



## UV-driven self-repair of cyclobutane pyrimidine dimers in RNA†

Sarah J. Crucilla,<sup>ab</sup> Dian Ding,<sup>cd</sup> Gabriella G. Lozano,<sup>a</sup> Jack W. Szostak,<sup>cdef</sup> Dimitar D. Sasselov<sup>ga</sup> and Corinna L. Kufner<sup>id\*<sup>a</sup></sup>

Cite this: *Chem. Commun.*, 2023, 59, 13603

Received 21st August 2023,  
Accepted 20th October 2023

DOI: 10.1039/d3cc04013e

rsc.li/chemcomm

Nucleic acids can be damaged by ultraviolet (UV) irradiation, forming structural photolesions such as cyclobutane-pyrimidine-dimers (CPD). In modern organisms, sophisticated enzymes repair CPD lesions in DNA, but to our knowledge, no RNA-specific enzymes exist for CPD repair. Here, we show for the first time that RNA can protect itself from photolesions by an intrinsic UV-induced self-repair mechanism. This mechanism, prior to this study, has exclusively been observed in DNA and is based on charge transfer from CPD-adjacent bases. In a comparative study, we determined the quantum yields of the self-repair of the CPD-containing RNA sequence, GAU = U to GAUU (0.23%), and DNA sequence, d(GAT = T) to d(GATT) (0.44%), upon 285 nm irradiation via UV/Vis spectroscopy and HPLC analysis. After several hours of irradiation, a maximum conversion yield of ~16% for GAU = U and ~33% for d(GAT = T) was reached. We examined the dynamics of the intermediate charge transfer (CT) state responsible for the self-repair with ultrafast UV pump – IR probe spectroscopy. In the dinucleotides GA and d(GA), we found comparable quantum yields of the CT state of ~50% and lifetimes on the order of several hundred picoseconds. Charge transfer in RNA strands might lead to reactions currently not considered in RNA photochemistry and may help understanding RNA damage formation and repair in modern organisms and viruses. On the UV-rich surface of the early Earth, these self-stabilizing

mechanisms likely affected the selection of the earliest nucleotide sequences from which the first organisms may have developed.

The integrity of genetic information is vital for all organisms on Earth. When exposed to ultraviolet (UV) light, nucleic acids can form photolesions, such as cyclobutane pyrimidine dimers (CPDs),<sup>1–4</sup> which can impede their function, cause mutations, or – in severe cases – lead to cell death.<sup>5,6</sup> In modern organisms, enzymes,<sup>7–9</sup> such as photolyases, repair such damage, *e.g. via* charge transfer (CT) of a photoelectron to the lesion. On the early Earth, an intramolecular process, UV-induced self-repair (Fig. 1A), is a likely precursor to enzymatic base repair.<sup>10,11</sup> Its mechanism resembles that of photolyase with regard to the CT.<sup>12</sup> In contrast to the enzymatic repair, the photoinduced electron transfer occurs intramolecularly from adjacent bases to the lesion.<sup>10–12</sup> For example, in a DNA sequence, d(GAT = T), with “=” representing the CPD lesion (Fig. 1A), a selective UV excitation of the G (and A) leads to the formation of a CT state.<sup>13,14</sup> From this long-lived CT state, a charge is then transferred to the d(T = T), which results in the cleavage of the dimer and a return of the CPD to undamaged d(TT).<sup>10,12</sup> The lifetime of the adjacent CT state directly affects the efficiency of the self-repair.<sup>10,14</sup>

This mechanism has been studied in-depth for DNA.<sup>10–12,18,19</sup> However, to our knowledge, no studies have been conducted in RNA, which could have preceded<sup>20</sup> or evolved simultaneously<sup>21,22</sup> to DNA. Here, we provide the first experimental evidence for UV-induced self-repair in RNA. We show the recovery of the short single-stranded RNA sequence GAUU from GAU = U via UV/Vis spectroscopy and high-performance liquid chromatography (HPLC). Complementary ultrafast UV-pump IR-probe experiments of the RNA sequence GA give insights into the underlying mechanism.

Our findings contribute to an understanding of the selection pressures acting on the first RNA sequences available on early Earth under exposure to sunlight. This sequence selectivity is likely to have affected RNA functions in prebiotic chemistry or as information carriers in translation or replication.

