Emergence of light-driven protometabolism upon recruitment of a photocatalytic cofactor by a self-replicator

Authors: Guillermo Monreal Santiago[1], Kai Liu[1], Wesley R. Browne[1], Sijbren Otto[1*]

Affiliations:

[1]Centre for Systems Chemistry, Stratingh Institute, University of Groningen, Nijenborgh 4, 9747 AG Groningen, Netherlands.

[*]Corresponding author. Email: s.otto@rug.nl

ABSTRACT

Establishing how life can emerge from inanimate matter is among the grand challenges of contemporary science. Chemical systems that capture life's essential characteristics: replication, metabolism and compartmentalization offer a route to understanding this momentous step. The synthesis of life, whether based on canonical biomolecules or fully synthetic molecules, requires the functional integration of these characteristics. Here we show how a system of fully synthetic self-replicating molecules, upon recruiting a cofactor, acquires the ability to transform thiols in its environment into disulfide precursors from which the molecules can replicate. The replicator recruits a cofactor, enhances its activity in oxidizing thiols into disulfides through photoredox catalysis and thereby accelerates replication by increasing the availability of the disulfide precursors. This positive feedback marks the emergence of a light-driven protometabolism in a system that bears no resemblance to canonical biochemistry and constitutes a significant step towards the highly challenging aim of creating a new and completely synthetic form of life.

All currently known life forms, diverse as they are, share the same set of building blocks and chemical reactions. Consequently, most research on the origins of life is focused on how these building blocks can be synthesized under prebiotically plausible conditions (*1*), or on how they can be combined to create a synthetic cell (*2*). However, there is no evidence to suggest that life in the universe, or even early life on Earth, *needs* to have a biochemistry similar to that with which we are familiar (*3*). This suggests that, in addition to the current focus on the origin of the *chemical structures* of known life, research into the emergence of the *functional characteristics* of life is of equal importance, also using system that are chemically distinct from the canonical biomolecules (RNA, DNA, sugars and proteins). While defining life is remarkably difficult (*4*), the current paradigm is that life requires the functional integration of three fundamental characteristics: self-replication, metabolism and compartmentalization (*5*). Self-replication refers to the ability of a system to produce (not necessarily exact) copies of itself. Compartmentalization refers to the ability to maintain co-localization of the components of a system. Metabolism has been defined as "the active or passive entrance of material and energy into the system which transforms them by chemical processes into its own internal constituents […], so that the chemical reactions result in a regulated and controlled increase of the inner constituents as well as in the energy supply of the system." (*6*).

Efforts aimed at integrating the functional characteristics of life have predominantly focused on enclosing (mostly enzymatic) reaction networks or a (self-)replicating entity within a compartment (*7-10*). Although the integration of metabolism with self-replication has been explored theoretically (*11*,*12*), experimental implementation is lacking, even though metabolic self-replication is a necessary step to go from "simple" replicators to complex living entities (*13*) and critical in a resource-poor environment (*14*). A major hurdle in achieving metabolic self-replication is associated with identifying replicating systems that can play a dual role: aside from catalyzing their own formation they should also be able to catalyze other (metabolic) reactions. Few systems exist that can do so: Lehman reported RNA sets, where both self-replication and protometabolism were based on the same phosphodiester chemistry (*15*), and Rebek developed synthetic self-replicators capable or organocatalysis, where replication and catalysis required different solvents (*16*).

We set out to integrate some of the features of metabolism into a previously described self-assembly-driven self-replicator (*17-18*). This is an attractive system as it is currently the only

self-replicator that is able to emerge spontaneously from a complex mixture and capable of exponential growth at the same time. Briefly, the replicator used in this work ($\mathbf{1}_6$) is a hexameric macrocycle made from building block **1** (Fig. 1). When **1** is oxidized in a buffered (pH = 8.2) aqueous solution, it gives rise to a small dynamic combinatorial library of disulfide-based macrocycles of different sizes (mostly $\mathbf{1}_3$ and $\mathbf{1}_4$), from which $\mathbf{1}_6$ can emerge and self-assemble into fibers. Once those fibers reach a critical length, they enter a templated growth/mechanical breakage cycle. In this way, $\mathbf{1}_6$ can grow exponentially as long as the concentration of its precursors (the disulfide-based macrocycles $\mathbf{1}_3$/$\mathbf{1}_4$) remains high enough. However, if an additional oxidant is not used, the oxidation of **1** (occurring slowly through reaction with dissolved oxygen) can become the rate-limiting step of the replication process.

We designed a rudimentary metabolism for this system by giving $\mathbf{1}_6$ the ability to promote the oxidation reaction of $\mathbf{1} \rightarrow \mathbf{1}_3/\mathbf{1}_4$, using light as an external energy source. This fulfills one of the main functions of metabolism: The use of external energy to synthesize the precursors of the self-replicating species starting from inert materials. Our strategy was inspired by cofactor recruitment encountered in present biochemistry (*19-21*) and involves the self-replicator binding and activating a small organic photosensitizer. This strategy was implemented using two different cofactors (**2** or **3**) as photosensitizers.

RESULTS AND DISCUSSION

Replicator-induced J-aggregation of Rose Bengal serves an activation mechanism for its photocatalytic activity

The first photosensitizer that we selected as a cofactor was Rose Bengal (**2**), an organic dye known to photocatalyze the oxidation of thiols to disulfides via generation of $^1O_2$ (*22*). Its hydrophobicity and two negative charges make it a good candidate for interacting with the fibers of $\mathbf{1}_6$, that present large hydrophobic pockets (*23*) (Fig. S1) and feature lysine residues that are positively charged at ca. pH 8. Spectral analysis confirmed an interaction of **2** with the macrocycles derived from **1** (Fig. 2A), but not with **1** itself, as the two thiolate groups give **1** a net negative charge (Fig. S2). Notably, a new red-shifted band exclusively appeared when **2** and replicator $\mathbf{1}_6$ were mixed. This sharp band, which exhibits circular dichroism (Fig. 2B), is indicative of the binding of **2** to the $\mathbf{1}_6$ fibers in a chiral, supramolecular J-aggregate (*24, 25*). The band is not present when the $\mathbf{1}_6$:**2** ratio is high (Fig. S3) or in the presence of $\mathbf{1}_3$/$\mathbf{1}_4$.

The appearance of this band allowed us to design a system where only the **2**-$\mathbf{1}_6$ complex is able to act as a photooxidation catalyst. By irradiating the system at wavelengths corresponding to the new band only, we ensured that **2** would be excited only when J-aggregated. Excited **2** then undergoes intersystem crossing, promoting $^1O_2$ generation, which in turn mediates thiol oxidation (Fig. 3C). Indeed, upon photo irradiation, the oxidation of **1** was accelerated 2.4 fold when both **2** and replicator $\mathbf{1}_6$ were present, compared to control experiments using only **2** or **2** and $\mathbf{1}_3$/$\mathbf{1}_4$ (Fig. 3A). (*26*) Samples that were not irradiated oxidized at a similar rate to the sample without **2**, confirming that the reaction proceeded via $^1O_2$-mediated photooxidation. Thus, catalysis of the reaction $\mathbf{1} \rightarrow \mathbf{1}_3/\mathbf{1}_4$ is an emergent property of the $\mathbf{1}_6$-**2** complex and cannot be attributed to any of the individual species acting in isolation.

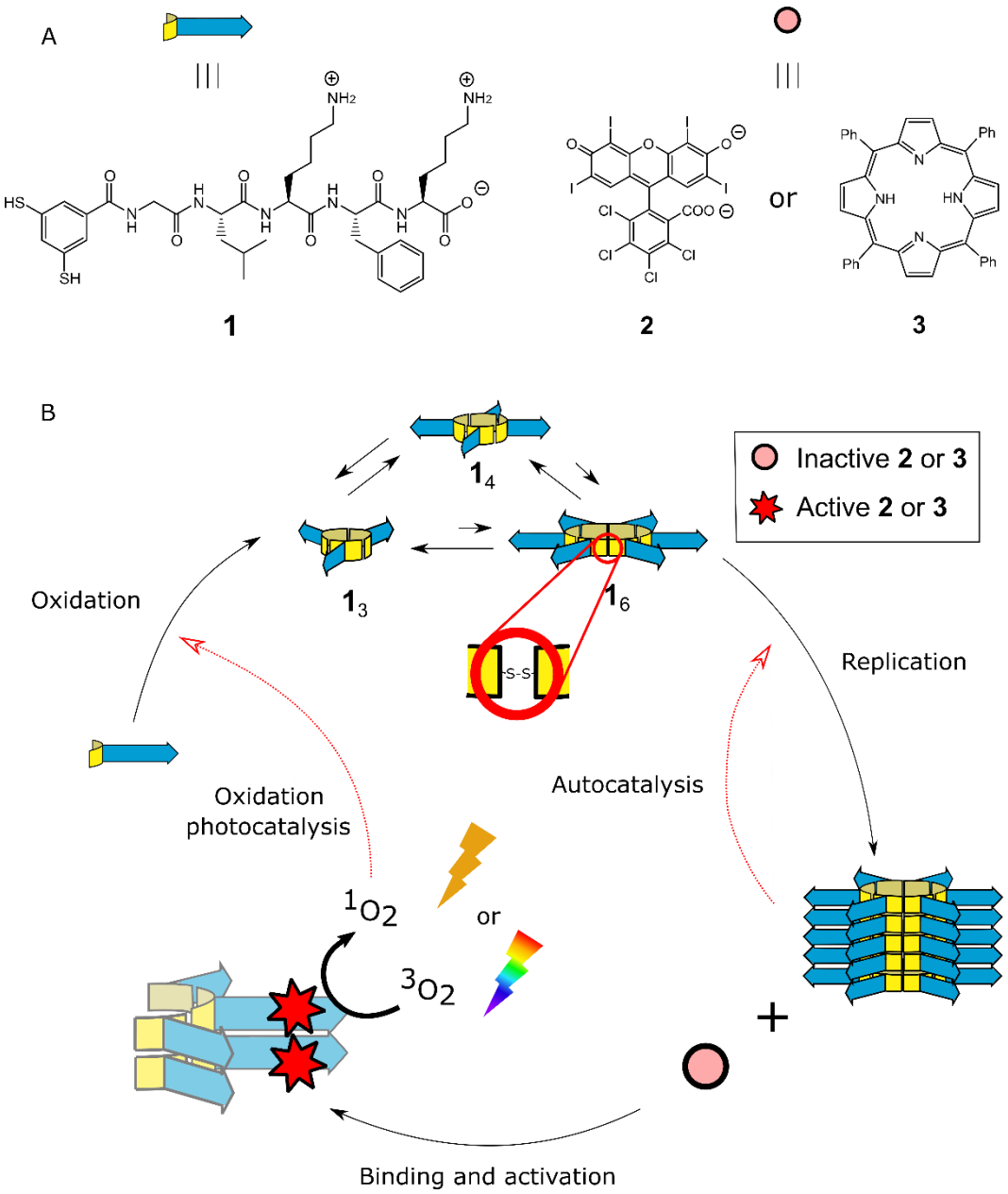

Fig. 1. Mechanism of protometabolic self-replication. ***(A)*** *Structures of building block* ***1****, Rose Bengal (****2****), and tetraphenylporphyrin (****3****).* ***(B)*** *The reaction network combining self-replication, cofactor recruitment, and* $^1O_2$*-mediated photo-oxidation. The oxidation of* ***1*** *leads to a DCL of macrocyclic disulfides, which provides the precursors for the replicator* $\mathbf{1}_6$*. The binding of* ***2*** *or* ***3*** *to the* $\mathbf{1}_6$ *fiber leads to emergent photocatalytic* $^1O_2$ *production upon photo-irradiation, which accelerates the oxidation process and leads to an autocatalytic feedback loop.*

