



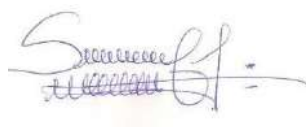


# Affidavit

Je soussigné, **Poutoum Palakiyém SAMIRE**, déclare par la présente que le travail présenté dans ce manuscrit est mon propre travail, réalisé sous la direction scientifique de **Frédéric BEISSON**, dans le respect des principes d'honnêteté, d'intégrité et de responsabilité inhérents à la mission de recherche. Les travaux de recherche et la rédaction de ce manuscrit ont été réalisés dans le respect à la fois de la charte nationale de déontologie des métiers de la recherche et de la charte d'Aix-Marseille Université relative à la lutte contre le plagiat.

Ce travail n'a pas été précédemment soumis en France ou à l'étranger dans une version identique ou similaire à un organisme examinateur.

Fait à Manosque, le 21 mars 2023



Cette œuvre est mise à disposition selon les termes de la [Licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International](https://creativecommons.org/licenses/by-nc-nd/4.0/).



# Abstract

Hydrocarbons (alkanes, alkenes) are molecules composed only of hydrogen and carbon atoms that are an integral part of daily human needs. Indeed, they are the basic compounds of fuels and are used in chemistry, as solvents, lubricants and in cosmetics. Almost all the hydrocarbons we use today are of fossil origin. However, pathways of hydrocarbon biosynthesis from fatty acids exist in many organisms and involve various enzymes. Unfortunately, these enzymes often have low turnovers and require catalytic conditions that are difficult to implement in industry. Thus, the understanding of the mechanisms of hydrocarbon-forming enzymes and their engineering has gained significant interest over the last 10 years. The last hydrocarbon-forming enzyme discovered (in 2017) is fatty acid photodecarboxylase (FAP), a protein found only in algae. This enzyme is particularly interesting because its reaction does not require electron donors but only a photon at each catalytic cycle. It is thus a photoenzyme, a rare type of enzyme. FAP has then attracted a lot of interest, especially for the photoconversion of fatty acids to hydrocarbons. My PhD work had three objectives: (i) Determine whether FAP activity was conserved in algae beyond the model green alga *Chlorella variabilis* NC64A where it was discovered, and see if some other putative FAP were likely to have a fatty acid specificity different than the FAP from *Chlorella variabilis* (CvFAP); (ii) Participate in a large multidisciplinary study involving an international consortium of laboratories which aimed at gaining insights into the structure and mechanism of CvFAP; (iii) Study the fatty acid specificity of CvFAP with a focus on short- and medium-chain fatty acids, which were thought to be poor substrates for FAP. In the first part, I have expressed in *E. coli* seven homologs of CvFAP from various algal groups and be able to obtain soluble active FAP for four of them: FAPs from *Ectocarpus siliculosus* (a brown macroalga), *Chondrus crispus* (a red macroalga), *Nannochloropsis gaditana* (a microalga close to brown algae) and *Galdieria sulphuraria* (a red microalga) were found to have conserved FAP activity but with distinct fatty acid specificities. In the second part, I have participated to production and purification of CvFAP on a large in scale for different biophysical approaches, performed biochemical, static crystallography and transient absorption spectroscopy experiments on CvFAP and on a mutant (R451K) important for the substrate stabilization. The whole study allowed us to elucidate the complete CvFAP mechanism. In the third part, I have studied the activity of CvFAP in vitro and shown that it can convert *n*-octanoic acid four times faster than *n*-hexadecanoic acid, its best substrate reported to date. I have also shown that in vivo this translates into a CvFAP -based production rate over ten-fold higher for *n*-heptane than for *n*-pentadecane. Experiments of time-resolved spectroscopy I have performed in collaboration have provided evidence that the high catalytic activity of FAP on *n*-octanoic acid is in part due to an autocatalytic effect of its *n*-heptane product, which fills the rest of the binding pocket of CvFAP. Finally, I have determined the effect of substrate concentration, pH (and light intensity) on the activity of CvFAP on *n*-hexanoic and *n*-butanoic acids. These results should guide future strategies of FAP selection, improvement and use for a light-driven production of medium- and short-chain hydrocarbons.

**Key words:** hydrocarbons, fatty acids, photoenzymes, biocatalysis, biofuels.



# Résumé

Les hydrocarbures (alcane, alcène), sont des molécules composées uniquement d'atomes d'hydrogène et de carbone qui font partie intégrante des besoins quotidiens de l'homme. En effet, ils sont les composés de base des carburants et sont utilisés en chimie, comme solvants et lubrifiants, ainsi qu'en cosmétique. La quasi-totalité des hydrocarbures que nous utilisons aujourd'hui sont d'origine fossile mais des voies de biosynthèse d'hydrocarbures à partir des acides gras existent dans de nombreux organismes. Malheureusement, ces enzymes ont souvent un faible rendement et nécessitent des conditions catalytiques difficiles à mettre en œuvre dans l'industrie. Ainsi, la compréhension des mécanismes des enzymes formant des hydrocarbures et leur ingénierie ont fait l'objet de nombreux travaux au cours des dix dernières années. La dernière enzyme productrice d'hydrocarbures à avoir été découverte (en 2017) est l'acide gras photodécarboxylase (acronyme anglais : FAP), une protéine que l'on ne trouve que dans les algues. Cette enzyme est particulièrement intéressante car sa réaction ne nécessite pas de donneurs d'électrons mais seulement un photon à chaque cycle catalytique. Il s'agit donc d'une photoenzyme, un type d'enzyme très rare. La FAP a suscité beaucoup d'intérêt en biocatalyse. Mon travail de doctorat avait trois objectifs : (i) Déterminer si l'activité de la FAP est conservée dans les algues au-delà de l'algue verte modèle *Chlorella variabilis* NC64A où elle a été découverte, et voir si d'autres FAP putatives sont susceptibles d'avoir une spécificité d'acide gras différente de la FAP de *Chlorella variabilis* (CvFAP) ; (ii) Participer à une vaste étude pluridisciplinaire impliquant un consortium international de laboratoires et visant à mieux comprendre la structure et le mécanisme de la CvFAP; (iii) Étudier la spécificité de substrat de la CvFAP en mettant l'accent sur les acides gras à chaîne courte et moyenne, considérés comme de mauvais substrats pour la FAP. Dans la première partie, j'ai exprimé dans *E. coli* sept homologues de la CvFAP provenant de divers groupes d'algues et j'ai pu obtenir une FAP active soluble pour quatre d'entre eux : Les FAP d'*Ectocarpus siliculosus* (macroalgue brune), de *Chondrus crispus* (macroalgue rouge), de *Nannochloropsis gaditana* (microalgue proche des algues brunes) et de *Galdieria sulphuraria* (microalgue rouge) se sont révélées avoir une activité FAP conservée mais avec des spécificités de conversion d'acides gras distinctes. Dans la deuxième partie, j'ai participé à la production et à la purification de la CvFAP à grande échelle pour différentes approches biophysiques. J'ai réalisé des expériences biochimiques, de cristallographie statique et de spectroscopie d'absorption transitoire sur la CvFAP et sur un mutant (R451K) important pour la stabilisation du substrat. L'ensemble de l'étude a permis d'élucider tout le mécanisme de la CvFAP. Dans la troisième partie, j'ai étudié l'activité de la CvFAP in vitro et montré qu'elle peut convertir l'acide *n*-octanoïque quatre fois plus vite que l'acide *n*-hexadécanoïque, son meilleur substrat rapporté à ce jour. J'ai également montré que, in vivo, cela se traduit par un taux de production (basé sur la CvFAP) plus de dix fois supérieur pour le *n*-heptane que pour le *n*-pentadécane. Les expériences de spectroscopie résolue dans le temps que j'ai réalisées en collaboration ont prouvé que la forte activité catalytique de la FAP sur l'acide *n*-octanoïque est en partie due à un effet autocatalytique de son produit, le *n*-heptane, qui remplit le reste de la cavité du site actif de la CvFAP. Enfin, j'ai déterminé l'effet de la concentration en substrat, du pH et de l'intensité lumineuse sur l'activité de la CvFAP pour la conversion des acides *n*-hexanoïque et *n*-butanoïque. Ces résultats devraient guider les futures stratégies de sélection, d'amélioration et d'utilisation de la FAP pour la production à la lumière d'hydrocarbures à chaîne courte ou moyenne.

