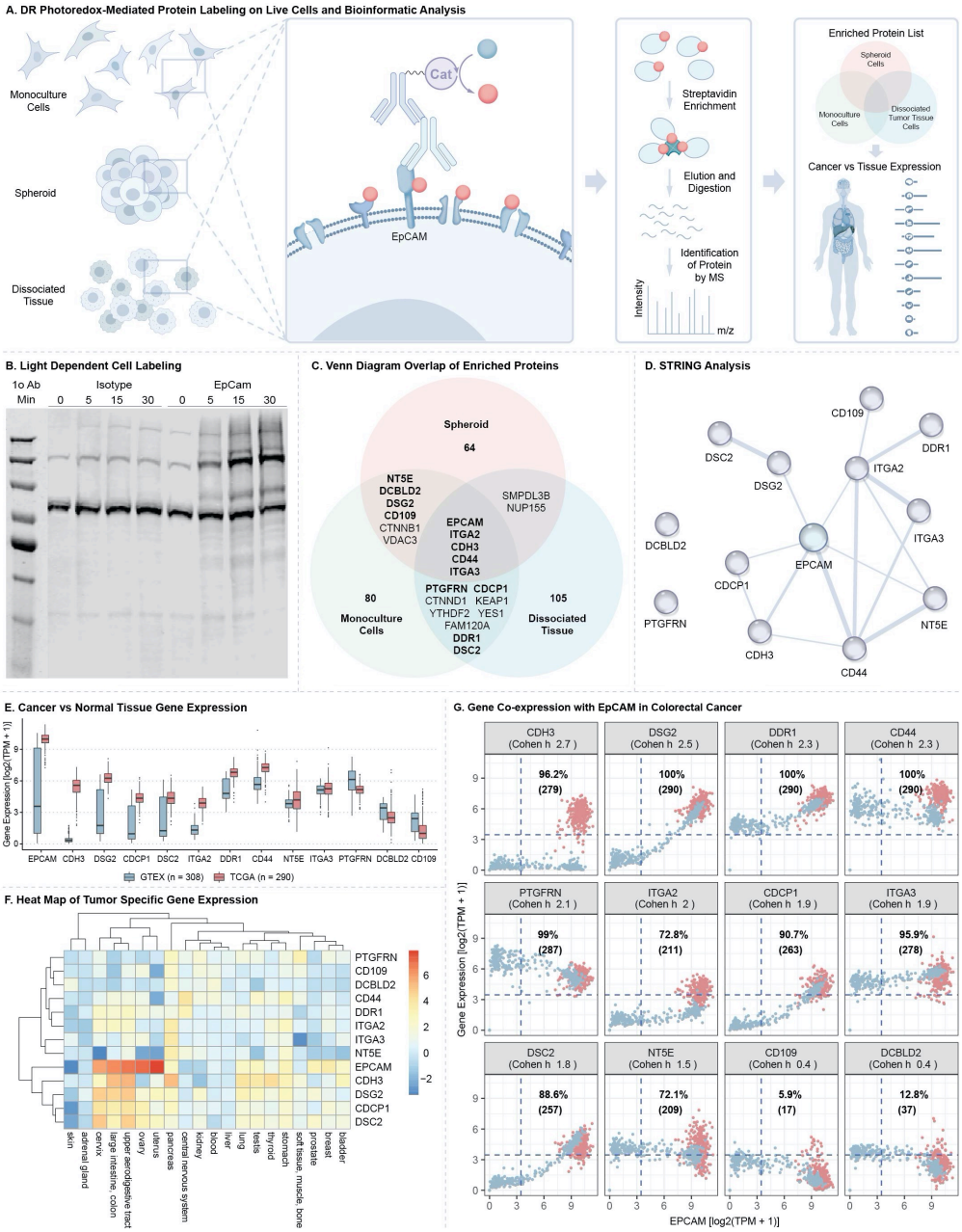

Figure 6. DR light-mediated photocatalytic labeling of cellular environments. A) Schematic depicting antibody-mediated proximity labeling of EpCAM on the cell surface of different HCT116 cell systems using the osmium photocatalyst system followed by mass spectrometry-based proteomic analysis to identify protein microenvironments for downstream bioinformatic analysis. B) Western blot analysis of EpCAM targeted labeling on the surface of HCT116 cells with an α-EpCAM or isotype primary antibody and a secondary antibody photocatalyst conjugate. Increased protein biotinylation is detected with increased DR-light exposure in the presence of the α-EpCAM antibody. C) Venn diagram analysis comparing statistically enriched surface proteins identified by proximity labeling in the three different cell systems (total number of enriched proteins in each system is indicated in the circle). Proteins with known surface