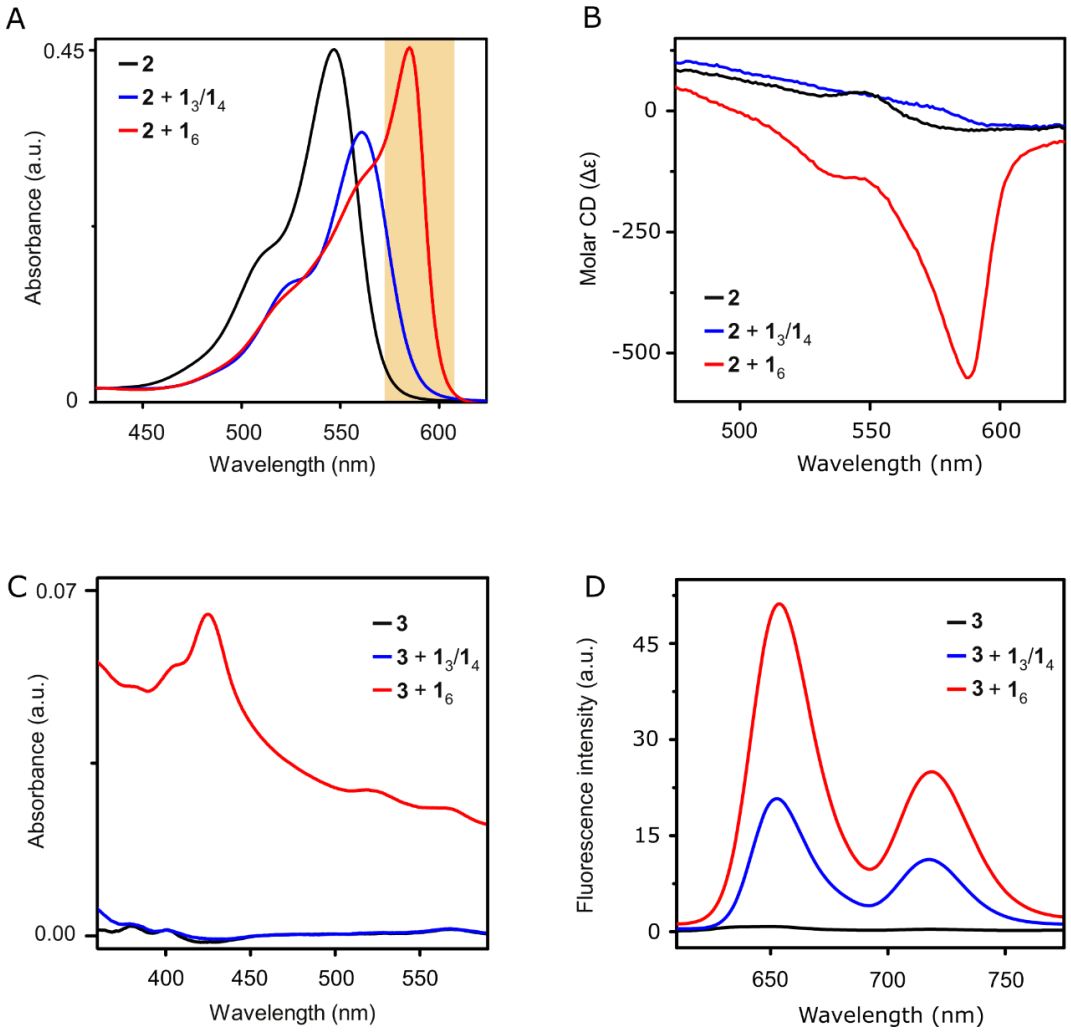

Fig. 2. Optical properties of 2 and 3 in different environments. *UV-Vis absorption spectra* ***(A)*** *and circular dichroism spectra* ***(B)*** *of* ***2*** *(4.0 µM) in buffer (black), in the presence of* ***1***$_3$/***1***$_4$ *(80 µM in* ***1****; blue) and in the presence of* ***1***$_6$ *(80 µM in* ***1****; red). The shaded area represents the range of wavelengths used for irradiation. UV-Vis absorption spectra* ***(C)*** *and fluorescence spectra (excitation at 420 nm)* ***(D)*** *of* ***3*** *(10 µM) in buffer (black) and in the presence of* ***1***$_3$/***1***$_4$ *(1.0 mM in* ***1****; blue) and* ***1***$_6$ *(1.0 mM in* ***1****; red).*

Tetraphenylporphyrin acts as a photocatalytic cofactor that can be activated for a wide range of wavelengths

A limitation of using **2** as a cofactor is that the irradiation source is restricted to a small and specific range of wavelengths. This restriction can be relaxed by using a different dye as a cofactor that can be excited with white light. Thus, we used tetraphenylporphyrin (**3**), which is a well-known $^1O_2$ photosensitizer (*28*). This nonpolar dye is poorly soluble in buffer (Fig. 2C), hampering its use as a photosensitizer in aqueous solution. However, in presence of **1**$_6$, **3** is solubilized, as indicated by the appearance of its Soret (425 nm) and Q bands (520, 566, and 660 nm) (Fig. 2C and S5). Compared with its spectrum in dimethyl formamide, the Soret band of **3** was red-shifted by 9 nm and broadened (Fig. S6), indicative of multiple binding sites on the **1**$_6$ fibers.

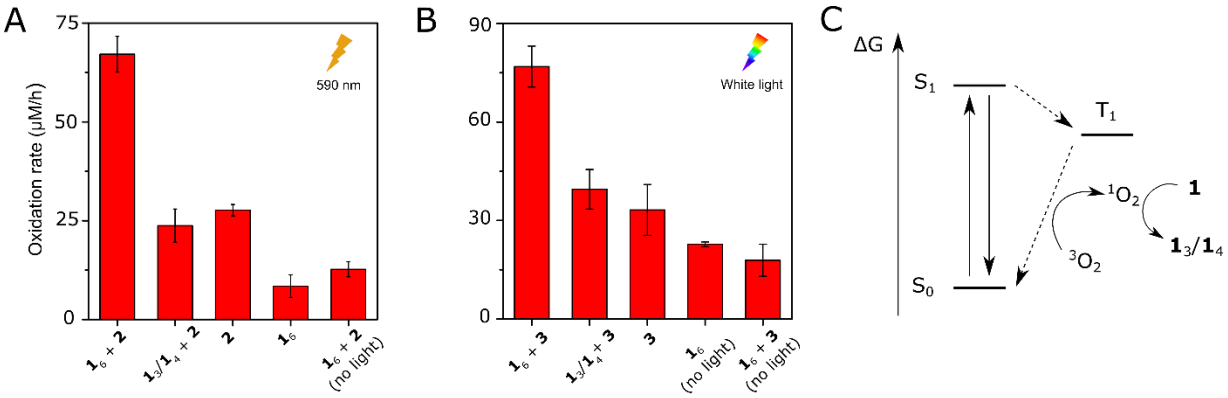

Fig. 3. Replicator 1_6 enhances photocatalytic oxidation of 1 mediated by 2 and 3. ***(A)*** *Initial rates of oxidation of* ***1*** *(200 μM) in the presence of* ***2*** *(2.5 μM) and* ***1_6*** *(80 μM) or* ***1_3/1_4*** *(200 μM).* ***(B)*** *The initial oxidation rates of* ***1*** *(700 μM) in the presence of* ***3*** *(10 μM) and* ***1_6*** *(50 μM) or* ***1_3/1_4*** *(150 μM). All samples were kept at 25 ˚C and irradiated with 590 nm light (for A) or a halogen lamp (for B). Rates were measured by monitoring the concentration of* ***1*** *over 2h using UPLC (Methods). The data shown is the average of the rates calculated for three (A) or two (B) independent samples, and the error bars correspond to their standard deviation.* ***(C)*** *Jablonski diagram showing the photochemical processes involved in the formation of $^1O_2$ (excitation from the ground state (S_0) to the singlet excited state (S_1), followed by intersystem crossing to the triplet excited state (T_1) which promotes singlet oxygen production) and the subsequent oxidation of thiols to small-ring disulfides.*

Binding of **3** to replicator $\mathbf{1}_6$ resulted in an increase in fluorescence from the dye (Fig. 2D) and also enhanced its production of $^1O_2$, manifested in the enhanced rate of oxidation of a fluorescent probe (*29*) (Fig. S7). In addition to solubilizing the dye, the hydrophobically driven binding of **3** to the replicator $\mathbf{1}_6$ may also alleviate the self-quenching of **3**, caused by its aggregation. Importantly, replicator $\mathbf{1}_6$ is better at solubilizing **3** than replicator precursors $\mathbf{1}_3$/$\mathbf{1}_4$ (Fig. 2D), which is presumably due to the larger hydrophobic binding sites offered by the $\mathbf{1}_6$ fibers. An advantage of this strategy is that it can be used to bind and activate other hydrophobic porphyrins (Fig. S8).

Upon binding to replicator $\mathbf{1}_6$ and irradiation, dye **3** induced a 2-fold enhancement in the rate of oxidation of **1** compared to the same dye in the presence of $\mathbf{1}_3$/$\mathbf{1}_4$ (Fig. 3B). Thus, like **2**, also **3** acts as a photocatalytic cofactor, albeit through a different activation mechanism than **2**. Importantly, in both systems the replicator accelerates the light-driven formation of $\mathbf{1}_3$/$\mathbf{1}_4$. Given that these macrocycles are the precursors for the replicator $\mathbf{1}_6$, this behavior corresponds to protometabolism.