**Mots clés :** hydrocarbures, acides gras, photoenzymes, biocatalyse, biocarburants.





# List of publications

1- Moulin SLY, Beyly-Adriano A, Cuiné S, Blangy S, Légeret B, Floriani M, Burlacot A, Sorigué D, **Samire PP**, Li-Beisson Y, Peltier G, Beisson F. Fatty acid photodecarboxylase is an ancient photoenzyme that forms hydrocarbons in the thylakoids of algae. *Plant Physiol.* 2021 Jul 6;186(3):1455-1472. doi: 10.1093/plphys/kiab168. PMID: 33856460; PMCID: PMC8260138.

2- Sorigué D, Hadjidemetriou K, Blangy S, Gotthard G, Bonvalet A, Coquelle N, **Samire P**, Aleksandrov A, Antonucci L, Benachir A, Boutet S, Byrdin M, Cammarata M, Carbajo S, Cuiné S, Doak RB, Foucar L, Gorel A, Grünbein M, Hartmann E, Hienerwadel R, Hilpert M, Kloos M, Lane TJ, Légeret B, Legrand P, Li-Beisson Y, Moulin SLY, Nurizzo D, Peltier G, Schirò G, Shoeman RL, Sliwa M, Solinas X, Zhuang B, Barends TRM, Colletier JP, Joffre M, Royant A, Berthomieu C, Weik M, Domratcheva T, Brettel K, Vos MH, Schlichting I, Arnoux P, Müller P, Beisson F. Mechanism and dynamics of fatty acid photodecarboxylase. *Science.* 2021 Apr 9;372(6538):eabd5687. doi: 10.1126/science.abd5687. PMID : 33833098.

3-**Poutoum P. Samire**, Bo Zhuang, Bertrand Légeret, Angel Baca-Porcel, Gilles Peltier, Damien Sorigué, Alexey Aleksandrov, Frédéric Beisson, Pavel Müller. Autocatalytic Effect Boosts the Production of Medium-Chain Hydrocarbons by Fatty Acid Photodecarboxylase. *Science Advances.* 9,eadg3881(2023).DOI:10.1126/sciadv.adg3881

4-**Poutoum P. Samire**, et al. Optimal photoconversion of short-chain fatty acids by fatty acid photodecarboxylase (*in preparation based on additional results described in chapter 5*)

5-Beisson et al. Biosynthesis and function of fatty acid-derived hydrocarbons (invited review by. K. Chapman, in preparation for *Progress in Lipid Research*).



# List of participation to congress/seminar and formations

## Congress/seminar

**Poster presentation** “Diversity of hydrocarbon synthesis in algae”. 9th European Symposium on Plant Lipids, 7-10 July 2019, Marseille, France.

**Oral presentation** “Autocatalytic Effect Boosts the Production of Medium-chain Hydrocarbons by Fatty Acid Photodecarboxylase”. Seminar Axis 4 (Bio-inspired engineering for energy, health and the environment) Club of the Institute for Integrative Biology of Cell (I2BC), online, 24<sup>th</sup> November 2022

## Formations

■ Formation "Anglais professionnel Niveau B2"- TREFLE FORMATION, 01 mars 2020 - 30 novembre 2021, Saint-Paul-lez-Durance.

■ MOOC Ethique de la Recherche, 29 novembre 2018, Université de Lyon.

■ Atelier “Cristallographie des Protéines : des données de diffraction à la carte de densité “, 11 au 15 juin 2019, CNRS, Station biologique de Roscoff.

■ Formation ImageJ, CEA Cadarache, 16 au 17 mai 2019, Saint-Paul-lez-Durance.

■ Formation en ligne “Research Integrity in Scientific Professions”, 04 mars 2019 - 29 avril 2019.



## Remerciements

Je tenais d'abord à remercier Carine Vergne-Vaxelaire et Catherine Sarazin pour avoir accepté de juger mon travail de thèse en tant que rapporteuses, ainsi que les autres membres examinateurs Frédéric Carrière et Pavel Müller pour leur participation au jury.

Ce travail a été possible grâce aux 4 années de travail acharné réalisé au sein de l'équipe EBMP (Environnement, Bioénergies, Microalgues et Plantes), anciennement LB3M (Laboratoire de Bioénergie et Biotechnologie des Bactéries et des Microalgues). J'ai eu la chance d'évoluer au sein de cette équipe et d'avoir interagit avec diverses personnes intégrées dans les différentes phases d'évolution de l'équipe et des remaniements de thématiques. De ce fait, je tiens à remercier en premier lieu les responsables de l'équipe : Gilles Peltier et Yonghua Li-Beisson pour m'avoir donné l'opportunité de réaliser mes travaux de recherche au sein de leur Laboratoire.

En plus de m'avoir permis d'intégrer son équipe de recherche, Gilles Peltier a été pour moi une oreille attentive et un regard critique pour apporter une perception neuve et constructive à la fois lors des expérimentations réalisées avec lui au MIMS mais aussi et surtout lors de nos échanges d'idées scientifiques. Je tenais à te témoigner toute ma reconnaissance pour ton implication dans mes travaux de thèse qui se sont avérés parfois à la fois complexes et 'interminables'.

Yonghua, au-delà de nous transmettre un regard rigoureux et pointu sur le plan scientifique, d'être attentive à nos besoins, tu es aussi une source de motivation par ta brillante carrière et ton expérience que tu n'as cessé de nous partager afin de nous alerter sur les exigences liées au monde de la recherche scientifique. Je t'adresse un sincère merci.

J'ai eu la chance d'être assez vite orienté et intégré aux diverses thématiques sur la FAP grâce à Damien Sorigué. Damien, tu as été pour moi comme un deuxième encadrant en m'édifiant sur les diverses techniques nécessaires pour l'accomplissement de mes travaux de thèse. Tu m'as permis de gagner en autonomie sur bon nombre de projets tout en ne rechignant jamais à m'apporter des idées nouvelles, à communiquer beaucoup de positivité et d'optimisme. Je te dis un grand merci.

Stéphanie Blangy, dès mes premiers pas dans le laboratoire, tu t'es impliquée beaucoup dans mon projet pour me donner les outils techniques et le 'mindset' nécessaires pour venir à bout de mes projets. Merci pour le baptême des deux semaines de purification enzymatique dans la chambre froide au tout début de ma thèse, eh oui, ça m'a été très utile par la suite. Merci pour ton énergie positive et les qualités humaines que tu communique.

Bertrand Légeret, tu fais partie des personnes sans lesquelles ma thèse n'aurait pas eu un si bon déroulé. Bertrand a toujours été présent que ce soit pour apporter son ingéniosité pour résoudre des problèmes techniques essentiellement liées à l'extraction et à la quantification des hydrocarbures ou pour mener des discussions scientifiques de qualité. Bertrand non seulement tu m'as permis de prendre conscience des limites de certaines de mes expérimentations, tu as également été une source d'idée dans le but de réaliser des travaux originaux en ne lésinant pas sur ton temps. Merci d'avoir été là pour ces discussions le plus souvent tardives et ta motivation pour aller vers des résultats plus concluants.

Frédéric Beisson, j'ai beaucoup apprécié la réalisation de ma thèse sous ta direction. Merci de m'avoir fait confiance sur le sujet de thèse et de m'avoir permis d'être autonome en me laissant la possibilité de développer mes propres idées ; Tu as su être présent surtout pour donner une nouvelle direction à mes expérimentations particulièrement lorsque les attendus étaient pour le moins complexes. J'ai particulièrement apprécié la qualité de nos discussions scientifiques, ton regard critique sur l'interprétation des résultats et tes idées originales toujours dans le but de produire un travail de qualité. Même si ma thèse a été plus longue que prévu, tu as su me communiquer la nécessité d'effectuer un travail complet en explorant divers aspects. Je suis fier du travail que j'ai réalisé avec toi, merci beaucoup.

C'est aussi l'occasion pour moi d'exprimer ma profonde gratitude aux autres membres de l'équipe qui ont su contribuer à la création d'un environnement favorable à mes travaux de recherche : Stéphan Cuiné, Véronique Cardettini, Marie Bertrand, Pascaline Auroy-Tarrago, Audrey Adriano-Beyly, Claire Sahut, Virginie Epting et Ge Pei

Je remercie également tous les étudiants avec qui j'ai pu interagir durant ma thèse et qui ont contribué à créer une atmosphère conviviale.

L'environnement pluridisciplinaire de ma thèse m'a permis de travailler et échanger avec divers chercheurs. Merci à Pavel Müller, de m'avoir introduit à l'univers de la spectroscopie cinétique. Ma collaboration fructueuse avec Pavel est à l'origine de belles publications décrites dans ce document ; à Pascal Arnoux pour les expériences de cristallographie et le traitement de données.