<sup>a</sup> Harvard-Smithsonian Center for Astrophysics, Harvard University, 60 Garden Street, Cambridge, MA 02138, USA.  
E-mail: corinna.kufner@cfa.harvard.edu

<sup>b</sup> Department of Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts 02138, USA

<sup>c</sup> Howard Hughes Medical Institute, Department of Molecular Biology and Center for Computational and Integrative Biology, Massachusetts General Hospital, Boston, Massachusetts 02114, USA

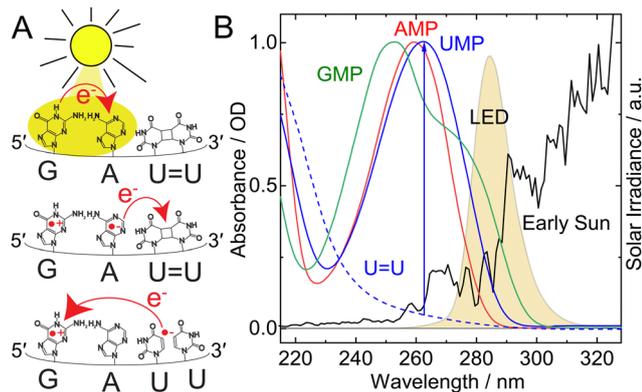
<sup>d</sup> Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, USA

<sup>e</sup> Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA

<sup>f</sup> Howard Hughes Medical Institute, Department of Chemistry, The University of Chicago, Chicago, Illinois 60637, USA

† Electronic supplementary information (ESI) available. See DOI: <https://doi.org/10.1039/d3cc04013e>





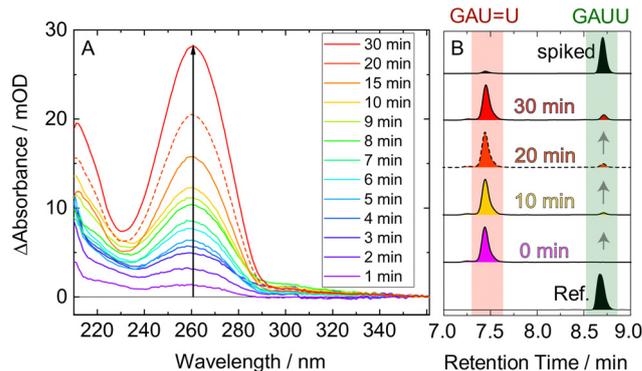
**Fig. 1** (A) Mechanism of UV-induced self-repair by electron transfer from an adjacent charge-transfer state. (B) UV spectra of the RNA monophosphate nucleotides AMP (red, 0.067 mM), UMP (blue solid, 0.075 mM), GMP (green, 0.072 mM) and the U = U photolesion (blue dashed, ~0.1 mM) in 50 mM phosphate buffer at pH 6.9 (concentrations determined from<sup>15</sup>). An LED with a central wavelength of 285 nm (orange) selectively excites the bases G and A in the damaged sequence GAU = U and triggers self-repair to the intact tetramer GAUU restoring the uridine absorption band (blue arrow). The solar spectrum on the surface of the early Earth is shown for comparison (black).<sup>16,17</sup>

In modern cells, RNA damage by UV light is less well studied and understood than DNA damage.<sup>23–26</sup> There are no known RNA-specific photolyases for CPD repair and cases of inferred RNA dimer photoreversal in plants and insects are rare.<sup>19,23,27–31</sup> Our results may be relevant to understanding how cells handle UV-damaged RNA with mechanisms very different from DNA repair activity.

Irradiation with an LED centered around 285 nm, 15 nm bandwidth (spectrum Fig. 1B, orange, setup Fig. S1, ESI<sup>†</sup>) allowed us to mimic prebiotically abundant (black)<sup>16,17,32</sup> long wavelength irradiation conditions with regard to the absorption of the nucleobases (green, red, blue). We exposed the RNA, GAU = U, and DNA single strand, GAT = T (see Fig. S2, S3 and supplementary data for the synthesis and characterization, ESI<sup>†</sup>), to 285 nm irradiation for increasing times. The 285 nm irradiation mainly excited the canonical bases, G (Fig. 1B, green) and A (red), due to the negligible absorbance of the CPD lesions in this wavelength range (blue dashed).<sup>10,33</sup> Upon self-repair of the CPD, we see recovery of the 263 nm absorbance band of the dipyrimidine, UU (Fig. 1B, blue arrow), and the 266 nm absorbance of d(TT), respectively.<sup>10</sup> Fig. 2A shows the UV absorbance difference spectra of the irradiated GAU = U minus the unirradiated reference. Increasing irradiation time led to an increased recovery of the 263 nm absorbance, indicating the self-repair of the RNA oligomer.