Photooxidative protometabolism yields the precursors of the replicator, increasing its replication efficiency

The dynamic behavior of our systems under conditions where self-replication and oxidative protometabolism could influence each other was explored. Dynamic combinatorial libraries made from thiol building block **1** were prepared in the presence of **2** or **3**, without added oxidant other than dissolved oxygen from the air, and stirred while irradiating with yellow light (590 nm) or a halogen lamp, respectively (Supplementary text). In both experiments, as soon as self-replicator $\mathbf{1}_6$ emerged, the rate of oxidation of thiol **1** increased (Fig. 4A and C;

top half of each graph shows the concentrations of **1** in black and $\mathbf{1}_6$ in red; bottom half of each graph shows the rates at which these species are consumed and produced, respectively), consistent with the previous experiments (Fig. 3A and B). As expected, in the controls without cofactor the oxidation rate did not change with the emergence of $\mathbf{1}_6$, remaining relatively low and decreasing slowly due to the depletion of **1** (Fig. 4B, D). Importantly, the enhanced rate of oxidation in the presence of cofactor also induced an increase in replication rate, constituting a second positive feedback loop in the system.

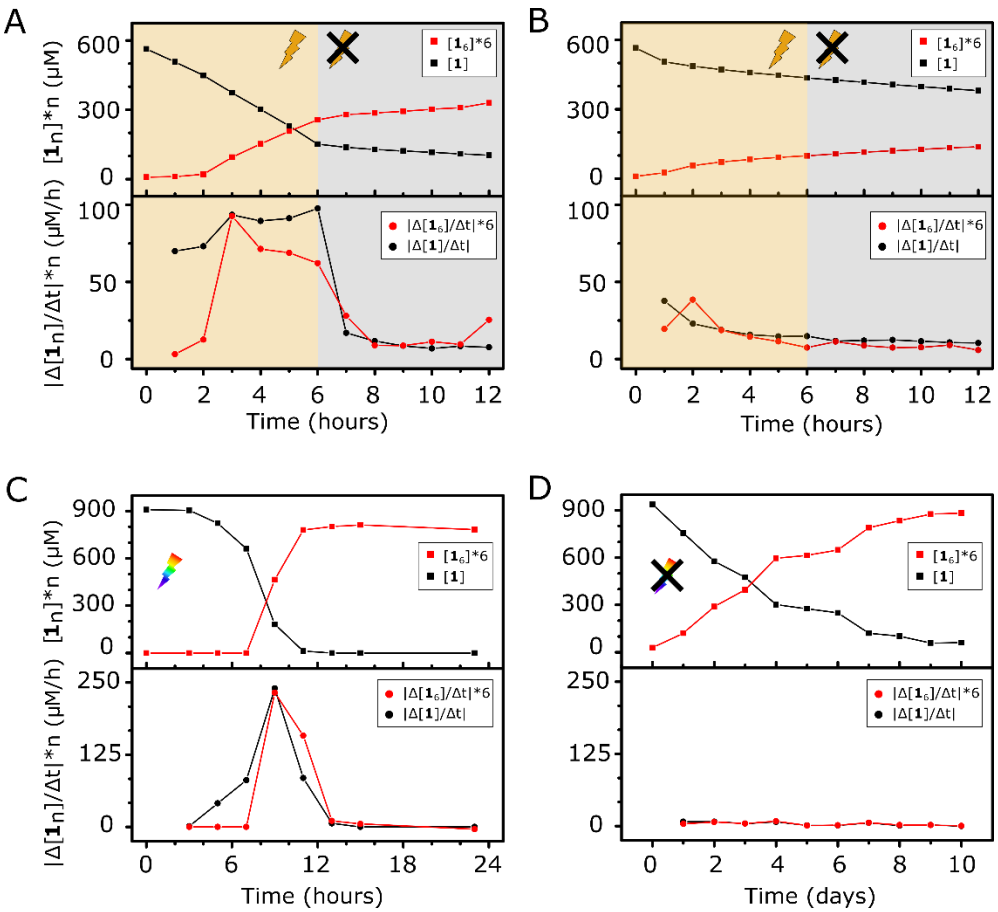

Fig. 4. Emergence of replicator 1_6 promotes photocatalytic production of precursors 1_3/1_4 which promotes replication. *Change in the concentrations of* ***1*** *and* $\mathbf{1}_6$ *in libraries in the presence* ***(A)*** *or absence of 1.0 μM* ***2 (B)****. Samples were stirred at 1200 rpm at 40 ˚C and irradiated with 590 nm light only for the first 6h, as marked by the orange color in A and B. Change in the concentrations of* ***1*** *and* $\mathbf{1}_6$ *in libraries in the presence* ***(C)*** *or absence of 5.0 μM* ***3 (D)****. These samples were stirred at 1200 rpm at 25 ˚C. The sample corresponding to panel C was irradiated with a halogen lamp during the whole process, and that corresponding to panel D was kept in the dark. The top half of each graph represents the concentrations of* ***1*** *(black) or* $\mathbf{1}_6$ *(red) in units of* ***1****. The bottom half of each graph represents the rates of oxidation (black) and replication (red), calculated by numerical differentiation of the data in the top halves. Note the difference in abscissa (time range) in 2C and 2D.*

We performed additional control experiments to confirm that the increase in oxidation and replication rates requires both photo irradiation and presence of replicator. Upon switching off photoirradiation mid-way during the experiment in Fig. 4A, both oxidation and replication rates dropped markedly, reaching values similar to those observed for control samples without

cofactor (Fig 4B) or those that were not irradiated (Fig. S9). Note that, by preventing the emergence of the replicator (by not stirring the libraries; replicator growth depends strongly on shear stress) (*18*), cofactor activation cannot take place, and the oxidation rate remains low throughout the experiment (Fig. S10).

CONCLUSIONS

In conclusion, we show how, upon photoirradiation, dyes can act as cofactors in a system of synthetic self-replicators to achieve photocatalytic conversion of molecules, present in the surroundings of the replicator, into precursors from which the replicators can grow, using light as an energy source. This positive feedback of a replicating system on the production of its own precursors can be regarded as protometabolism, but still falls short of complete metabolism since the external energy is not stored nor used to perform an endergonic reaction. Yet, we feel that the emergence of such protometabolism constitutes an important milestone in the de-novo synthesis of life. It opens the way for protometabolic activity to be further optimized and extended through selection by Darwinian evolution, since protometabolism is directly coupled to replication.

Photocatalytic feedback loops of the kind developed here are also attractive components for systems chemistry, as the rates of photochemical processes can be tuned independent of other chemical processes in the system by manipulating light intensity. The ability to orthogonally manipulate the rates of different reaction pathways of a complex chemical system is essential for homing in on the right part of parameter space that harbors specific emergent phenomena (such as, for example, reaction waves or oscillations).

METHODS

Chemicals and irradiation sources

Building block **1** was obtained from Cambridge Peptides Ltd (Birmingham, UK). Dyes **2** and **3** were obtained from Sigma-Aldrich (purities 95% and 99%, respectively). The irradiation at 590 nm was performed with 5 mm LEDs purchased from Kingbright (model L53-LYD) through a local distributor (OKAPHONE, Netherlands), connected to a 5V power source using a homemade setup. The light intensity of the LEDs is 2 mcd with a viewing angle of 60˚, their emission wavelengths are 590 ± 12.5 nm, and their power dissipation is 105 mW, according to the supplier. Halogen lamps (5 W) whose emission encompass the entire visible region (wavelengths higher than 350 nm) were used to irradiate systems containing **3**.

Library preparation

Stock solutions of $\mathbf{1}_6$ and $\mathbf{1}_3$-$\mathbf{1}_4$ were prepared by dissolving building block **1** in borate buffer (pH = 8.2, see Supplementary Methods) to a final concentration of 4.0 mM and either stirring at 1200 rpm at 40 °C (for $\mathbf{1}_6$) or adding 1 equiv. of $NaBO_4$ (for $\mathbf{1}_3$-$\mathbf{1}_4$). Their conversion was monitored by UPLC until completion. The stock solutions were then mixed with stock

solutions of **2** or **3** and diluted with borate buffer to the concentrations indicated in the text for each experiment.

UPLC analysis

UPLC analysis was performed on a Waters Acquity UPLC-H class system equipped with a PDA detector. All analyses were performed using a reversed-phase UPLC column (Aeris Peptide 1.7 µm XB-C18 x 2.10 mm, Phenomenex). The column temperature was kept at 35 °C, and the sample plate was kept at 25 °C, unless otherwise specified. UV absorbance was monitored at 254 nm. For each injection, 10 µL of sample was injected. The peaks were assigned based on controls or previous results (*17*). Examples of typical chromatograms for the libraries studied are shown in the supplementary information (Fig. S12). Only the main components of the system (**1**, $\mathbf{1}_3$, $\mathbf{1}_4$, $\mathbf{1}_6$) were taken into account for the calculations and graphs shown in the text.

ACKNOWLEDGEMENTS

We thank B. M. Matysiak for performing mass spectrometry measurements. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie Grant Agreement (No. 642192) and was supported by ERC (AdG 741774), NWO (VICI grant), and the Dutch Ministry of Education, Culture and Science (Gravitation program 024.001.035). K. L. acknowledges support from Simons Foundation (Award ID: 553330) and Marie Skłodowska-Curie grant (No. 786350).

AUTHOR CONTRIBUTIONS

S.O., G.M.S. and K.L. conceived the experiments; G.M.S. performed the experiments related to Rose Bengal; K.L. performed the experiments related to porphyrins; W.R.B. performed the experiments related to IR luminescence of singlet oxygen; G.M.S, K.L. and S.O. wrote the manuscript.

COMPETING INTERESTS

Authors declare no competing interests.

DATA AVAILABILITY

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

REFERENCES AND NOTES

1. J. D. Sutherland, Studies on the origin of life - the end of the beginning. *Nat. Rev. Chem.* **1**, 0012 (2017).

2. J. W. Szostak, D. P. Bartel, P. L. Luisi, Synthesizing life. *Nature.* **409**, 387 (2001).

3. S. Benner, A. Ricardo, M. Carrigan. Is there a common chemical model for life in the universe? *Curr. Opin. Chem. Biol.* **8**, 672-689. (2004).

4. J. Gayon, C. Malaterre, M. Morange, F. Raulin-Cerceau, S. Tirard, Defining life: conference proceedings. *Orig. Life Evol. Biosph.* **40**, 119-243. (2010).

5. K. Ruiz-Mirazo, C. Briones, A. de la Escosura, Prebiotic systems chemistry: new perspectives for the origins of life. *Chem. Rev.* **114**, 285-366 (2013).