Je ne peux finir mes remerciements, sans évoquer une personne qui a été d'un soutien indéfectible et permanent pour moi. Bien que loin de la paillasse et de la sphère scientifique, mon épouse, Louise a su me reconforter dans les moments difficiles entre les échecs de manipulations et les longues attentes des résultats de soumission d'article. Je dédie également cette thèse à celle par qui j'ai vu le jour, ma mère, Pamh Landa.

# Contents

Abstract .....	4
Résumé .....	6
List of publications .....	8
List of participation to congress/seminar and formations .....	10
Remerciements .....	12
List of figures .....	18
List of tables .....	20
List of abbreviations .....	22
Chapter 1: Bibliographic introduction and objectives of the thesis .....	24
Bibliographic introduction .....	26
A-Fatty acids.....	26
1-Structure and nomenclature .....	26
2- Solubility and self-assembly in aqueous solutions .....	29
2.1-Ionization state .....	29
2.2-Solubility in water .....	31
2.3-Self-assembly into higher order structures.....	34
B-Fatty-acids derived hydrocarbons.....	37
1-Definition, structure, and nomenclature .....	37
2-Physico-chemical properties.....	38
2.1-Melting and boiling point.....	39
2.2-Vapor density .....	40
2.3-Flammability .....	41
3-Occurrence in living organisms and biological roles .....	41
3.1-In insects .....	41
3.2-In plants and algae.....	42
3.3-In cyanobacteria and other bacteria.....	42
4-Diverses applications, economic importance, and possible alternatives .....	43
4.1-Diverses applications and economic importance .....	43
4.2. Challenges and possible alternatives .....	44
C-Hydrocarbon-forming enzymes that use fatty acids and derivates.....	46
1-ADO.....	47
2-OleABCD .....	49
3-OleTje.....	51

4-Ols.....	52
5-CER1/CER3.....	52
6-UndA and UndB.....	56
7-FAP.....	58
7.1-Discovery.....	58
7.2-Structure and mechanism.....	61
7.3-Biochemical properties.....	62
7.3.1-Substrate specificity.....	62
7.3.2-Interfacial properties.....	64
7.4-Biodiversity.....	65
7.6-Use of FAP for biotechnological purposes.....	67
7.6.1-Use of wild type CvFAP.....	67
7.6.2-Use of CvFAP mutants.....	71
7.7-Fatty acid photodecarboxylase limitations.....	74
7.7.1-Inhibition/ Inactivation by light and thermostability.....	74
7.7.2-Physical state of the substrate.....	74
8-Summary of the different engineering techniques used to boost alkane production.....	75
References.....	78
Objectives of the thesis.....	86
A-Explore the biodiversity of FAPs.....	86
B-Contribute to a multidisciplinary study on the mechanism of FAP.....	86
C-Characterize the substrate specificity of FAP and optimize the photoconversion of SC and MC fatty acids substrates.....	87
Chapter 2: Exploring the biodiversity of fatty acid photodecarboxylase.....	88
Objectives, personal contribution, and main results.....	90
Article 1:.....	92
Fatty acid photodecarboxylase is an ancient photoenzyme that forms hydrocarbons in the thylakoids of algae.....	92
Chapter 3: Study of the mechanism of fatty acid photodecarboxylase.....	136
Objectives, personal contribution and main results.....	138
Article 2:.....	140
Mechanism and dynamics of fatty acid photodecarboxylase.....	140
Chapter 4: Substrate specificity of fatty acid photodecarboxylase and autocatalytic effect.....	156
Objectives, personal contribution and main results.....	158
Article 3:.....	160



Autocatalytic Effect Boosts the Production of Medium-Chain Hydrocarbons by Fatty Acid Photodecarboxylase .....	160
Chapter 5: Optimization of short-chain fatty acids photoconversion .....	190
Introduction: context and objectives .....	192
Results and Discussion .....	193
Effect of pH and substrate concentration.....	193
Evidence for CvFAP inactivation on some substrates.....	196
Light-dependent inactivation of FAP.....	198
Material and methods.....	200
Conclusions .....	200
References.....	201
General conclusion and Perspectives .....	208
Abstract.....	210



# List of figures

<b>Figure 1:</b> Examples of fatty acids structures and nomenclature.....	27
<b>Figure 2A:</b> Molecular state of fatty acids in an aqueous solution below their CMC.....	29
<b>Figure 2B:</b> Effect of chain length on the pKa of saturated fatty acids. ....	30
<b>Figure 2C:</b> Influence of the number and the configuration of unsaturations of C18 fatty acids on their pKa .....	30
<b>Figure 3:</b> Classification of biologically active lipids.....	31
<b>Figure 4:</b> Logarithmic plot of the molar solubility C1-C18 <i>n</i> -alkanoic acids and their sodium salts in water and methanol .....	32
<b>Figure 5A:</b> Schematic summary of the physical states formed by C10-C18 fatty acids in excess water as a function of pH and temperature (T). ....	35
<b>Figure 5B:</b> Schematic representation of fatty acid partitioning in an aqueous environment containing other lipids or fatty acid-binding molecules. ....	36
<b>Figure 6:</b> Examples of major types of hydrocarbon structures.....	38
<b>Figure 7:</b> Physical properties of alkanes (C1 to C50).....	40
<b>Figure 8:</b> Major distillation fractions from crude petroleum. ....	44
<b>Figure 9:</b> Biofuel production from different generations.....	45
<b>Figure 10:</b> Pathways of hydrocarbons synthesis originally proposed in the cyanobacterium <i>S.elongatus</i> .....	48
<b>Figure 11:</b> Proposed pathway for alkene biosynthesis from condensation of fatty acids.....	50
<b>Figure 12:</b> Proposed mechanisms for OleTje including fatty acid decarboxylation and $\alpha$ - or $\beta$ -hydroxylation and desaturation routes.....	51
<b>Figure 13:</b> Domain organization and proposed mechanism of the olefin synthase encoded by <i>Ols</i> .....	52
<b>Figure 14:</b> Synthesis of VLC Alkanes in INVSur4# Yeast Coexpressing Various Combinations of CER1 together with CER3, CYTB5-B, and LACS1 .....	54
<b>Figure 15:</b> Proposed Biochemical model in which Arabidopsis CER1 and CER3 proteins Act Synergistically with cytochrome b5 for very-long-chain Alkane Synthesis.....	56

<b><u>Figure 16:</u></b> Identification of UndB gene responsible of 1-undecene production by screening.	57
<b><u>Figure 17:</u></b> Schematic representation of different steps of fatty acids conversion to corresponding C-1, 1-alkene by UndA .....	58
<b><u>Figure 18:</u></b> Light dependency of the <i>C. variabilis</i> alkane synthase. Activity of the purified enzyme under successive light conditions.....	60
<b><u>Figure 19:</u></b> Structure of CvFAP. A. Overall architecture of the enzyme in complex with FAD and a fatty acid. ....	61
<b><u>Figure 20:</u></b> First Proposed photocycle of CvFAP based on time resolved spectroscopy approach.....	62
<b><u>Figure 21:</u></b> Interfacial Recognition Site and dipolar moment in CvFAP.....	65
<b><u>Figure 22:</u></b> Unrooted phylogenetic tree of the GMC oxidoreductase family (including FAPs). .....	66
<b><u>Figure 23:</u></b> Variation in FAP amino acids surrounding the carboxylate moiety of the fatty acid. ....	71
<b><u>Figure 24:</u></b> Designed mutants of CvFAP. The references of the original articles are in blue ..	72
<b><u>Figure 25:</u></b> Radical based deactivation of CvFAP .....	75
<b><u>Figure 26:</u></b> Fatty acid-derived hydrocarbon biosynthesis pathways.....	76
<b><u>Figure 27:</u></b> Schematic representation of the various metabolic and enzyme engineering techniques used to boost hydrocarbon production.....	77

# List of tables

<b><u>Table 1</u></b> : Systematic and trivial names of some fatty acids.....	28
<b><u>Table 2</u></b> : Temperature dependency of the solubility in water of C6-C18 linear saturated fatty acids.....	33
<b><u>Table 3</u></b> : Effect of chain length and temperature on critical micellar concentration. ....	36
<b><u>Table 4</u></b> : Cyanobacterial species selected for the subtractive genome approach used to identify ADO and AAR.....	47
<b><u>Table 5</u></b> : Substrate conversion by wt CvFAP. ....	63
<b><u>Table 6</u></b> : Summary of different studies made on CvFAP wild type protein. ....	70
<b><u>Table 7</u></b> : Resume of mutants designed on CvFAP and principal characteristics. ....	73