expression[37] are highlighted in bold. D) STRING protein interaction network analysis of significantly enriched membrane proteins from EpCAM-targeted labeling in different HCT116 cell systems. Line thickness between nodes correlates with experimental evidence of interactions from the STRING database. E) Tumor (TCGA) vs Normal (GTEx) log-fold change of the genes enriched in two or more cell systems highlighted in panel C. F) Heatmap showing log_2 fold-change between tumor vs normal gene expression across the indicated tumor types. G) Scatterplots showing co-expression between EpCAM and the indicated gene in primary tumor colon adenocarcinoma samples from TCGA (red, n = 290) and healthy colon and large intestine samples from GTEX (blue, n = 308). Co-expression is denoted as % tumor samples that expressed both genes at TPM > 10, indicating medium/high expression level (blue, dashed line). The effect size of tumor versus normal co-expression difference is provided as Cohen h for each gene pair.

Conclusion

In summary, we developed a novel photoredox strategy to generate high-energy nitrenes via low-energy light that proceeds through a redox-neutral electron transfer process. The resulting azide activation yields only triplet nitrene and operates solely when osmium photocatalyst and red light are used. These prerequisites enable the use of our photocatalytic system for targeted labeling applications wherein a photocatalyst can be directed to a cellular region of interest for localized activation of aryl azide probes. This strategy has particularly important implications for profiling biological environments with high spatiotemporal resolution. Accordingly, we applied this photocatalytic system for the targeted labeling of EpCAM on the surface of live colorectal cancer cells to identify not only EpCAM, but a small subset of neighboring surface proteins. The successful utilization of this platform technology opens the possibility to probe other membrane proteins of interest for the development of cell surface-focused therapeutics and diagnostics. We further anticipate that this mode of triplet nitrene activation will uncover new chemical reactivity and invite new applications for synthesis and

chemical biology. Accordingly, efforts to further develop photocatalytic-based bioorthogonal activation modes that make use of low-energy light in tissue environments are underway.

ORCID

N.E.S.T. 0000-0002-4063-8880; K.A.R. 0000-0001-6244-0878; J.L.W. 0000-0002-4937-9651; A. K. O. 0000-0002-0113-8306; D.R.R. 0000-0002-5265-5637; R.C.O 0000-0002-2920-9532; O.O.F. 0000-0002-5525-1304; T.R. 0000-0001-6287-8669

Acknowledgements

Research was supported by the NIH National Institute of General Medical Sciences (R01-GM125206) and gifts from Merck & Co., Inc., Kenilworth, New Jersey, USA. J.L.W was funded in part by the Columbia Center for Computational Electrochemistry (CCCE). This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562[38]. In particular, we used San Diego Computing Center's Expanse resources under allocation ID COL151.

Author contributions

N.E.S.T., K.A.R., R.C.O., O.O.F., and T.R. conceived of the work. N.E.S.T., K.A.R., A.K.O., R.C.O., O.O.F., and T.R. designed and executed experiments. N.E.S.T., K.A.R., R.C.O., O.O.F., and T.R. interpreted results. J.L.W and D.R.R. designed and executed chemistry-based computations. J.L.W. and D.R.R. interpreted chemistry-based computational results. N.E.S.T., K.A.R., R.C.O., O.O.F., and T.R. wrote the manuscript with input from all authors.

Author Information

K.A.R., A.K.O., R.C.O., and O.O.F. are employees of Merck Sharp & Dohme Corp., a subsidiary of Merck & Co., Inc., Kenilworth, NJ, USA. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to R.C.O (rob.oslund@merck.com), O.O.F (olugbeminiyi.fadeyi@merck.com), and T.R. (tr2504@columbia.edu). [4]Present address: Department of Chemistry, Princeton University, Princeton, New Jersey 08544, United States

Data availability

All relevant data are included in the manuscript and supplementary information.