6. Ganti, T. "The criteria of life" in *The Principles of Life* (Oxford University Press, 2003).

7. M. D. Hardy, J. Yang, J. Selimkhanov, C. M. Cole, L. S. Tsimring, N. K. Devaraj. Self-reproducing catalyst drives repeated phospholipid synthesis and membrane growth. *Proc. Natl. Acad. Sci.* **27**, 8187-8192 (2015).

8. P. M. Gardner, K. Winzer, B. G. Davis, Sugar synthesis in a protocellular model leads to a cell signalling response in bacteria. *Nat. Chem.* **1**, 377-383 (2009).

9. K. Adamala, J. W. Szostak, Nonenzymatic template-directed RNA synthesis inside model protocells. *Nat. Chem.* **5**, 495 (2013).

10. N. Ichihashi, K. Usui, Y. Kazuta, T. Sunami, T. Matsuura, T. Yomo. Darwinian evolution in a translation-coupled RNA replication system within a cell-like compartment. *Nat. Commun.* **4**, 2494 2013

11. T. Czárán, B. Könnyű, E. Szathmáry, Metabolically coupled replicator systems: overview of an RNA-world model concept of prebiotic evolution on mineral surfaces. *J. Theor. Biol.* **381**, 39-54 (2015).

12. B. Könnyű, T. Czárán, The evolution of enzyme specificity in the metabolic replicator model of prebiotic evolution. *PLoS One.* **6**, e20931 (2011).

13. A. Pross The driving force for life's emergence: kinetic and thermodynamic considerations. *J. Theor. Biol.* **220**, 393-406 (2003).

14. N. Wagner, A. Pross, E. Tannenbaum, Selection advantage of metabolic over non-metabolic replicators: A kinetic analysis. *Biosystems.* **99**, 126-129 (2010).

15. S. Arsene, S. Ameta, N. Lehman, A. D. Griffiths, P. Nghe. Coupled catabolism and anabolism in autocatalytic RNA sets. *Nucleic Acids Res.* **46**, 9660-9666, (2018).

16. S. Kamioka, D. Ajami, J. Rebek, Autocatalysis and organocatalysis with synthetic structures *Proc. Natl. Acad. Sci. USA.* **107**, 541–544 (2010).

17. M. Colomb-Delsuc, E. Mattia, J. W. Sadownik, S. Otto. Exponential self-replication enabled through a fibre elongation/breakage mechanism. *Nat. Commun*. **6**, 7427. (2015)

18. J. M. Carnall et al., Mechanosensitive self-replication driven by self-organization. *Science.* **327**, 1502-1506 (2010).

19. G. Caetano-Anollés, K. M. Kim, D. Caetano-Anollés, The phylogenomic roots of modern biochemistry: origins of proteins, cofactors and protein biosynthesis. *J. Mol. Evol.* **74**, 1–34 (2012).

20. G. Wächtershäuser, Evolution of the first metabolic cycles. *Proc. Natl. Acad. Sci. USA.* **87**, 200-204 (1990).

21. R. Croce, H. Van Amerongen, Natural strategies for photosynthetic light harvesting. *Nat. Chem. Biol.* **10**, 492 (2014).

22. T. Tankam, K. Poochampa, T. Vilaivan, M. Sukwattanasinitt, S. Wacharasindhu. Organocatalytic visible light induced S-S bond formation for oxidative coupling of thiols to disulfides. *Tetrahedron.* **72**, 788–793 (2016).

23. P. W. J. M. Frederix, J. Idé, Y. Altay, G. Schaeffer, M. Surin, D. Beljonne, A. S. Bondarenko, T. L. C. Jansen, S. Otto, S. J. Marrink. Structural and spectroscopic properties of assemblies of self-replicating peptide macrocycles. *ACS Nano.* **11**, 7858-7868 (2017).

24. Würthner, T. E. Kaiser, C. R. Saha-Möller, J-Aggregates: from serendipitous discovery to supramolecular engineering of functional dye materials. *Angew. Chem. Int. Ed.* **50**, 3376-3410 (2011).

25. Piotr Bilski, Ralph N. Holt, Colin F. Chignell. Premicellar aggregates of Rose Bengal with cationic and zwitterionic surfactants. *J. Photochem. Photobiol. A.* **110**, 67-24 (1997)

26. The observed oxidation rates in presence of $\mathbf{1}_6$ and $\mathbf{1}_3$-$\mathbf{1}_4$ are lower than the ones that could be expected by the differences in absorbance only. IR luminescence (*27*) shows that the quantum yield of $^1O_2$ formation by **2** is considerably reduced when $\mathbf{1}_6$ or $\mathbf{1}_3$/$\mathbf{1}_4$ are present (Fig. S4). This causes a decrease in oxidation rate with the formation of the **2**-$\mathbf{1}_6$ complex, that counters the increase in absorbance and partially explains the observed difference (Table S1).

27. A. A. Krasnovsky. Photoluminiscence of singlet oxygen in pigment solutions. *Photochem. Photobiol.* **29**, 29–36 (1979).

28. R. W. Redmond, J. N. Gamlin. A compilation of singlet oxygen yields from biologically relevant molecules. *Photochem. Photobiol.* **70**, 391-475 (1999).

29. B. A. Lindig, M. A. J. Rodgers, A. P. Schaap. Determination of the lifetime of singlet oxygen in D_2O using 9,10-Anthracenedipropionic acid, a water-soluble probe. *J. Am. Chem. Soc.* **102**, 5590-5593 (1980).

30. M. C. A. Stuart, J. C. van de Pas, J. B. F. N. Engberts, The use of Nile Red to monitor the aggregation behavior in ternary surfactant-water-organic solvent systems. *J. Phys. Org. Chem.* **18**, 929-934. (2005)

Supplementary information for:

Emergence of light-driven protometabolism upon recruitment of a photocatalytic cofactor by a self-replicator

Authors: Guillermo Monreal Santiago[1], Kai Liu[1], Wesley R. Browne[1], Sijbren Otto[1*].

Affiliations:

[1]Centre for Systems Chemistry, Stratingh Institute, University of Groningen, Nijenborgh 4, 9747 AG Groningen, Netherlands.

[*]Corresponding author. Email: s.otto@rug.nl

Contents:

Materials and Methods

Materials

All reagents, solvents and buffer salts were purchased from commercial sources and used without further purification. Building block **1** was obtained from Cambridge Peptides Ltd (Birmingham, UK). Dyes **2** and **3** were obtained from Sigma-Aldrich (purities 95% and 99%, respectively). The temperature of the samples was kept constant by using a Thermo Fisher compact dry bath with a custom milled aluminium block to accommodate HPLC vials (with a dimension of 12 x 32 mm). The irradiation at 590 nm was performed with 5 mm LEDs purchased from Kingbright (model L53-LYD) through a local distributor (OKAPHONE, Netherlands), connected to a 5V power source using a homemade setup. The light intensity of the LEDs is 2 mcd with a viewing angle of 60˚, their emission wavelengths are 590 ± 12.5 nm, and their power dissipation is 105 mW, according to the supplier. Halogen lamps (5 W) whose emission encompass the entire visible region (wavelengths higher than 350 nm) were used to irradiate systems containing **3**. UPLC analysis was performed on a Waters Acquity UPLC-H class system equipped with a PDA detector. All analyses were performed using a reversed-phase UPLC column (Aeris Peptide 1.7 µm XB-C18 x 2.10 mm, Phenomenex). The column temperature was kept at 35 °C, and the sample plate was kept at 25 °C, unless otherwise specified. UV absorbance was monitored at 254 nm. For each injection, 10 µL of sample was injected. UPLC-MS analysis were performed on a Waters Acquity UPLC H-class system coupled to a Waters Xevo-G2 TOF, operated in positive electrospray ionization mode. Capillary, sampling cone and extraction cone were kept at 2.5 kV, 20 V and 4 V respectively. Source and desolvation temperatures were set as 150 º C and 500 º C. Nitrogen was used as cone and desolvation gas (5 L/h and 800 L/h respectively). The UV-Vis spectra shown in Figure S5 were recorded using a Specord 210 Plus spectrophotometer, and all the others were recorded using a Jasco V-630 UV spectrophotometer. CD spectra were recorded using a Jasco J-815 CD spectrometer, keeping the high tension values around 300 V for all of the measurements. Fluorescence spectra were recorded using a JASCO FP6200 fluorimeter. Solutions for UV, CD and fluorescence were prepared and measured in polystyrene cuvettes, from Brand GMBH (Werrheim, Germany), using the corresponding buffer as a blank.

UPLC analysis

The following solvents and gradient were used for UPLC analysis. The flow rate was kept at 0.3 ml/h.

Samples containing **2**		
Time (min)	% MeCN/ 0.1 % TFA	% H_2O/ 0.1 % TFA
0	10	90
1	10	90
1,3	25	75
3	28	72
11	40	60
11,5	95	5
15	95	5
15,5	10	90
20	10	90

Samples containing **3**		
Time (min)	% MeCN/ 0.1 % TFA	% H_2O/ 0.1 % TFA
0	10	90
1	10	90
1,3	25	75
3	28	72
11	40	60
11,5	95	5
12	95	5
12,5	10	90
17	10	90

The peaks were assigned based on controls or previous results (*17*). Examples of typical chromatograms for the libraries studied are shown below (Fig. S12). Only the main components of the system (**1**, $\mathbf{1}_3$, $\mathbf{1}_4$, $\mathbf{1}_6$) were taken into account for the calculations and graphs shown in the text.

Preparation of borate buffer

For the experiments related to **3** and their corresponding controls, sodium borate buffer was prepared by dissolving $Na_2B_4O_7$ in doubly distilled water to a concentration of 50 mM (200 mM in boron atoms). The pH of the buffer was adjusted to 8.2 by addition of HCl (1 M).

For the experiments involving **2** and their corresponding controls, borate buffer with lower ionic strength was used. To prepare it, B_2O_3 (From Sigma-Aldrich) was dissolved in doubly distilled water and diluted to a final concentration of 25 mM (50 mM in boron atoms) after adjusting the pH to 8.2.

Preparation of $\mathbf{1}_6$

A stock solution of $\mathbf{1}_6$ was typically prepared by dissolving building block **1** in the corresponding borate buffer to a final concentration of 4.0 mM, and stirring at 1200 rpm at 40 ˚C for 24-48h, while following the reaction by UPLC.

Preparation of $\mathbf{1}_3$-$\mathbf{1}_4$

Stock solutions of $\mathbf{1}_3$-$\mathbf{1}_4$ were prepared by mixing freshly prepared concentrated solutions of **1** and $NaBO_4$ and diluting to a final concentration of 4.0 mM of each of them. The oxidation level was monitored by UPLC, typically reaching 100% in less than 30 minutes.