## List of abbreviations

**AAR:** Acyl-ACP Reductase

**ACP:** Acyl Carrier Protein

**ADO:** Aldehyde Deformylating Oxygenase

**AHDA:** 2-Hydroxydodecanoic Acid

**Arg:** Arginine

**Asn:** Asparagine

**ATP:** Adenosine Triphosphate

**BHDA:** 3-Hydroxydodecanoic Acid

**BLASTp:** Protein Basic Local Alignment Search Tool

**BP:** Boiling Point

**cdNA:** complementary Desoxyribonucleic Acid

**CER:** Eceriferum

**CMC:** Critical Micellar Concentration

**CO:** Carbon monoxide

**CoA:** Coenzyme A

***C. reinhardtii*:** *Chlamydomonas reinhardtii*

***C. variabilis*:** *Chlorella variabilis* NC64A

**CvFAP:** Fatty Acid Photodecarboxylase of *Chlorella variabilis* NC64A

**CW-EPR:** Continuous Wave-Electron Paramagnetic Resonance

**CYP:** Cytochrome P450

**DMSO:** Dimethyl Sulfoxide

**EPR:** Electron Paramagnetic Resonance

**FAD:** Flavin Adenine Dinucleotide

**FAP:** Fatty Acid Photodecarboxylase

**FRISM:** Focused Rational Iterative Site-Specific Mutagenesis

**GC-MS:** Gas Chromatography–Mass Spectrometry

**GMC oxidoreductase:** Glucose Methanol Choline oxidoreductase

**Gly:** Glycine

**IRS:** Interfacial Recognition Site

**LED:** Light Emitting Diode

**LC:** Long Chain

**LEL:** Lower Explosive Limit

**MC:** Medium Chain

**MIMS:** Membrane Inlet Mass Spectrometry

**MP:** Melting Point

**NADPH:** Nicotinamide Adenine Dinucleotide Phosphate

**NMR:** Nuclear Magnetic Resonance

**OD:** Optical Density

**OleTje:** Olefin Terminal synthase from *Jeotgalicoccus spp.*

**Ols:** Olefin Synthase

**PC:** Phosphatidylcholine

**pH:** Potential of Hydrogen

**PPT:** Phosphinothricin

**QM/MM:** Quantum Mechanics/Molecular Mechanics

**SC:** Short Chain

**sp:** species

**SPME:** Solid Phase Micro-Extraction

**TAG:** TriAcylGlycerol

**Tes:** Thioesterase

**TON:** Turnover Number

**VLC:** Very Long Chain

**IUPAC:** International Union of Pure and Applied Chemistry

**VOCs:** Volatil Organic Compounds



# **Chapter 1 : Bibliographic introduction and objectives of the thesis**



# Bibliographic introduction

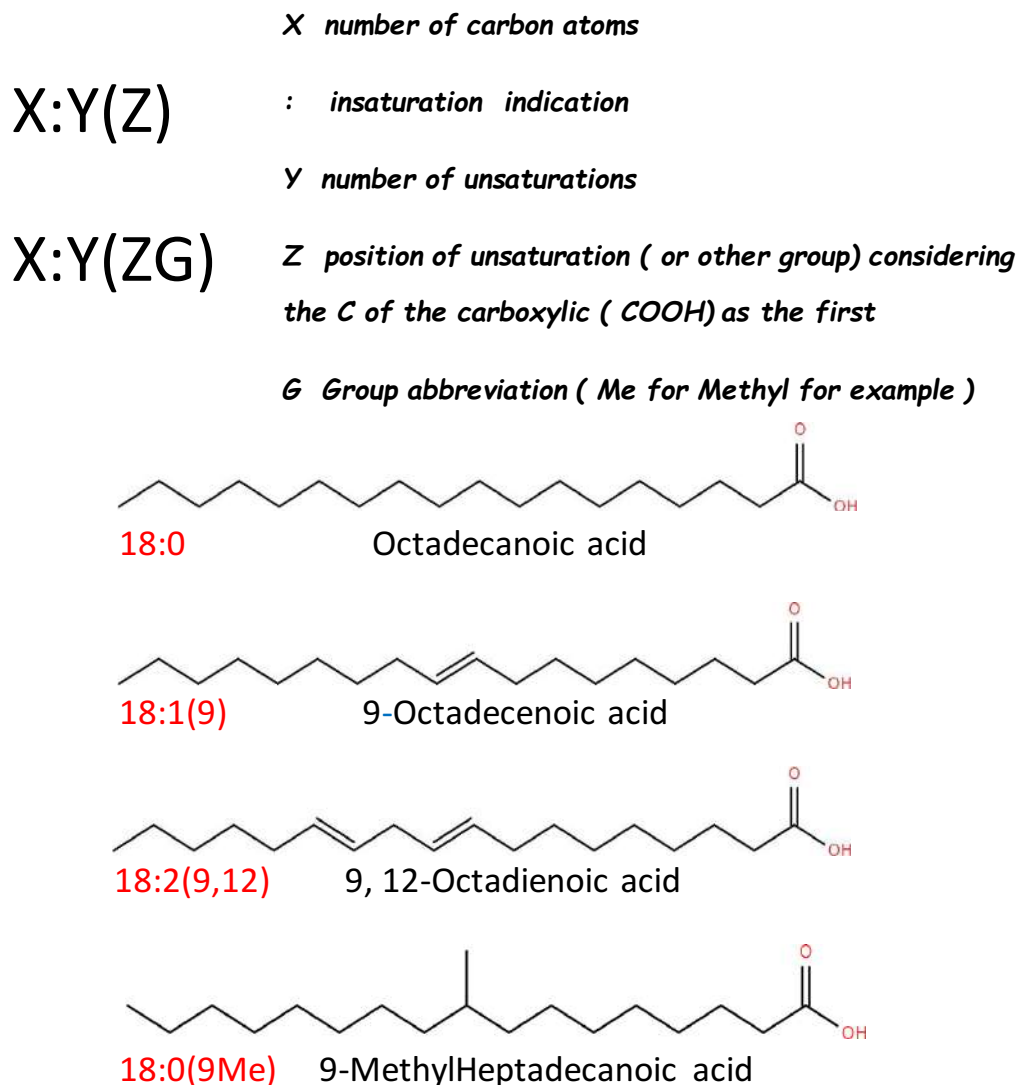
## A-Fatty acids

### 1-Structure and nomenclature

Fatty acids are organic compounds with a carboxylic acid group (-COOH) present at the end of an aliphatic chain (**Figure 1**). The fatty acid chain can be linear or branched with alkyl groups, saturated or unsaturated (with one or more C-C double or even triple bonds). Depending on the length of the carbon chain, we can distinguish short-chain (SC) fatty acids (1 to 6 C), medium-chain (MC) fatty acids (7 to 12C), long-chain (LC) fatty acids (12 to 20 C) and very-long-chain (VLC) fatty acids (>20C).

Fatty acids are ubiquitous building blocks of membrane lipids, being esterified to glycerol in glycerolipids, sterols in sterol esters and LC bases in sphingolipids (Liu and Benning [2013](#)). Fatty acids can also exist under a non-esterified form (“free fatty acids”) as small pools in cells or in biological fluids. Hundreds of fatty acids can be found in nature, with plants collectively synthesizing over 450 structurally distinct fatty acids (Ohlrogge et al. [1993](#)). The most common fatty acids have a linear carbon chain with an even number of total carbon atoms (most often 16 or 18). Among other functions in living cells, fatty acids are important energy sources (storage lipids) and are precursors for lipid mediators involved in cell signaling (e.g., oxylipins).

Concerning the nomenclature, fatty acids are named by a symbolic writing, a trivial name, or a systematic name. The symbolic and systematic names consider the carbon chain length, the presence of unsaturation or other groups (e.g., methyl) and their position whereas the trivial name is more a common name. Considering a fatty acid with 18(or 19) carbon atoms, containing or not double bonds and methyl groups, the rule of nomenclature is as shown in **Figure 1**. Trivial and systematic names for linear saturated fatty acids C1-C24 are shown in **Table 1**.



**Figure 1: Examples of fatty acids structures and nomenclature.** Abbreviations and systematic names are shown in red and black respectively. For each double bond of unsaturated fatty acids, there are two structural configurations called *cis* and *trans*. Fatty acids occurring naturally are mostly *cis* isomers. Only fatty acids with double bonds in *trans* configuration are represented here. When configuration is known, *cis* and *trans* isomers are noted c and t after the double bond positions (e.g., oleic acid is 18:1(9c) and its *trans* isomer is 18:1(9t)). In the *trans* configuration, the carbon chain is straight while the *cis* fatty acids are curved.