In a complementary approach, we performed analytical HPLC analysis of the unirradiated and irradiated samples. Fig. 2B shows the analytical HPLC chromatograms of the sequence GAU = U at increasing times of 285 nm irradiation at an average power of 0.4 mW (see Fig. S4 for the corresponding chromatograms of d(GAT = T), ESI<sup>†</sup>).

The purple plot in Fig. 2B displays the HPLC analysis of the unirradiated damaged sequence GAU = U with a peak at a retention time of 7.4 min. Upon irradiation, the undamaged peak recovered



**Fig. 2** (A) UV/Vis absorbance change spectra of the sequence GAU = U upon irradiation at 285 nm. Due to self-repair, the 263 nm absorbance of the sample increases. (B) Analytical HPLC analysis of the GAU = U self-repair. The purple plot shows the unirradiated (0 min) starting material, GAU = U, which recovers during irradiation. The undamaged sequence, GAUU, is shown in black at the bottom for reference. The top black plot shows the 30 min sample spiked with undamaged GAUU. The full range chromatograms from 0 min to 15 min are shown in Fig. S5 (ESI<sup>†</sup>).

gradually. The chromatogram of the undamaged GAUU, with a peak at 8.7 min, is added at the bottom as a reference. Its shape and location match the repair peak. After 30 min of irradiation the sample was spiked with undamaged GAUU, which strengthened the signal at 8.7 min (top black plot, scaled down for comparison). The spiked peak matches the peak of the recovered sequence from the irradiation. The HPLC analysis shows a clear recovery of the intact sequence GAUU from GAU = U, as a result of the absorption of increasing doses of 285 nm photons. These findings confirm the results obtained by UV spectroscopy (Fig. 2A).

The quantum yield of a photochemical process is a measure of the efficiency of a photoinduced reaction. It is defined as the number of product molecules, which here is intact GAUU, divided by the number of photons absorbed by the starting material, here GAU = U.<sup>34,35</sup> The number of product molecules, GAUU can be calculated from the absorbance change at 263 nm in Fig. 2A and the number of absorbed photons from the photometric setup (see ESI<sup>†</sup> for details on the quantum yield determination). The absorbance change at 263 nm (arrow) is plotted as a function of the absorbed dose for the oligonucleotide GAU = U in Fig. S6 (ESI<sup>†</sup>). The positive sign of the slope is indicative of the self-repair of the sequence. The slope is linearly proportional to the quantum yield of the self-repair. Table S1 (ESI<sup>†</sup>) lists the slopes for both oligonucleotides determined from linear fits (Fig. S6, red, ESI<sup>†</sup>) and the respective quantum yields. The quantum yield for the self-repair of the RNA sequence, GAU = U, is ~0.23% and for the DNA sequence, d(GAT = T), ~0.44%. In both oligonucleotides, the G and A are the main absorbers of the 285 nm photons and exceed the weak direct absorption of T = T and U = U, which can lead to the direct photoreversal of the CPD.<sup>36–38</sup> Previously, the photoreversal of the DNA oligonucleotide, GAT = T, has been reported to be ~0.25% at a narrow-band 290 nm excitation (Δλ = 3 nm, FWHM).<sup>10</sup> The higher quantum yield of ~0.44% in this study can be attributed to the lower excitation wavelength of 285 nm at a larger (Δλ = 15 nm, FWHM) bandwidth.



From a photophysical perspective, the quantum yield characterizes the efficiency of the repair process. On early Earth, the maximum total yield of the repair is most significant for the preferential survival of certain sequences over others. We therefore conducted long-term irradiations over several hours to reach the photostationary equilibrium of the oligonucleotides, where the rate of damage is equal to the rate of self-repair. The equilibrium was reached after conversion of  $\sim 33\%$  d(GAT = T) to d(GATT), and after  $\sim 16\%$  of GAU = U to GAUU (Fig. S7, ESI<sup>†</sup>). Consequently, in the absence of repair enzymes, a significant fraction of both intact sequences will be recovered by UV-induced self-repair.