Determination of oxidation rates using **2** and 590 nm light

Libraries were prepared by dissolving building block **1** (200 μM) and the different components of the system ($\mathbf{1}_3$-$\mathbf{1}_4$, $\mathbf{1}_6$ and/or **2**, at the concentrations indicated in the main text) in 1.0 mL of borate buffer (pH = 8.2, see above) in a 12 x 32 mm HPLC vial. A 5 mm yellow LED was fitted directly through a Teflon septum screw cap, and the temperature was kept at 25 ˚C using a dry bath. The cap and dry bath shielded the vials from any other irradiation source than the LED. The concentration of the different components of the library was monitored periodically for 2h by injecting the undiluted sample onto an UPLC system. The UPLC peak area was converted to concentration of the different components with a coefficient determined from a calibration curve ($R^2 > 0.99$). The oxidation rate of each sample was determined by linear regression of the concentration of **1** over time ($R^2 > 0.95$). The data shown is the average of three independent samples, and the error bars correspond to their standard deviation.

Determination of oxidation rates using **3** and a halogen lamp

Aliquots of 200 μL $\mathbf{1}_6$ or $\mathbf{1}_3$-$\mathbf{1}_4$ (1.0 mM in **1**) were incubated with **3** (4.0 μM) in sodium borate buffer for 2h, followed by the addition of 300 μL of **1** (0.50 mM). The resulting solution was irradiated using a 5 W halogen lamp at a distance of 3 cm and stirred at 1200 rpm. The concentration of **1** was measured by UPLC before and after 2h of irradiation. The sample was diluted 50 times in doubly distilled water before injection. The peak area of **1** was converted to concentration with a coefficient determined from a calibration curve ($R^2 > 0.99$). The data shown is an average of two independent samples, and the error bars correspond to their standard deviation. The control experiments were performed using the same conditions, except for the absence of $\mathbf{1}_6$, **3**, or light.

Emergence experiments using **2** (corresponding to Figures 4 and S9-S14)

Libraries were prepared by dissolving **1** (around 500 μM) with or without **2** (1.0 μM) in 1.0 mL of borate buffer (pH = 8.2, see above) in a 12 x 32 mm HPLC vial. A 5 mm yellow LED was fitted directly through a Teflon septum screw cap for the irradiated samples, and the temperature was kept at 40 ˚C using a dry bath. The samples were covered so the only irradiation source was the LED. Stirring was performed at 1200 rpm, using Teflon coated magnetic stirring bars (5 x 2 mm, VWR). The concentration of the different species was monitored by UPLC, injecting directly the undiluted sample. The UPLC peak area was converted to concentration of the different components using a coefficient determined from a calibration curve ($R^2 > 0.99$). The oxidation and replication rates were calculated by numerical differentiation.

Emergence experiments using **3** (corresponding to Figures 4 and S9-S14)

Libraries were prepared by dissolving **1** (1.0 mM) with or without **3** (20 μM) in 1.0 mL of borate buffer (pH 8.2, see above) in a 12 x 32 mm HPLC vial. The samples were irradiated with a 5 W halogen lamp from a distance of 3 cm and stirred at 1200 rpm, using Teflon coated magnetic stirring bars (5 x 2 mm, VWR). The samples were monitored periodically by UPLC analysis of an aliquot that was diluted 50 times with doubly distilled water. To correct for the error caused by dilution, the concentration of the different library members was calculated based on the ratio between each peak and the total peak area of the disulfide peaks (6-12 minutes) in the chromatogram.

Determination of the $^1O_2$ generation quantum yield of **2**

Solutions of **2** (4.0 μM) and **1**, $\mathbf{1}_3$/$\mathbf{1}_4$, or $\mathbf{1}_6$ (80 μM in **1**), prepared using D_2O as a solvent, were excited at 532 nm diode (45 mW), and $^1O_2$ phosphorescence collected by a fibre optic coupled to a Shamrock 163 spectrograph (ANDOR) and dispersed onto a iDUS-InGaAs-512 diode array detector.

Supplementary Text

The libraries including **2** and their controls were set up using a less concentrated buffer (see above). This typically increases the time required for the emergence of $\mathbf{1}_6$, so the temperature of these libraries was increased in the emergence experiments (as the ones shown in Fig. 4 A and B) to compensate for this effect and reduce experiment times.

The time required for the emergence of $\mathbf{1}_6$ varied in different repeats of the experiment. We attribute this difference to the stochastic nature of the nucleation of fibers. The emergence of $\mathbf{1}_6$ was always followed by higher oxidation and replication rates in the irradiated libraries, compared to unirradiated controls (Fig. S11).

The photooxidation reactions described here converted the thiols to mostly (>90%) disulfides, as observed in the chromatograms below (Fig. S12), and by studying the evolution of the total concentration of thiols and disulfides over time (Fig. S13). Only small amounts of overoxidized $\mathbf{1}_2$ were observed as a side-product (Fig. S14). The photobleaching of the cofactors cannot be completely discarded, since their low concentration made them hard to quantify by UPLC and their UV spectra changed with the concentration of other species. However, since catalytic activity was still observed after long irradiation times, the extent of photobleaching was limited.

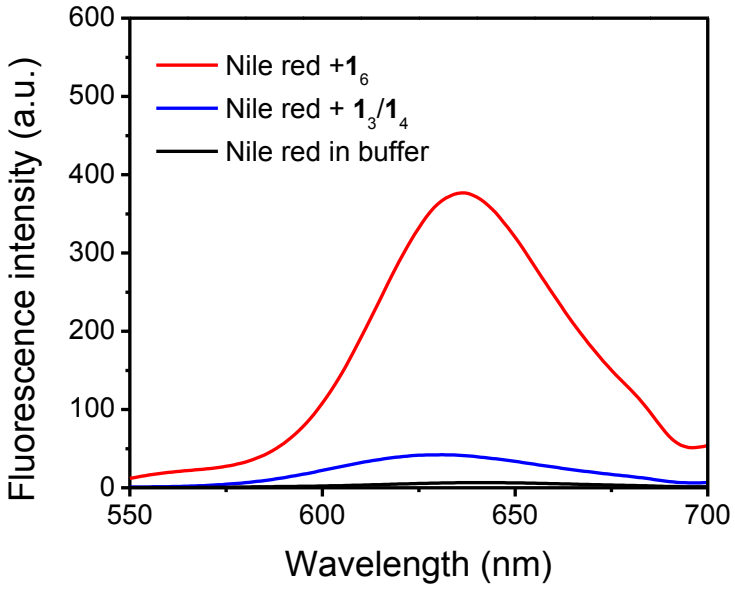

Fig. S1.

Fluorescence of Nile Red (5 μM) in absence and presence of the macrocycles derived from **1** (1.0 mM) with excitation at 530 nm. Enhanced fluorescence indicates the presence of hydrophobic binding sites (*30*).

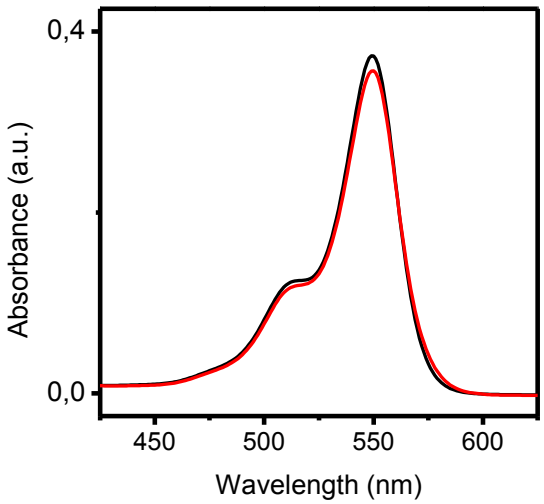

Fig. S2.

Visible absorption spectra of **2** (4.0 μM) in the absence (black) and presence of 80 μM of **1** and 40 μM of TCEP (red)

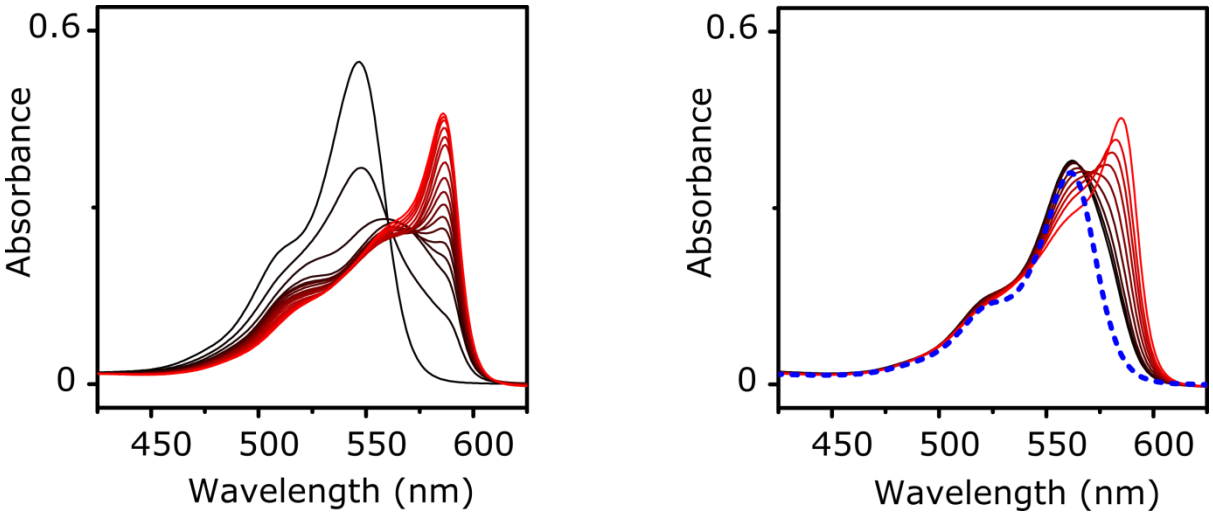

Fig. S3.

Visible absorption spectra of **2** (4.0 μM) in the presence of increasing amounts of $\mathbf{1}_6$. The spectra have been separated in two panels for clarity: Left – from 0 (black) to 16 (red) equivalents of $\mathbf{1}_6$ (in terms of **1**). Right – from 20 (red) to 100 (black) equivalents. In the right panel, the spectrum of **2** (4.0 μM) in presence of 100 equivalents of $\mathbf{1}_3$/$\mathbf{1}_4$ has also been included in a blue dashed line.