NUMBER OF CARBONS	SYSTEMATIC NAMES	TRIVIAL NAMES
1	Methanoic acid	Formic acid
2	Ethanoic acid	Acetic acid
3	Propanoic acid	Propionic acid
4	<i>n</i> -Butanoic acid	Butyric acid
5	<i>n</i> -Pentanoic acid	Valeric acid
6	<i>n</i> -Hexanoic acid	Caproic acid
7	<i>n</i> -Heptanoic acid	Enanthic acid
8	<i>n</i> -Octanoic acid	Caprylic acid
9	<i>n</i> -Nonanoic acid	Pelargonic acid
10	<i>n</i> -Decanoic acid	Capric acid
11	<i>n</i> -Undecanoic acid	Undecylic acid
12	<i>n</i> -Dodecanoic acid	Lauric acid
13	<i>n</i> -Tridecanoic acid	Tridecylic acid
14	<i>n</i> -Tetradecanoic acid	Myristic acid
15	<i>n</i> -Pentadecanoic acid	Pentadecylic acid
16	<i>n</i> -Hexadecanoic acid	Palmitic acid
17	<i>n</i> -Heptadecanoic acid	Margaric acid
18	<i>n</i> -Octadecanoic acid	Stearic acid
19	<i>n</i> -Nonadecanoic acid	Nonadecylic acid
20	<i>n</i> -Eicosanoic acid	Arachidic acid
21	<i>n</i> -Heneicosanoic acid	Heneicosylic acid
22	<i>n</i> -Docosanoic acid	Behenic acid
23	<i>n</i> -Tricosanoic acid	Tricosylic acid
24	<i>n</i> -Tetracosanoic acid	Lignoceric acid

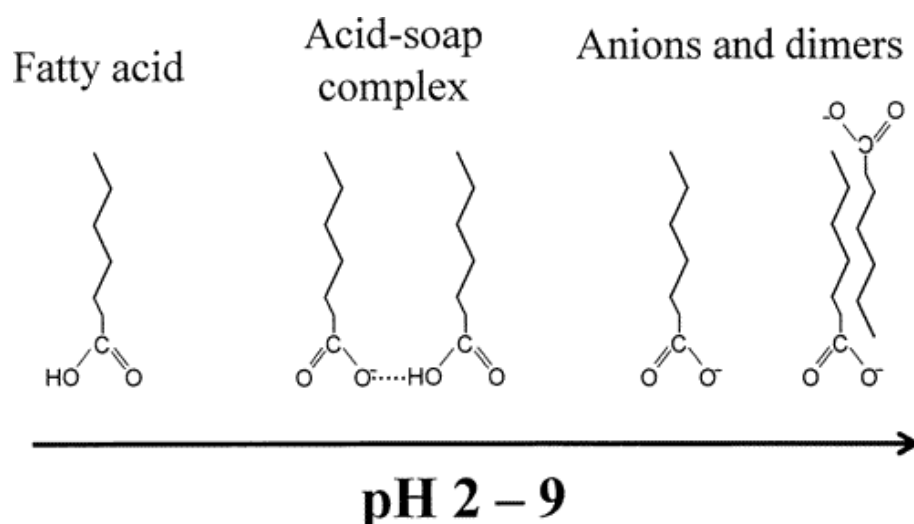
**Table 1: Systematic and trivial names of some fatty acids.**

Only linear saturated fatty acids (also called saturated straight-chain fatty acids or in the systematic nomenclature *n*-alkanoic acids) have been included because those are the only ones used in the experimental part of this manuscript.

## 2- Solubility and self-assembly in aqueous solutions

### 2.1-Ionization state

In aqueous solutions, fatty acids exist in two forms: the acid form R-COOH (protonated) and its conjugate base R-COO<sup>-</sup> (deprotonated). The proportion of these two forms is variable in solution according to the pH. The pK<sub>a</sub> defines the pH at which there is acid-base equilibrium, i.e., the R-COOH and R-COO<sup>-</sup> forms are in equal proportions. At pH < pK<sub>a</sub>, there is a predominance of the acid form and vice versa. The pH thus defines the coexistence of monomeric forms R-COO<sup>-</sup> and R-COOH as well as dimeric forms (**Figure 2A**). At low fatty acid concentration, i.e., below their critical micellar concentrations (CMC, see [section 2.2](#)) fatty acids are only under a soluble monomeric/dimeric form, when the pH increases, the R-COOH form progressively deprotonates into R-COO<sup>-</sup> and we have a mixture of these two forms as well as an acid-soap complex.

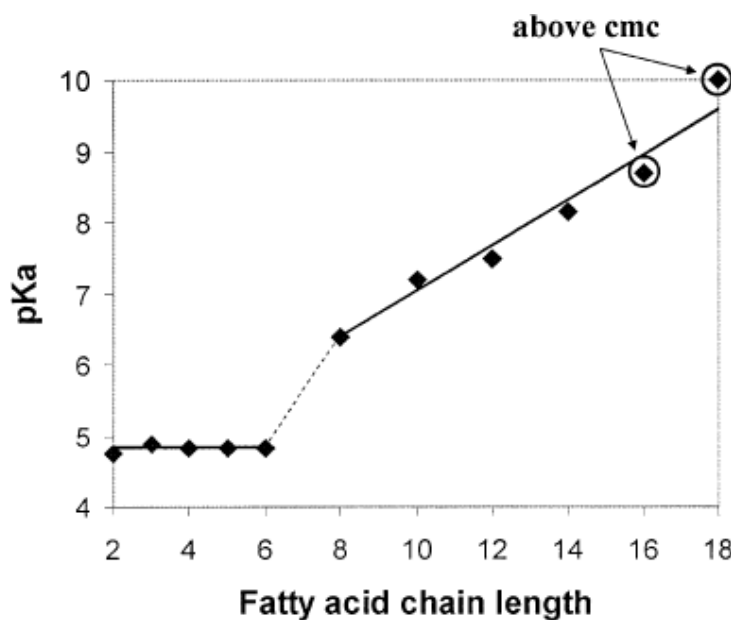


**Figure 2A: Molecular state of fatty acids in an aqueous solution below their CMC.** The different molecular states occurring from pH 2-9 are illustrated (From Kanicky and Shah [2003](#)).

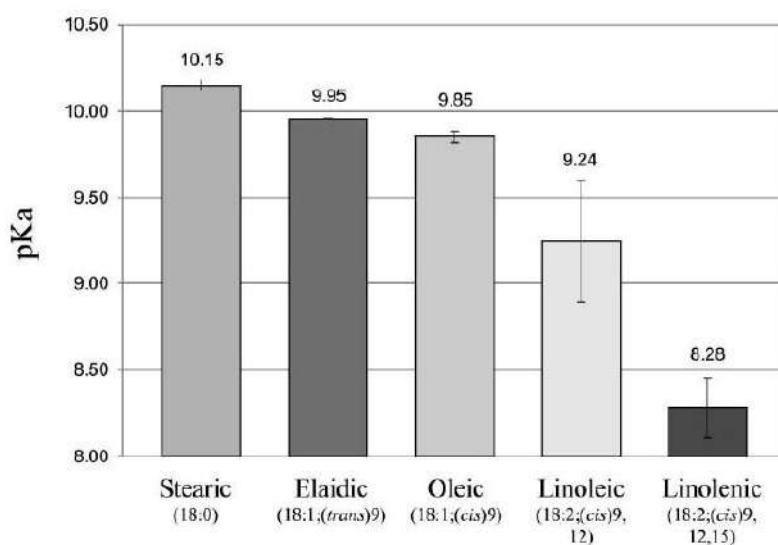
The apparent pK<sub>a</sub> of C2-C6 fatty acids is around 4.75, while the apparent pK<sub>a</sub> of C8-C18 fatty acids increases noticeably with chain length when measured with acid-base titration technique (Kanicky and Shah [2003](#)) (**Figure 2B**).

It should be noted that the number of unsaturations in the fatty acid chain also impacts the apparent pK<sub>a</sub> (Kanicky and Shah [2002](#)) (**Figure 2C**): pK<sub>a</sub> decreases with increasing double bonds. The fact that this decrease in pK<sub>a</sub> relates to the fatty acid melting temperature and the

area per molecule of the fatty acid monomolecular film spread at the air-water interface, suggests that as intermolecular distances in monolayer increases, pKa decreases due to reduced cooperation between adjacent carboxyl groups.