References

[1] Kotzyba-Hibert, F., Kapfer, I. & Goeldner, M. Recent Trends in Photoaffinity Labeling. *Angew. Chem. Int. Ed.* **34**, 1296–1312 (1995).
[2] Murale, D. P., Hong, S. C., Haque, Md. M. & Lee, J.-S. Photo-affinity labeling (PAL) in chemical proteomics: a handy tool to investigate protein-protein interactions (PPIs). *Proteome Science* **15**, 14 (2017).
[3] Sadhu, K. K., Eierhoff, T., Römer, W. & Winssinger, N. Photoreductive Uncaging of Fluorophore in Response to Protein Oligomers by Templated Reaction in Vitro and in Cellulo. *J. Am. Chem. Soc.* **134**, 20013–20016 (2012).
[4] Geng, J. *et al.* Switching on prodrugs using radiotherapy. *Nat. Chem.* **13**, 805–810 (2021).
[5] Holland, J. P., Gut, M., Klingler, S., Fay, R. & Guillou, A. Photochemical Reactions in the Synthesis of Protein–Drug Conjugates. *Chem. Eur. J.* **26**, 33–48 (2020).
[6] Fantoni, N. Z., El-Sagheer, A. H. & Brown, T. A Hitchhiker's Guide to Click-Chemistry with Nucleic Acids. *Chem. Rev.* **121**, 7122–7154 (2021).
[7] Deb, T., Tu, J. & Franzini, R. M. Mechanisms and Substituent Effects of Metal-Free Bioorthogonal Reactions. *Chem. Rev.* **121**, 6850–6914 (2021).
[8] Heiss, T. K., Dorn, R. S. & Prescher, J. A. Bioorthogonal Reactions of Triarylphosphines and Related Analogues. *Chem. Rev.* **121**, 6802–6849 (2021).
[9] Chin, J. W. *et al.* Addition of *p*-Azido-L-phenylalanine to the Genetic Code of Escherichia coli. *J. Am. Chem. Soc.* **124**, 9026–9027 (2002).
[10] N. Kotani, J. Gu, T. Isaji, K. Udaka, N. Taniguchi, K. Honke, Biochemical visualization of cell surface molecular clustering in living cells. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 7405–7409 (2008).

[11] N. P. Gritsan, M. S. Platz, Kinetics, Spectroscopy, and Computational Chemistry of Arylnitrenes. *Chem. Rev.* **106**, 3844–3867 (2006).

[12] B. A. DeGraff, D. W. Gillespie, R. J. Sundberg, Phenyl nitrene. Flash photolytic investigation of the reaction with secondary amines. *J. Am. Chem. Soc.* **96**, 7491–7496 (1974).

[13] Li, Y. Zhuo., Kirby, J. P., George, M. W., Poliakoff, Martyn. & Schuster, G. B. 1,2-Didehydroazepines from the photolysis of substituted aryl azides: analysis of their chemical and physical properties by time-resolved spectroscopic methods. *J. Am. Chem. Soc.* **110**, 8092–8098 (1988).

[14] A. Marcinek, M. S. Platz, Deduction of the activation parameters for ring expansion and intersystem crossing in fluorinated singlet phenylnitrenes. *J. Phys. Chem.* **97**, 12674–12677 (1993).

[15] L.-H. Liu, M. Yan, Perfluorophenyl Azides: New Applications in Surface Functionalization and Nanomaterial Synthesis. *Acc. Chem. Res.* **43**, 1434–1443 (2010).

[16] S. Xie, M. Sundhoro, K. N. Houk, M. Yan, Electrophilic Azides for Materials Synthesis and Chemical Biology. *Acc. Chem. Res.* **53**, 937–948 (2020).

[17] R. Poe, K. Schnapp, M. J. T. Young, J. Grayzar, M. S. Platz, Chemistry and kinetics of singlet pentafluorophenylnitrene. *J. Am. Chem. Soc.* **114**, 5054–5067 (1992).

[18] N. Soundararajan, M. S. Platz, Descriptive photochemistry of polyfluorinated azide derivatives of methyl benzoate. *J. Org. Chem.* **55**, 2034–2044 (1990).

[19] A. Marcinek, M. S. Platz, S. Y. Chan, R. Floresca, K. Rajagopalan, M. Golinski, D. Watt, Unusually long lifetimes of the singlet nitrenes derived from 4-azido-2,3,5,6-tetrafluorobenzamides. *J. Phys. Chem.* **98**, 412–419 (1994).