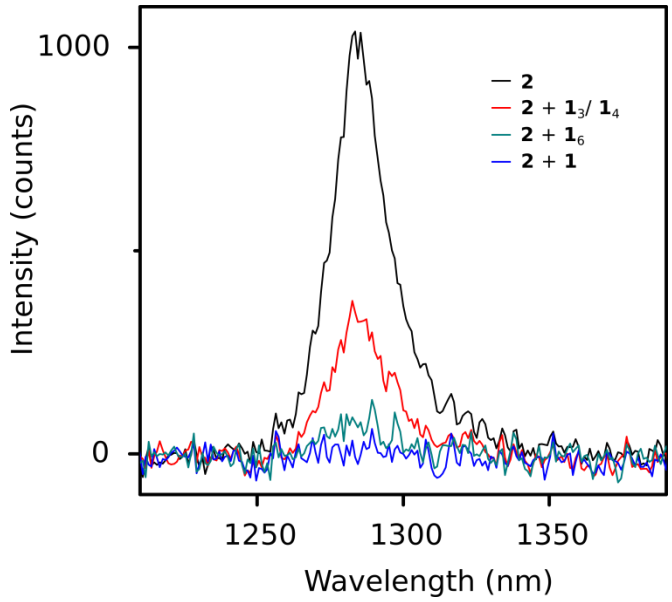

Fig. S4

$^1O_2$ phosphoresence by sensitisation by **2** in the absence or presence of **1** or macrocycles formed from **1**, monitored through the NIR signal corresponding to $^1O_2$ luminescence (*27*). The spectra were recorded with solutions containing **2** (4.0 μM) and **1**, **1**$_3$/**1**$_4$ or **1**$_6$ (80 μM). Dye **2** retains its activity as photosensitizer in the presence of **1**$_3$/**1**$_4$ or **1**$_6$, albeit with reduced efficiency. In the presence of **1**, as expected, $^1O_2$ emission was not observed as it reacts with the thiols of **1**. The samples were prepared by diluting 4.0 mM stock solutions prepared in borate buffer (pH = 8.2) with D_2O, and irradiated with at 532 nm (45 mW)

Table S1.

Calculated quantum yields of $^1O_2$ generation and expected relative rates of photooxidation with **2** in presence of **2** (4.0 μM) and $\mathbf{1}_3/\mathbf{1}_4$ or $\mathbf{1}_6$ (80 μM). The observed oxidation rates in presence of $\mathbf{1}_3/\mathbf{1}_4$ or $\mathbf{1}_6$ are around 2-3 times lower than the values expected based only on the effect of the macrocycles on absorbance and quantum yield, indicating that these macrocycles influence the reaction in more ways than by just increasing the absorbance of **2** or decreasing its Φ_Δ.

Sample	**2**	**2** + $\mathbf{1}_6$	**2** + $\mathbf{1}_3/\mathbf{1}_4$
Relative absorbance (532 nm)	1	0.56	0.67
Relative luminescence (integrated)	1	0.1	0.3
$^1O_2$ generation quantum yield, Φ_Δ	0.76 (*28*)	0.1	0.4
Relative absorbance (577 - 603 nm)	1	39.2	6.7
Expected relative photooxidation rate ($Abs_{577\text{-}603} * \Phi_\Delta$), normalized to **2**	1	5	3
Experimental relative photooxidation rate	1	2.4	0.9

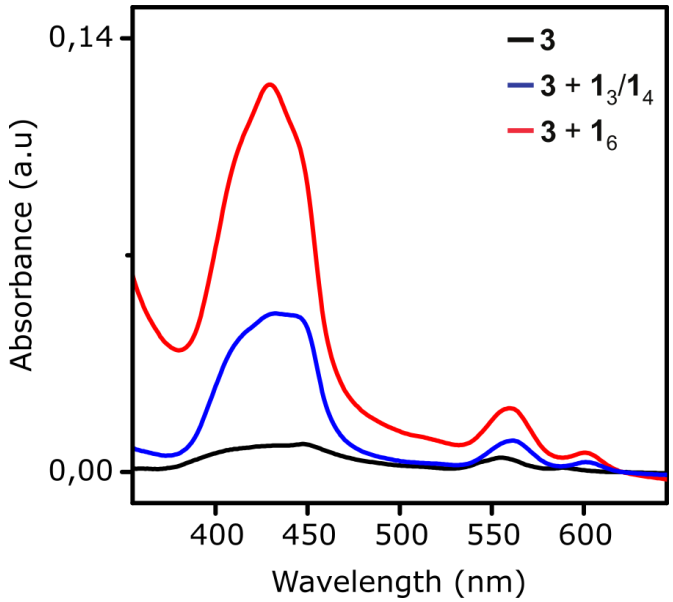

Fig. S5

Visible absorption spectra of **3** (10 μM) in buffer (black), and in the presence of **1**$_3$/**1**$_4$ (1.0 mM in **1**; blue) and **1**$_6$ (1.0 mM in **1**; red). In order to show that the increased absorbance of **3** in presence of **1**$_6$ is not related to scattering, this effect was minimized by recording these spectra in a different spectrophotometer (see Materials), using a wider slit (4 nm), a shorter light path (4.5 mm), and placing the cuvette immediately before the detector.

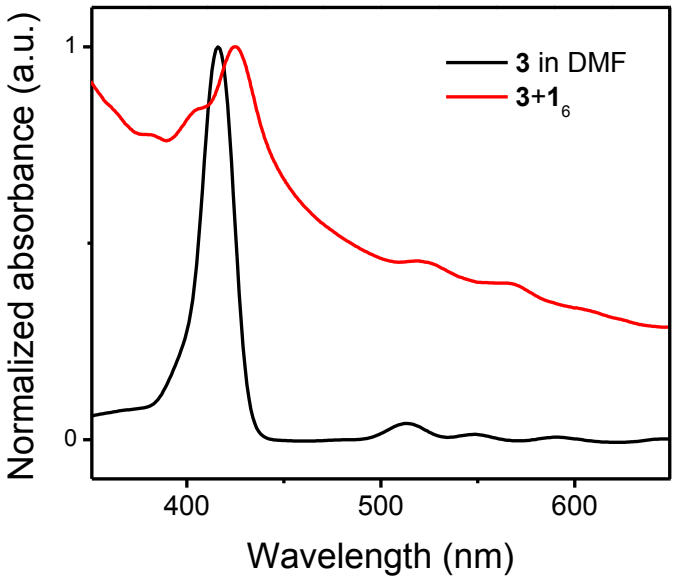

Fig. S6

Visible absorption spectra of **3** (10 μM) in DMF (black) and in borate buffer, in presence of $\mathbf{1}_6$ (1.0 mM in **1**, red).

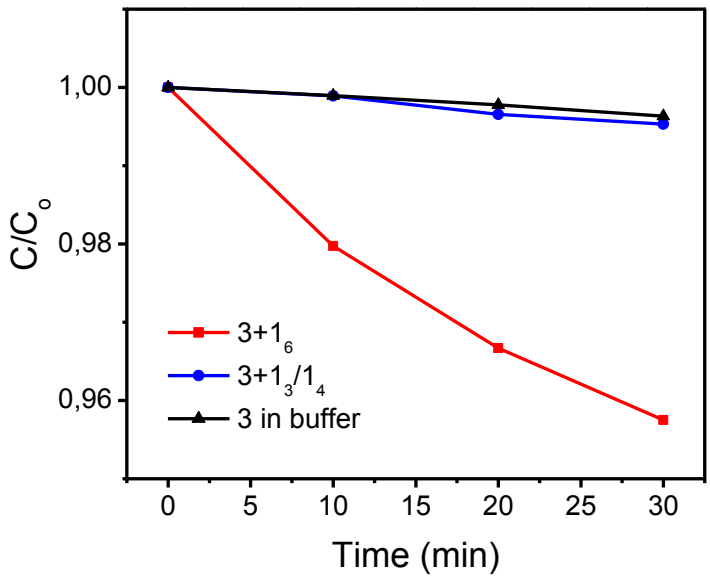

Fig. S7

Relative $^1O_2$-generation efficiency of **3** (10 μM) in aqueous solutions in the presence of $\mathbf{1}_6$ or $\mathbf{1}_3/\mathbf{1}_4$ (1.0 mM in **1**) by using anthracene-9,10-dipropionic acid (4.0 μM) as an ROS probe. The samples were irradiated using a halogen lamp. The absorption decrease at 378 nm was measured by UV–vis absorption spectrophotometry.

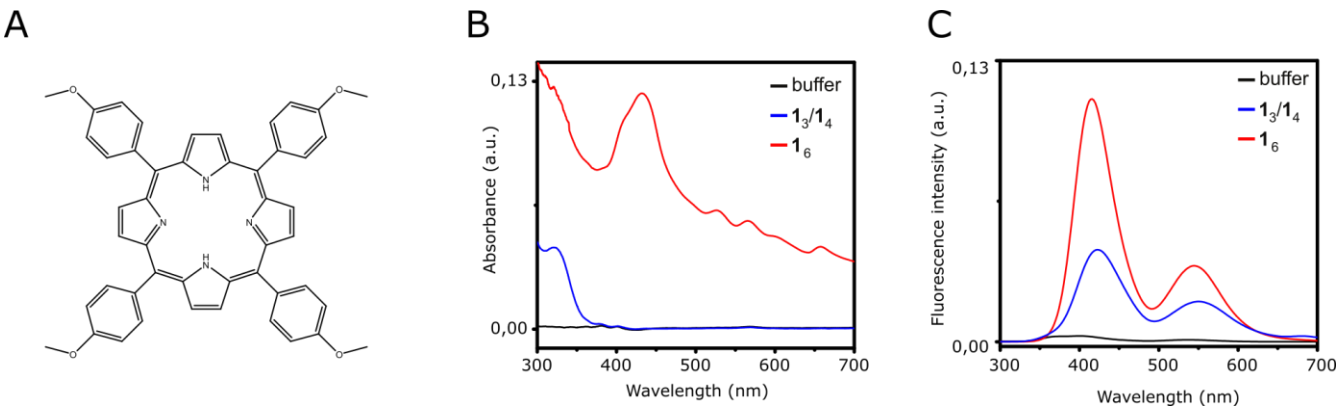

Fig. S8

a) Molecular structure of 5,10,15,20-tetrakis(4-methoxyphenyl)porphyrin, b) UV/vis absorption spectra and c) fluorescence spectra (excitation at 420 nm) of **3** (10 μM) in buffer (black) and in the presence of $\mathbf{1}_3/\mathbf{1}_4$ (1.0 mM in **1**; blue) or $\mathbf{1}_6$ (1.0 mM in **1**; red).