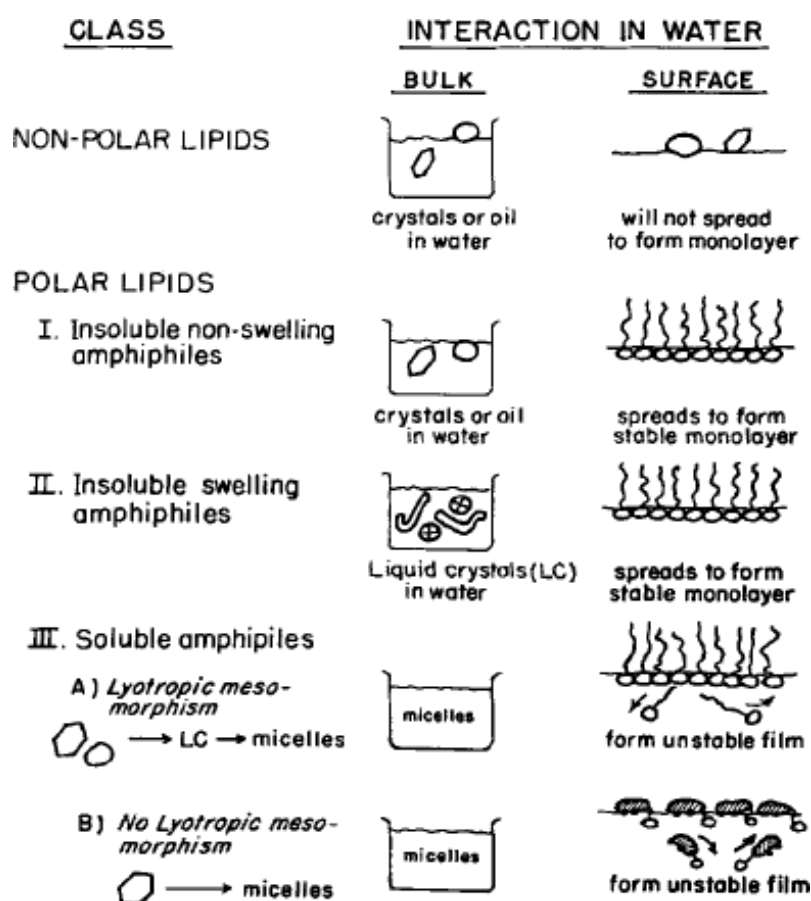
**Figure 2B: Effect of chain length on the pKa of saturated fatty acids.** Fatty acids were used at concentrations below CMC except for C16 and C18 fatty acids which were slightly above CMC. The pKa of sodium soaps were determined at 25°C $\pm$ 1 by acid-base titration (From Kanicky and Shah [2003](#)).



**Figure 2C. Influence of the number and the configuration of unsaturations of C18 fatty acids on their pKa.** The pKa were determined by acid-base titration at 20°C (From Kanicky and Shah [2002](#)).

## 2.2-Solubility in water

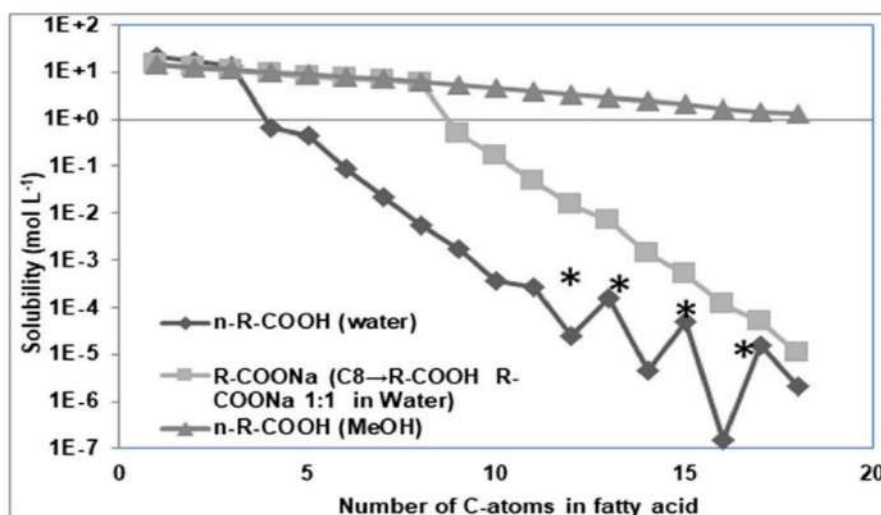
An important physical property of fatty acids is their amphipathic nature: the -COOH group is polar while the carbon chain is nonpolar. In the classical lipid classification of Small, long-chain fatty acids (such as oleic acid) are polar lipids of class II (insoluble nonswelling amphiphiles), meaning that they are virtually insoluble in the bulk in water and, depending on the temperature, form either crystals or oil in water (**Figure 3**) (Small, [1968](#)). However, long-chain fatty acids have a surface solubility because they can be spread at the air-water interface and form stable monomolecular films. By contrast, soaps of long-chain fatty acids (e.g., sodium oleate for oleic acid) are polar lipids of class IIIA (soluble amphiphiles of type A). Soaps can also form stable monolayers on the surface.



**Figure 3: Classification of biologically active lipids.** Column 1 gives the three major lipid classes. Column 2 depicts the physical state of the lipid in bulk aqueous system and on the surface of water. Wavy lines represent the aliphatic tails of lipid molecules, and open circles represent polar heads. Cross-hatched area molecules of Type IIIB represent the steroid nucleus of the bile salt (From Small [1968](#)).



The total carbon number of the chain (chain length) greatly influences the solubility of fatty acids in water and other solvents, i.e., the shorter the fatty acid, the more it will be soluble in water (**Figure 4**). Only SC fatty acids including methanoic, ethanoic and propanoic acids are miscible with water at very high concentrations. In methanol, by contrast, longer fatty acids are soluble even at higher concentration and C-atoms number.



**Figure 4: Logarithmic plot of the molar solubility C1-C18 *n*-alkanoic acids and their sodium salts in water and methanol.** Conditions were  $7 \leq \text{pH} \leq 9$  and a temperature of 25 °C. The values of *n*-undecanoate, *n*-tridecanoate, *n*-pentadecanoate, and *n*-heptadecanoate marked with asterisk (\*) have been averaged (From Ziogas et al. [2020](#)).

The pH (and thus the ionization state) as well as the presence of sodium or potassium ions will also influence fatty acid solubility (**Figure 4**). Fatty acids have a very low solubility in water under their protonated form (low pH) whereas they are relatively more soluble under their deprotonated form (higher pH). In the presence of sodium or potassium ions, the solubility increases compared to the corresponding undissociated acids in aqueous solution at 25°C, from 3 mM to 8 M for *n*-octanoic acid, and from 0.1 μM to 20 μM for *n*-hexadecanoic acid. The acid and sodium salt forms coexist in aqueous solution. For example, for *n*-octanoic acid, at neutral pH, the acid and sodium salt forms are found in a ratio of 1:1 (Cistola [1988](#)). The determination of the solubility of fatty acids is therefore complex because it is strongly dependent on multiple parameters, including the length of the carbon chain, the pH, the presence of ion (sodium and potassium) etc.

Additionally, it has been shown that bivalent cation also highly affects the solubility of fatty acids. In fact, they can interact with deprotonated fatty acids (especially with LC fatty acids) to form soap complexes, which decreases their solubility (Ayala-Bribiesca et al. [2017](#)).

It is well known that temperature greatly influences the solubility of organic matter in general. For fatty acids (SC, MC and LC), the higher the temperature, the higher the solubility (Ralston and Hoerr [1942](#), Romero et al. [2009](#)) (**Table 2**). The interdependence between temperature and solubility is shown by the van't Hoff equation:

$$\frac{\partial \ln X_2}{\partial (1/T)} = - \frac{\Delta_{\text{sol}} H_m^{\text{app}}}{R}$$

$X_2$  is the solubility of the solute in the saturated solution expressed in mole fraction,  $\Delta_{\text{sol}} H_m^{\text{app}}$  is the molar apparent enthalpy of solution,  $R$  is the universal gas constant ( $8.314 \text{ J} \cdot \text{mol}^{-1} \cdot \text{K}^{-1}$ ) and  $T$  is the absolute temperature. This equation is used to define the behavior of solubility as a function of absolute temperature of various compounds. It is assumed that for fatty acids, The temperature dependence of solubility is related to the mole fraction of the solute.