[20] Schnapp, K. A., Poe, R., Leyva, E., Soundararajan, N. & Platz, M. S. Exploratory photochemistry of fluorinated aryl azides. Implications for the design of photoaffinity labeling reagents. *Bioconjugate Chem.* **4**, 172–177 (1993).

[21] Keana, J. F. W. & Cai, S. X. New reagents for photoaffinity labeling: synthesis and photolysis of functionalized perfluorophenyl azides. *J. Org. Chem.* **55**, 3640–3647 (1990).

[22] K. Kanakarajan, Raymond. Goodrich, M. J. T. Young, Soundara. Soundararajan, M. S. Platz, EPR spectroscopy of triplet aryl nitrenes covalently bound to α-chymotrypsin. Application of low-temperature methods to photoaffinity labeling. *J. Am. Chem. Soc.* **110**, 6536–6541 (1988).

[23] Ryu, K. A., Kaszuba, C. M., Bissonnette, N. B., Oslund, R. C. & Fadeyi, O. O. Interrogating biological systems using visible-light-powered catalysis. *Nat Rev Chem* **5**, 322–337 (2021).

[24] H. Huang, S. Banerjee, K. Qiu, P. Zhang, O. Blacque, T. Malcomson, M. J. Paterson, G. J. Clarkson, M. Staniforth, V. G. Stavros, G. Gasser, H. Chao, P. J. Sadler, Targeted photoredox catalysis in cancer cells. *Nat. Chem*. **11**, 1041–1048 (2019).

[25] J. B. Geri, J. V. Oakley, T. Reyes-Robles, T. Wang, S. J. McCarver, C. H. White, F. P. Rodriguez-Rivera, D. L. Parker, E. C. Hett, O. O. Fadeyi, R. C. Oslund, D. W. C. MacMillan, Microenvironment mapping via Dexter energy transfer on immune cells. *Science*. **367**, 1091–1097 (2020).

[26] Wright, M. H.; Sieber, S. A. Chemical Proteomics Approaches for Identifying the Cellular Targets of Natural Products. *Nat. Prod. Rep.* **33**, 681–708 (2016).

[27] Liu, J., Hadad, C. M. & Platz, M. S. The Reaction of Triplet Nitrenes with Oxygen: A Computational Study. *Org. Lett.* **7**, 549–552 (2005).

[28] Y. Chen, A. S. Kamlet, J. B. Steinman, D. R. Liu, A biomolecule-compatible visible-light-induced azide reduction from a DNA-encoded reaction-discovery system. *Nat. Chem*. **3**, 146–153 (2011).

[29] E. P. Farney, T. P. Yoon, Visible-Light Sensitization of Vinyl Azides by Transition-Metal Photocatalysis. *Angew. Chem. Intl. Ed*. **53**, 793–797 (2014).

[30] Wang, H. *et al.* Selective Mitochondrial Protein Labeling Enabled by Biocompatible Photocatalytic Reactions inside Live Cells. *JACS Au* **1**, 1066–1075 (2021).

[31] Ravetz, B. D. *et al.* Development of a Platform for Near-Infrared Photoredox Catalysis. *ACS Cent. Sci.* **6**, 2053–2059 (2020).

[32] This qualitative observation was validated via a localized approximation of coupled cluster (DLPNO-CCSD(T_0)). See: Ásgeirsson, V. *et al.* Nudged Elastic Band Method for Molecular Reactions Using Energy-Weighted Springs Combined with Eigenvector Following. *J. Chem. Theory Comput.* **17**, 4929–4945 (2021).

[33] Bissonnette, N. B. *et al.* Design of a Multiuse Photoreactor To Enable Visible-Light Photocatalytic Chemical Transformations and Labeling in Live Cells. *ChemBioChem* **21**, 3555–3562 (2020).
[34] Keller, L., Werner, S. & Pantel, K. Biology and clinical relevance of EpCAM. *Cell Stress* **3**, 165–180 (2019).
[35] Macdonald, J. *et al.* EpCAM Immunotherapy versus Specific Targeted Delivery of Drugs. *Cancers* **10**, 19 (2018).
[36] Szklarczyk, D. *et al.* The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* **49**, D605–D612 (2021).
[37] Bausch-Fluck, D. *et al.* The in silico human surfaceome. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E10988–E10997 (2018).
[38] Towns, J. *et al.* XSEDE: Accelerating Scientific Discovery. *Computing in Science Engineering* **16**, 62–74 (2014).