In the DNA sequence, d(GAT = T), excitation at 290 nm has been shown to promote sequence selective self-repair *via* a photolyase-like mechanism:<sup>10</sup> Selective excitation of the dG (and dA) leads to the formation of a charge transfer (CT) state, d(G<sup>•+</sup>A<sup>•-</sup>), with a lifetime of several hundred picoseconds.<sup>13,14</sup> From the long-lived CT state, a charge is then transferred to the T = T, which results in the self-repair.<sup>10,12</sup> The lifetime of the adjacent CT state directly affects the efficiency of the self-repair.<sup>10,14</sup> To understand the underlying mechanism of RNA self-repair, we used ultrafast 260 nm pump, mid-IR probe spectroscopy to study the lifetimes and quantum yields of the transient CT states of the RNA dinucleotide, GA, and DNA dinucleotide, d(GA). Fig. 3 shows the transient absorption change spectra in the mid-infrared range of GA as a function of delay time (see Fig. S8 for d(GA) data, ESI<sup>†</sup>). The negative (blue) bands at 1574 cm<sup>-1</sup> and 1667 cm<sup>-1</sup> correspond to the ground state bleach (GSB) of G and the band at 1619 cm<sup>-1</sup> can be assigned to the GSB of A.<sup>13</sup> The positive (red) bands around 1593 cm<sup>-1</sup> and 1705 cm<sup>-1</sup> are indicative of the G<sup>•+</sup> radical

cation and between 1500 cm<sup>-1</sup> and 1550 cm<sup>-1</sup> for the A<sup>•-</sup> radical anion.<sup>10,13</sup> Global fitting analysis reveals a lifetime of the CT state of 390 ± 120 ps and a quantum yield of 58 ± 30% for GA. In the case of d(GA) the CT lifetime is 400 ± 130 ps (Fig. S8, ESI<sup>†</sup>) and the quantum yield is 60 ± 30%, which agrees with previous work.<sup>13,14</sup>

The similar quantum yields and lifetimes of the CT states point towards similar base stacking of the dinucleotides regardless of the sugar backbone. Despite the similarity of the adjacent CT states in GA and d(GA), the quantum yield of the self-repair of the RNA sequences, GAU = U, is about a factor of 2 lower than of the DNA sequence, d(GAT = T). This difference may originate from the photoreversibility of d(T = T) compared to U = U and from the structure of the backbone. Previously, dT and dU containing CPD lesions have been reported to have very similar geometrical structures.<sup>39</sup> The conformational freedom on the sugar moiety of the *cis-syn* CPD lesions<sup>39</sup> and differences in the flexibility of RNA and DNA backbones<sup>40</sup> could lead to different stacking which could affect the self-repair quantum yields.

In a surface environment on early Earth, UV-driven self-repair likely coincided with the formation of various photolesions. Among these, CPDs and photohydrates are the most frequent RNA lesions with quantum yields on the 0.5–2% level.<sup>4,41–43</sup> Self-repair on a similar order of magnitude (a factor of  $\sim 4$  lower for the sequence GAU = U) may have protected sequences from photolysis. Further studies are required to determine the sequence and wavelength dependence, and the influence of the molecular conformation on the self-repair efficiency. However, our findings indicate that self-repair could be an underlying mechanism for the selection of the RNA sequences that survive exposure to sunlight.

We have shown that UV light induces self-repair in the RNA sequence, GAU = U, with a similar mechanism to self-repair in the DNA sequence d(GAT = T). Our UV/Vis spectroscopic data and HPLC analysis show a clear recovery of the intact sequence GAUU following 285 nm irradiation. Ultrafast UV-pump, mid-IR-probe spectroscopy revealed that the repair is initiated by electron transfer from a long-lived CT state. The photoreversibility of the lesions and the backbone conformation may contribute to the efficiency of the self-repair. In the absence of enzymatic repair, under continuous exposure to UV irradiation, a significant fraction of the intact sequence can be recovered by self-repair (up to  $\sim 16\%$  of GAUU from GAU = U). Therefore, self-repair might be important for the photostability of RNA, both on early Earth and in modern organisms. On early Earth, self-repair could have affected the set of RNA sequences that survived exposure to sunlight.

In modern cells, CPD lesions in DNA play a central role in UV-induced carcinogenesis.<sup>44</sup> Average prokaryotic and eukaryotic cells contain 4–6 times more RNA than DNA and significant UV-damage occurs to RNA as well.<sup>23,45</sup> While the higher RNA turnover in cells is thought to ameliorate the deleterious effects of RNA photolesions, it is notable that accumulation of damaged RNA has been observed in aging and disease, indicating that turnover alone is not sufficient.<sup>23</sup> Our study lays the groundwork for understanding