Number of C atoms	0.0°C		20.0°C		30.0°C		45.0°C		60.0°C	
	g/L	mM	g/L	mM	g/L	mM	g/L	mM	g/L	mM
6	8.64	74.482	9.68	83.448	10.19	87.845	10.95	94.397	11.71	100.948
7	1.90	14.615	2.44	18.769	2.71	20.846	3.11	23.923	3.53	27.154
8	0.44	3.056	0.68	4.722	0.79	5.486	0.95	6.597	1.13	7.847
9	0.14	0.886	0.26	1.646	0.32	2.025	0.41	2.595	0.51	3.228
10	0.095	0.552	0.15	0.872	0.18	1.047	0.23	1.337	0.27	1.570
11	0.063	0.339	0.093	0.500	0.11	0.591	0.13	0.699	0.15	0.806
12	0.037	0.185	0.055	0.275	0.063	0.315	0.075	0.375	0.087	0.435
13	0.021	0.098	0.033	0.154	0.038	0.178	0.044	0.206	0.054	0.252
14	0.013	0.057	0.020	0.088	0.024	0.105	0.029	0.127	0.034	0.149
15	0.0076	0.031	0.012	0.050	0.014	0.058	0.017	0.070	0.020	0.083
16	0.0046	0.018	0.0072	0.028	0.0083	0.032	0.010	0.039	0.012	0.047
17	0.0028	0.010	0.0042	0.016	0.0055	0.020	0.0069	0.026	0.0081	0.030
18	0.0018	0.006	0.0029	0.010	0.0034	0.012	0.0042	0.015	0.0050	0.018

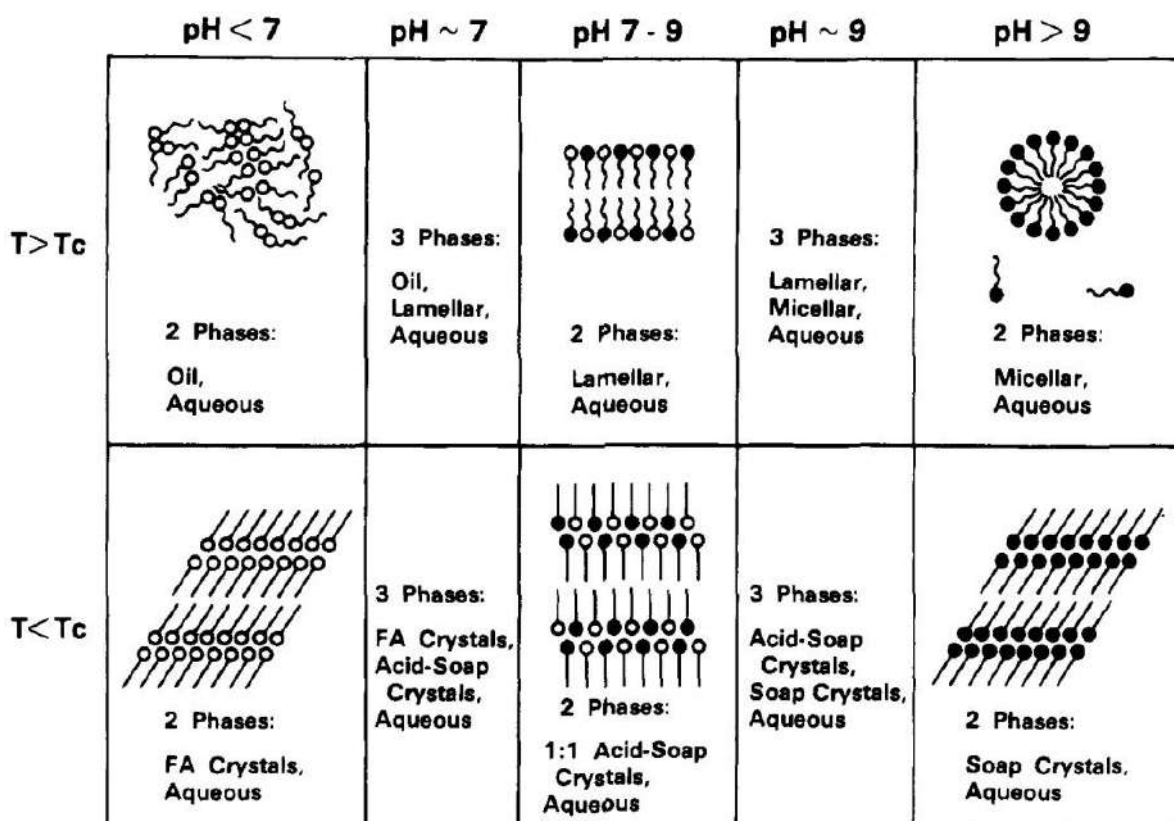
**Table 2: Temperature dependency of the solubility in water of C6-C18 linear saturated fatty acids.** Flasks containing acid-water samples were shaken in a constant temperature bath for 2-4 days. Solution was removed from the flasks by means of a small, fine sintered pyrex glass filter sealed in a glass tube. The solutions of C6 to C11 acids were titrated with standardized  $\text{Ba}(\text{OH})_2$

using phenolphthalein as the indicator. The less soluble samples from C12 acid upwards in the series, were titrated conductimetrically (Adapted from Ralston and Hoerr [1942](#)).

### 2.3-Self-assembly into higher order structures

When fatty acids are dissolved in water (or aqueous solutions in general), chain length of the fatty acid, ionization state (or pH values of the solution) and the presence of fatty acids binding molecules will govern the higher order structures into which they will self-organize beyond their solubility limit. Different structures can be found including lamellar liquid crystalline phase, hexagonal liquid crystalline phase, micellar solution, and monomers in aqueous solution (or true/ideal aqueous solution) (**Figure 5A**).

**Figure 5A** summarizes the results of a systematic study that used titration curve profiles determination coupled with physicochemical techniques, including X-ray diffraction and  $^{13}\text{C}$  NMR spectroscopy, to study the physical states of C10-C18 fatty acids in water (Cistola et al. [1988](#)). The authors were able to distinguish five structuration regions. These regions were divided considering the pH and melting temperature. At  $T > T_c$ , for pH values  $< 7$ , an oil and aqueous phase can be distinguished ; at pH  $\sim 7$ , a three-phase region containing oil, lamellar, and aqueous phases; between pH 7 and 9, a two-phase region containing lamellar and aqueous phases; at pH  $\sim 9$ , a three-phase region containing lamellar ,micellar, and aqueous phases; and at pH values  $> 9$ , a two-phase region containing micellar and aqueous phases.



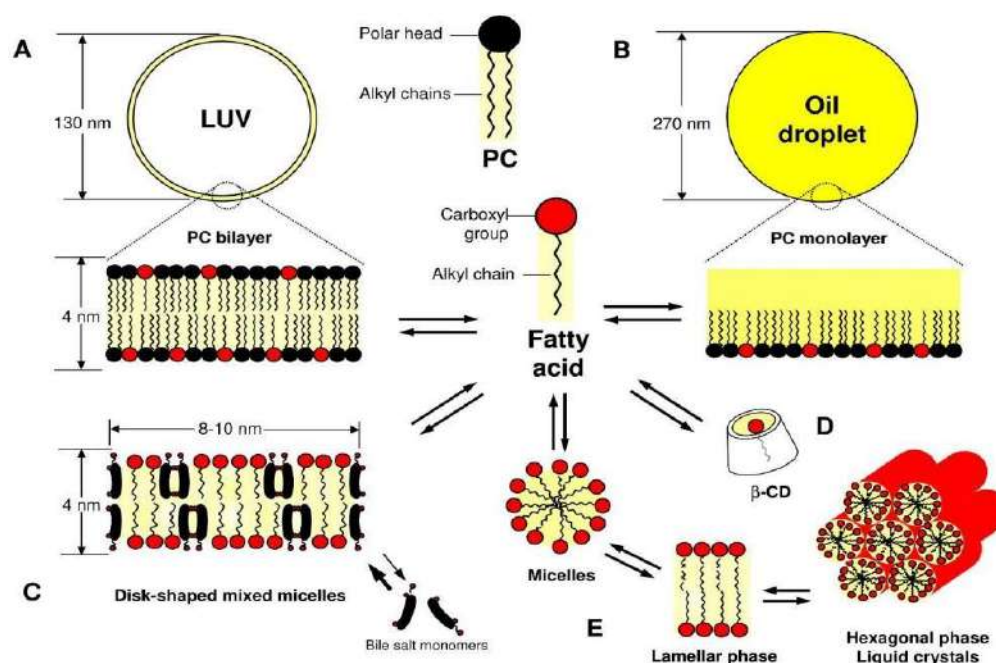
**Figure 5A:** Schematic summary of the physical states formed by C10-C18 fatty acids in excess water as a function of pH and temperature (T). The fatty acid hydrocarbon chain melting temperature in water (T<sub>c</sub>) is defined as the temperature beyond which there is a change of state from solid to liquid. (T<sub>c</sub> increases with the length of the fatty acid chain and differs for fatty acids, 1:1 acid-soaps and soaps). The aqueous phase is a saturated solution of fatty acids, acid-soap or soap and the concentration of these molecules varies with ionization state and fatty acid chain length. The closed circles represent ionized carboxylate groups and the open circles protonated carboxyl groups. The straight lines and the curved lines represent the ordered and disordered hydrocarbon chains of the fatty acids respectively (From Cistola et al. [1988](#)).