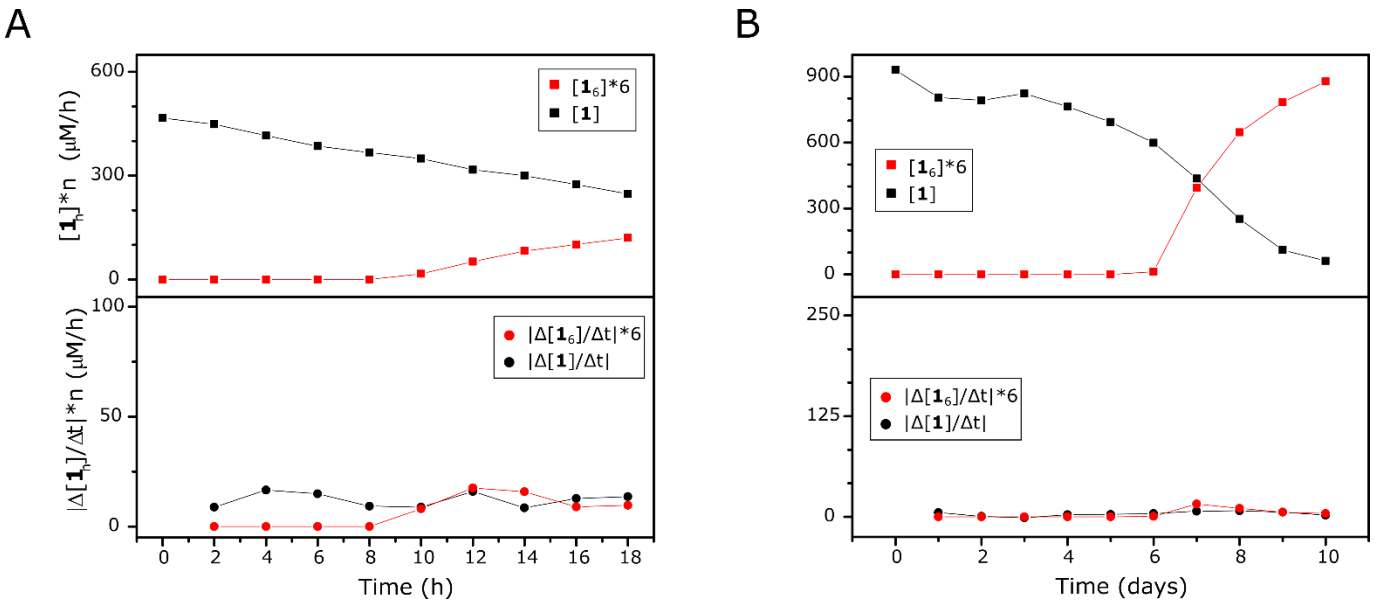

Fig. S9

Evolution over time of the concentrations of **1** and $\mathbf{1}_6$ in non-irradiated libraries made from **1** and **2** (A) or **1** and **3** (B). The top halves of the graphs show the concentration of **1** (black squares) and $\mathbf{1}_6$ (red squares, in units of **1**). The bottom halves of the graphs represent the oxidation (black circles) and replication (red circles) rates in the system, calculated by numerical differentiation of the curves in the top halves. Both libraries were prepared exactly as their equivalents in Figure 4, but they were kept in the dark.

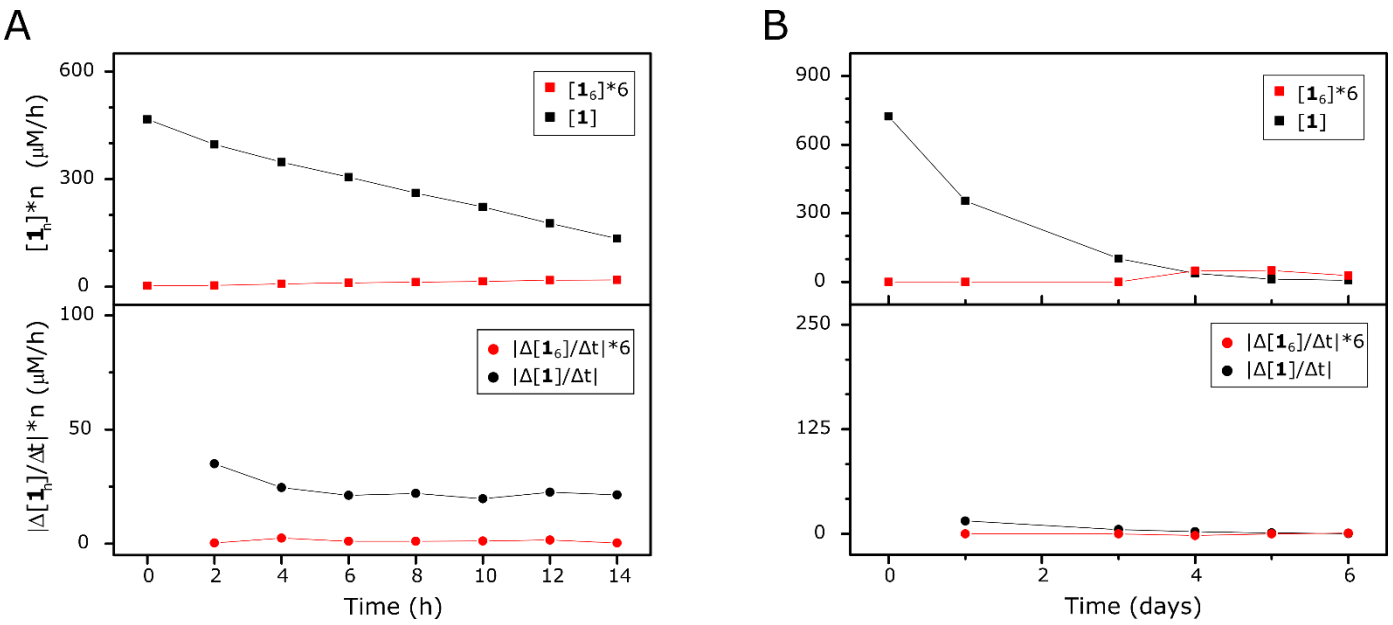

Fig. S10

Evolution over time of the concentrations of **1** and $\mathbf{1}_6$ in unstirred libraries made from **1** and **2** (A) or **1** and **3** (B). The top halves of the graphs show the concentration of **1** (black squares) and $\mathbf{1}_6$ (red squares, in units of **1**). The bottom halves of the graphs represent the oxidation (black circles) and replication (red circles) rates in the system, calculated by numerical differentiation of the curves in the top halves. Both libraries were prepared exactly as their equivalents in Figure 4, but they were not stirred.

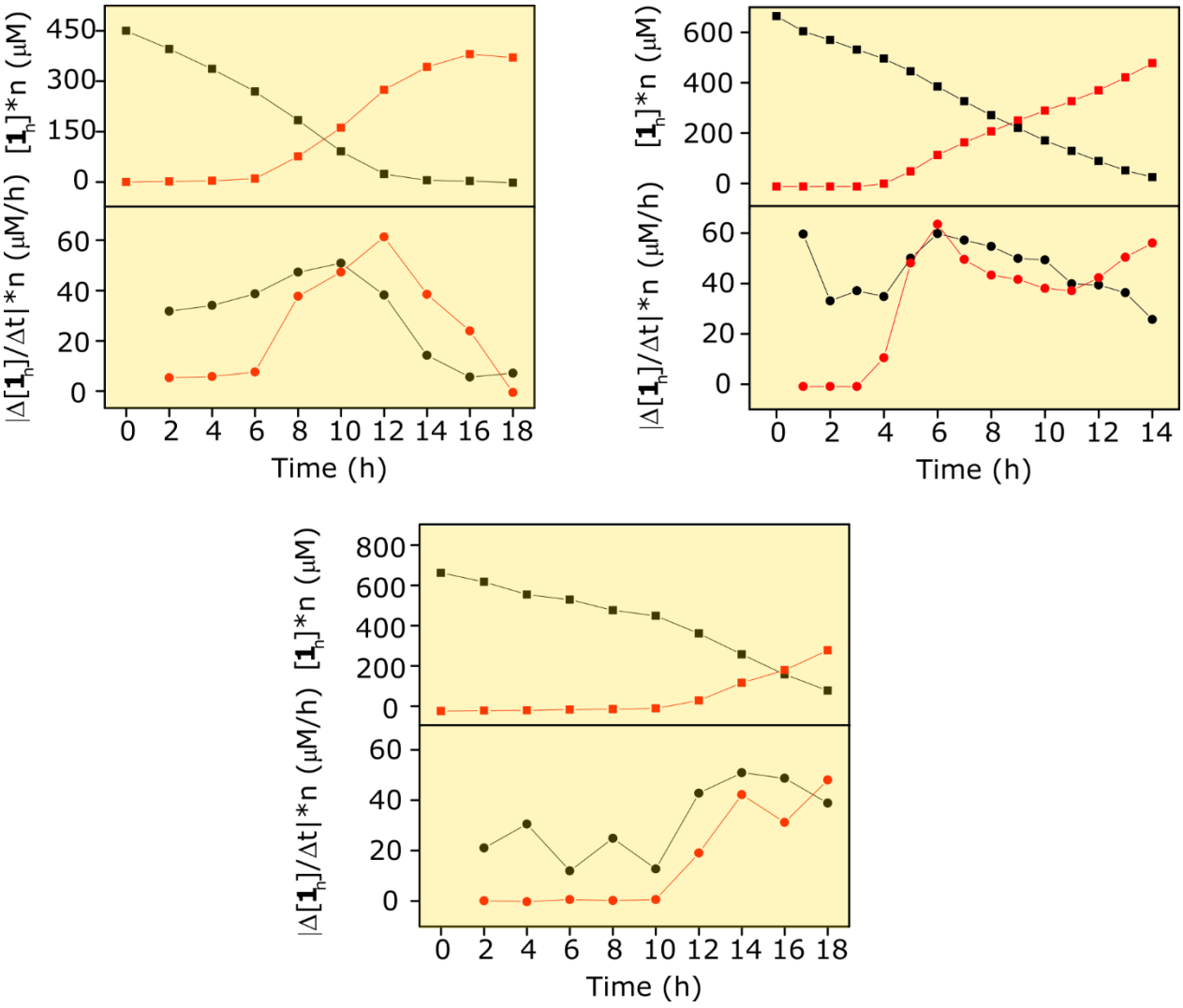

Fig. S11

Evolution over time of the concentrations of **1** and $\mathbf{1}_6$ in three repeats of the same experiment, containing **1** and **2**. The top halves of the graphs show the concentration of **1** (black squares) and $\mathbf{1}_6$ (red squares, in **1** units). The bottom halves of the graphs represent the oxidation (black circles) and replication (red circles) rates in the system, calculated by numerical differentiation of the two curves in the top halves. All libraries were prepared using the same conditions as for the experiments shown in Figure 4A.