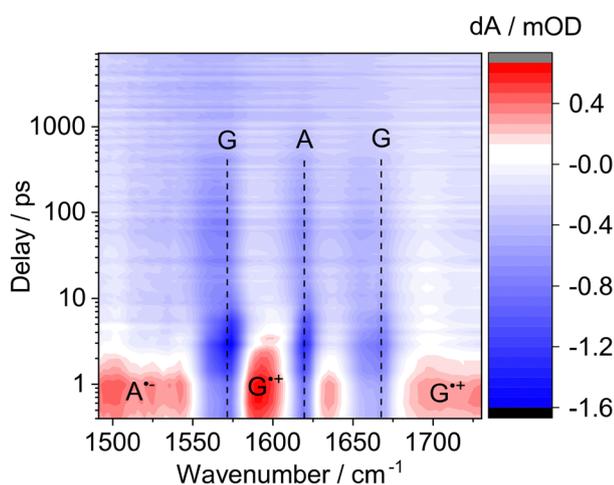


Fig. 3 Transient mid-infrared absorption difference spectra of the RNA dinucleotide GA, excited at 260 nm. The negative (blue) bands can be assigned to the bleaching (*i.e.* recovery after initial photodepletion) of the ground state of G ( $\sim 1574$  cm<sup>-1</sup> and  $\sim 1667$  cm<sup>-1</sup>) and A (1619 cm<sup>-1</sup>). The positive bands (red) indicate the presence of a long-lived CT state on the order of 390 ps, which is long enough to initiate self-repair: The bands around 1593 cm<sup>-1</sup> and 1705 cm<sup>-1</sup> are characteristic of the G<sup>•+</sup> radical cation and the broad signature between 1500 cm<sup>-1</sup> and 1550 cm<sup>-1</sup> of the A<sup>•-</sup> radical anion.<sup>10,13,14</sup>



the protection of RNA from UV-induced damage. Future studies with additional sequences will be required to assess the generality of the self-repair process in RNA.

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

This work was supported by the Simons Foundation (SCOL 290360FY18 to D.D.S., SCOL 290363 to J. W. S) and by the National Science Foundation (Grant No. 1933505 to D. D. S., DGE 2140743 to S. J. C.). The authors thank Fanny Ng for assistance conducting the experiments, Dr Zoe R. Todd for feedback on the manuscript. The authors are grateful to Dr Sandra L. Wolin, Dr Stewart H. Shuman, and Dr Aziz Sancar for helpful discussions on RNA repair enzymes. They acknowledge the Harvard Origin of Life Initiative.

## Conflicts of interest

There are no conflicts to declare.

## References

- 1 A. Rörsch, R. Beukers, J. Ijlstra and W. Berends, *Recl. Trav. Chim. Pays-Bas*, 1958, **77**, 423–429.
- 2 A. Downes and T. P. Blunt, *Proc. R. Soc. London*, 1877, **26**, 488–500.
- 3 H. Johns, S. Rapaport and M. Delbrück, *J. Mol. Biol.*, 1962, **4**, 104–114.
- 4 P. Swenson and R. Setlow, *Photochem. Photobiol.*, 1963, **2**, 419–434.
- 5 N. Buron, O. Micheau, S. Cathelin, P.-O. Lafontaine, C. Creuzot-Garcher and E. Solary, *Invest. Ophthalmol. Visual Sci.*, 2006, **47**, 4221–4230.
- 6 I. Tessman, M. A. Kennedy and S.-K. Liu, *J. Mol. Biol.*, 1994, **235**, 807–812.
- 7 H. Harm and C. S. Rupert, *Mutat. Res., Fundam. Mol. Mech. Mutagen.*, 1970, **10**, 291–306.
- 8 R. Setlow and W. Carrier, *Proc. Natl. Acad. Sci. U. S. A.*, 1964, **51**, 226.
- 9 G. B. Sancar, F. W. Smith, R. Reid, G. Payne, M. Levy and A. Sancar, *J. Biol. Chem.*, 1987, **262**, 478–485.
- 10 D. B. Bucher, C. L. Kufner, A. Schlueter, T. Carell and W. Zinth, *J. Am. Chem. Soc.*, 2016, **138**, 186–190.
- 11 R. Szabla, M. Zdrowowicz, P. Spisz, N. J. Green, P. Stadlbauer, H. Kruse, J. Šponer and J. Rak, *Nat. Commun.*, 2021, **12**, 3018.
- 12 R. Szabla, H. Kruse, P. Stadlbauer, J. Šponer and A. L. Sobolewski, *Chem. Sci.*, 2018, **9**, 3131–3140.
- 13 D. B. Bucher, B. M. Pillis, T. Carell and W. Zinth, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 4369–4374.
- 14 C. L. Kufner, W. Zinth and D. B. Bucher, *ChemBioChem*, 2020, **21**, 1–6.
- 15 M. J. Cavaluzzi and P. N. Borer, *Nucleic Acids Res.*, 2004, **32**, e13.
- 16 S. Ranjan and D. D. Sasselov, *Astrobiology*, 2017, **17**, 169–204.
- 17 S. Rugheimer, A. Segura, L. Kaltenegger and D. Sasselov, *Astrophys. J.*, 2015, **806**, 137.
- 18 Z. Pan, J. Chen, W. J. Schreier, B. Kohler and F. D. Lewis, *J. Phys. Chem. B*, 2011, **116**, 698–704.
- 19 D. J.-F. Chinnapen and D. Sen, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 65–69.
- 20 L. E. Orgel, *Crit. Rev. Biochem. Mol. Biol.*, 2004, **39**, 99–123.
- 21 K. Le Vay and H. Mutschler, *Emerging Top. Life Sci.*, 2019, **3**, 469–475.
- 22 S. Bhowmik and R. Krishnamurthy, *Nat. Chem.*, 2019, **11**, 1009–1018.
- 23 E. J. Wurtmann and S. L. Wolin, *Crit. Rev. Biochem. Mol. Biol.*, 2009, **44**, 34–49.
- 24 B. Singer, *Virology*, 1971, **45**, 101–107.
- 25 P. Harriman and H. Zachau, *J. Mol. Biol.*, 1966, **16**, 387–403.
- 26 M. S. Jordanov, D. Pribnow, J. L. Magun, T.-H. Dinh, J. A. Pearson and B. E. Magun, *J. Biol. Chem.*, 1998, **273**, 15794–15803.
- 27 H. Jäckle and K. Kalthoff, *Photochem. Photobiol.*, 1980, **32**, 749–761.
- 28 C. P. Selby and A. Sancar, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 17696–17700.
- 29 A. Sancar, *Chem. Rev.*, 2003, **103**, 2203–2238.
- 30 K. J. Hughes, X. Chen, A. M. Burroughs, L. Aravind and S. L. Wolin, *Cell Rep.*, 2020, **33**, 108527.
- 31 H. Werbin, R. C. Valentine and A. McLaren, *Photochem. Photobiol.*, 1967, **6**, 205–213.
- 32 I. Cnossen, J. Sanz-Forcada, F. Favata, O. Witasse, T. Zegers and N. F. Arnold, *J. Geophys. Res.: Planets*, 2007, **112**, E2.
- 33 W. J. Schreier, P. Gilch and W. Zinth, *Annu. Rev. Phys. Chem.*, 2015, **66**, 497–519.
- 34 I. U. o. P. a. A. Chemistry, Quantum Yield,  $\Phi$ , 2014.
- 35 C. L. Kufner, D. B. Bucher and D. D. Sasselov, *ChemSystemsChem*, 2022, e202200019.
- 36 M. R. Holman, T. Ito and S. E. Rokita, *J. Am. Chem. Soc.*, 2007, **129**, 6–7.
- 37 H. Johns, M. Pearson, J. LeBlanc and C. Helleiner, *J. Mol. Biol.*, 1964, **9**, 503-IN501.
- 38 F. Garces and C. Davila, *Photochem. Photobiol.*, 1982, **35**, 9–16.
- 39 T. M. Koning, J. J. van Soest and R. Kaptein, *Eur. J. Biochem.*, 1991, **195**, 29–40.
- 40 A. Noy, A. Perez, F. Lankas, F. J. Luque and M. Orozco, *J. Mol. Biol.*, 2004, **343**, 627–638.
- 41 H. Görner, *J. Photochem. Photobiol., B*, 1991, **10**, 91–110.
- 42 E. V. Khoroshilova and D. N. Nikogosyan, *J. Photochem. Photobiol., B*, 1990, **5**, 413–427.
- 43 I. Brown and H. Johns, *Photochem. Photobiol.*, 1968, **8**, 273–286.
- 44 J. Jans, W. Schul, Y.-G. Sert, Y. Rijkssen, H. Rebel, A. P. Eker, S. Nakajima, H. Van Steeg, F. R. De Gruijl and A. Yasui, *Curr. Biol.*, 2005, **15**, 105–115.
- 45 S. E. Beck, R. A. Rodriguez, M. A. Hawkins, T. M. Hargy, T. C. Larason and K. G. Linden, *Appl. Environ. Microbiol.*, 2016, **82**, 1468–1474.