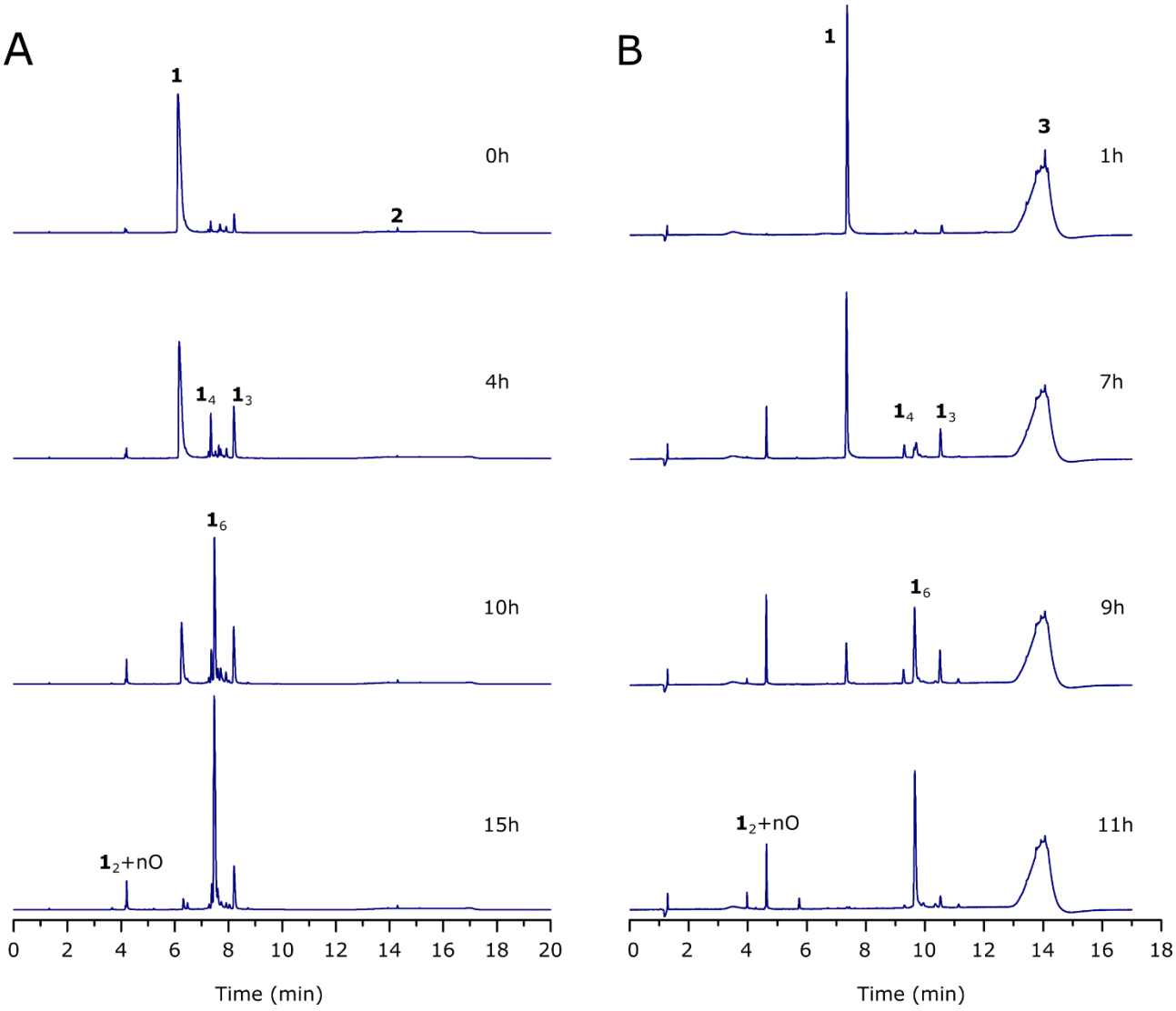

Fig. S12

UPLC chromatograms (monitored at 254 nm) of typical stirred libraries of A) **1** + **2** under 590 nm irradiation, kept at 40 °C and B) **1** + **3**, irradiated with a halogen lamp at room temperature (in the same conditions as the ones in Figures 4 A and C, respectively). Building block **1**, and the disulfides formed by its oxidation, elute between 6 and 12 minutes and have similar molar absorptivities at 254 nm (*17*). The evolution over time of the total peak area between 6 and 12 minutes is shown in Fig. S13, and the mass spectrum corresponding to the peak at ~4 minutes is shown in Fig. S14.

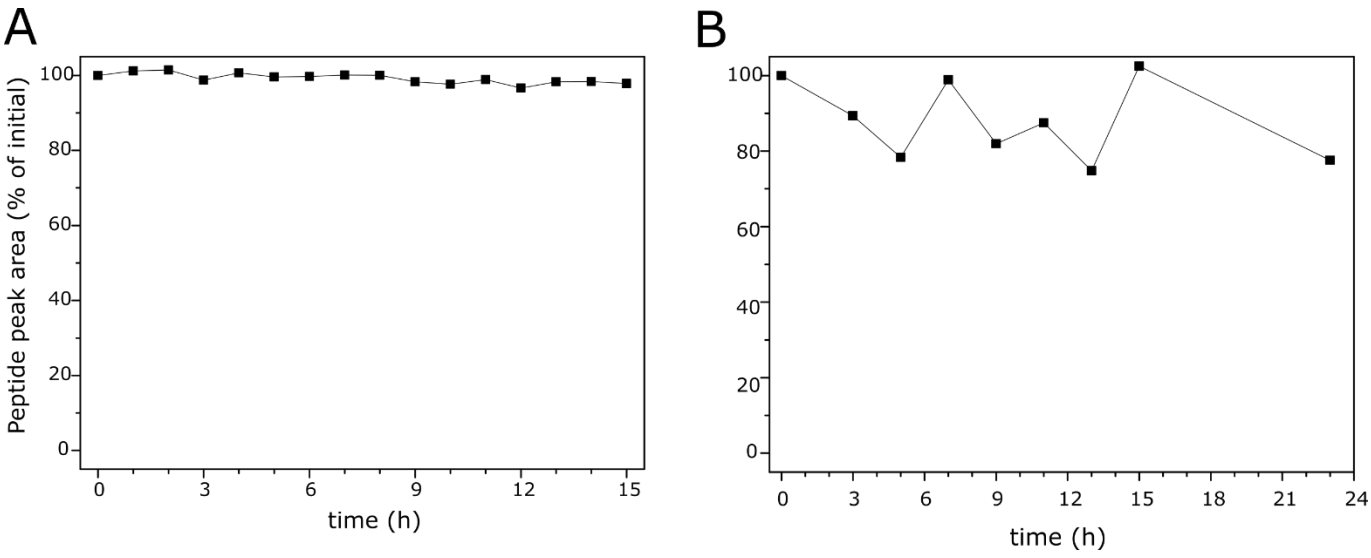

Fig. S13

Peak area between 6 and 12 minutes (corresponding to **1** and its disulfide-based derivatives) in successive chromatograms of irradiated libraries a) containing **1** + **2**, kept at 40 °C and b) containing **1** + **3** and kept at room temperature (the libraries were prepared using the same conditions as the ones in Figure 4). The values are represented as % of the initial peak area, showing that in both cases the amount of oxidation products other than disulfides is negligible. The noise in the sample containing **3** is higher due to sampling (see Methods).

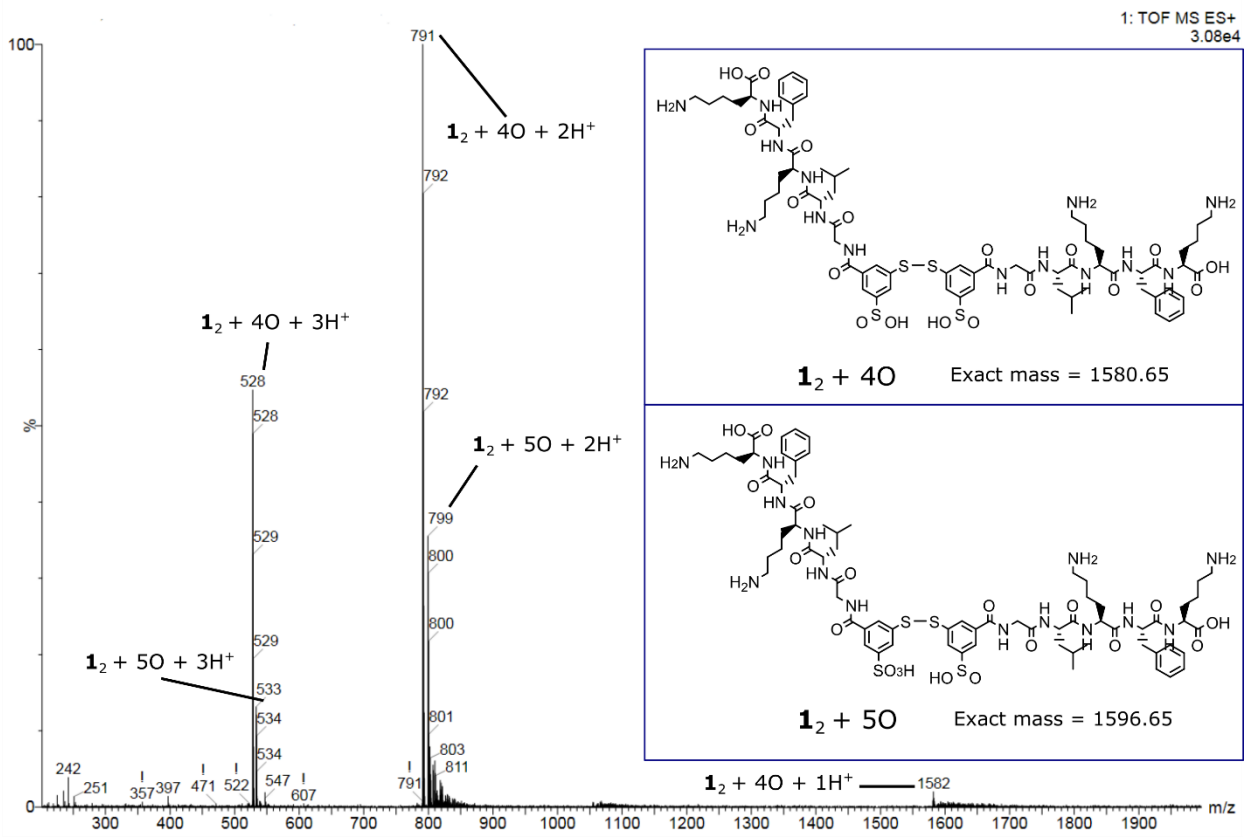

Fig. S14

Mass spectrum of the peak formed in photooxidized libraries, with retention time ~4 minutes. Possible isomers of the ions that can be detected.