At a specific temperature, the fatty acid concentration below which micelles fall apart and the molecules are present in an ideal aqueous solution is defined as the critical micellar concentration (CMC). Usually, the CMC increases with increasing temperature. In contrast, CMC will tend to decrease with the carbon chain length (**Table 3**).

Number of C atoms	mmol/L	
	25°C	45°C
8	390	450
10	98	118
12	25.5	30.5
14	6.6	7.4
16	-	1.9

**Table 3: Effect of chain length and temperature on critical micellar concentration.** The CMC values were determined by refraction technique. (-) indicate non available data. (Adapted From Kleven, [1947](#)).

When other complex supramolecular structures containing lipids are present (liposomes, oil droplets, etc.), there is also a partitioning of fatty acids between the aqueous phase and the membrane-like structures (**Figure 5B**).



**Figure 5B: Schematic representation of fatty acid partitioning in an aqueous environment containing other lipids or fatty acid-binding molecules.** (A) Large Unilamellar Vesicles (LUVs); (B) Phosphatidylcholine (PC)-stabilized oil-in-water emulsion; (C) Mixed micelles with amphiphiles like bile salts; (D) Complex with fatty acid-binding molecules like  $\beta$ -cyclodextrin ( $\beta$ -CD); (E) Self-aggregation in the form of micelles, lamellar or hexagonal phases when  $\text{pH} > \text{pK}_a$ . Depending on parameters like pH, counter ions and concentration, the non-esterified fatty acids can be present in these different assemblies, some of which may be in equilibrium (From Aselmeyer et al [2021](#)).

## B-Fatty-acids derived hydrocarbons

### 1- Definition, structure, and nomenclature

Hydrocarbons are organic compounds composed only of hydrogen and carbon atoms. They are the simplest class of organic compounds and are the major component of several important substances such as natural gas, petroleum, and other fossil fuels.

Hydrocarbons are also synthesized by many living organisms. Two major types of hydrocarbons can be distinguished based on their biosynthetic origin: isoprenoid hydrocarbons (belonging to terpenes), which derive from the metabolism of the C<sub>5</sub> substrates dimethylallyl pyrophosphate and isopentenyl diphosphate (Oldfield and Lin [2012](#)) and fatty acid-derived hydrocarbons (Liu and Li [2020](#)). Here we will only deal with the latter (mostly linear alkanes and alkenes), which have mostly found and studied in some bacteria, in insects and in many photosynthetic organisms (plants, microalgae and cyanobacteria).

From the point of view of chemical structure, several groups of hydrocarbons can be distinguished (**Figure 6**):

1) Aliphatic hydrocarbons, which include linear, branched, or cyclic (non-aromatic) hydrocarbons. Depending on the type of C-C bonds, we can distinguish (n is the number of carbon atoms):

(i) Alkanes: composed of saturated chains having only simple bonds between carbon atoms, their general formula is C<sub>n</sub>H<sub>(2n+2)</sub>. Methane (CH<sub>4</sub>) represents the most basic alkane example.

(ii) Alkenes: unsaturated hydrocarbons that have simple bonds between carbons atoms but also a double bond; their general formula is C<sub>n</sub>H<sub>2n</sub> and

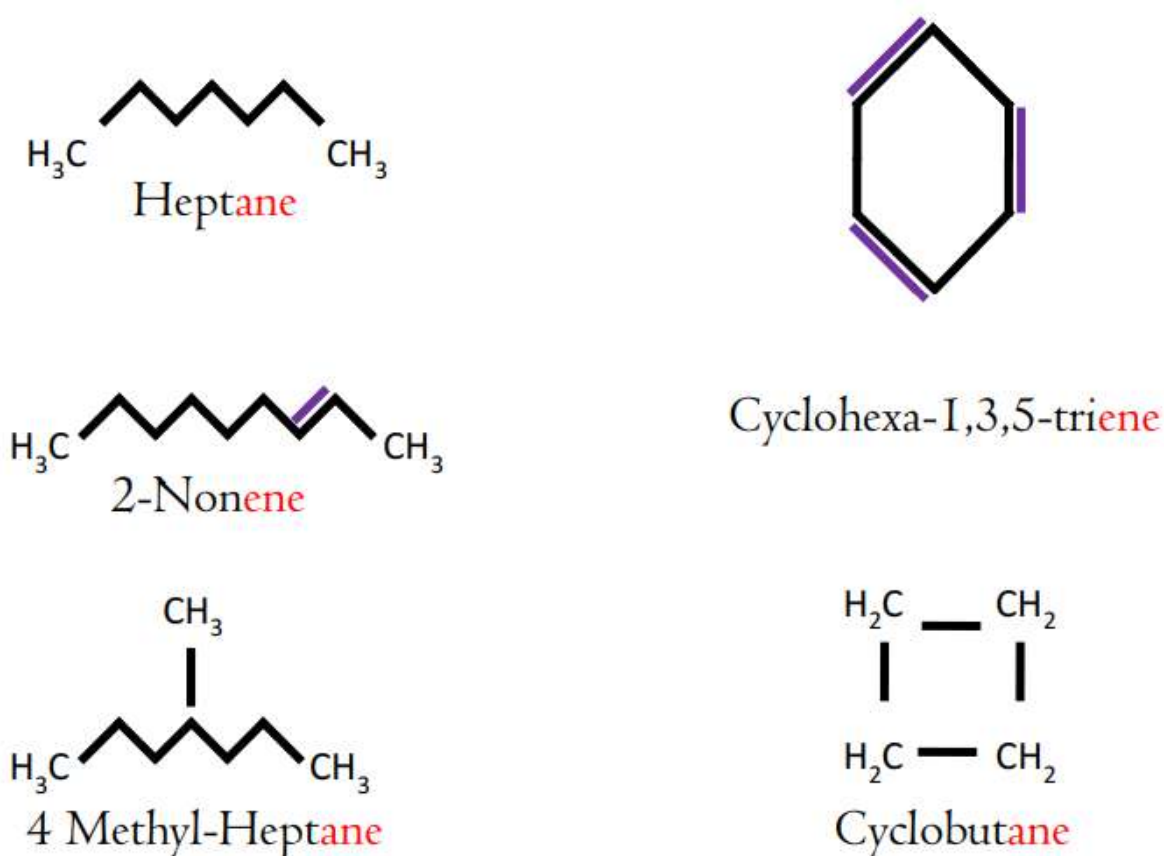
(iii) Alkynes: containing a triple bond; their general formula is C<sub>n</sub>H<sub>(2n-2)</sub>.

2) Cyclic aromatic hydrocarbons: this class is characterized by the presence of atoms with alternating double bonds also called aromatic ring. The most common example of an aromatic hydrocarbon is cyclohexa-1,3,5-triene (commonly called benzene, C<sub>6</sub>H<sub>6</sub>). (**Figure 6**)

The nomenclature of hydrocarbons follows some standard ruled from the International Union of Pure and Applied Chemistry (IUPAC). The IUPAC nomenclature system for hydrocarbons is based principally on two parameters including: (i) the number of carbon atoms in the longest chain and (ii) the type and the position of any functional groups or double/triple bonds in the chain. For alka(e)(y)nes, the name is based on the number of carbon atoms in the molecule. First, If the hydrocarbon is linear ("normal") the name should be preceded by the prefix *n*-. To indicate the number of carbon atoms, prefixes "meth-"(1C), "eth-"(2C), "prop-"(3C), "but-"(4C), and so on, are used. Then follow the suffixes "-ane","-ene" or "-yne" to indicate that the molecule is an alkane, alkene, or alkyne respectively. For example, a linear

alkane with seven carbon atoms is named "*n*-heptane". Additionally, for alkenes and alkynes, the position of the double or triple bond is indicated by a number that corresponds to the lowest-numbered carbon atom involved in the bond. For example, the hydrocarbon with nine carbon atoms and a double bond between the second and third carbon atoms from one end is named "2-nonene" and not "7-nonene". Hydrocarbons containing alkyl groups follow the same nomenclature rules as for hydrocarbons with double or triple bonds.

For aromatic hydrocarbons, the name is based on the prefix *cyclo* and the number of carbon atoms in the ring, followed by the suffix "-ene" to indicate that the molecule contains a double bond.



**Figure 6: Examples of major types of hydrocarbon structures.** The prefixes are in black and the suffixes in red color.

## 2- Physico-chemical properties

Like fatty acids, Hydrocarbons are insoluble in water and highly miscible with organic solvents (non-polar) like ether. Hydrocarbons properties are mostly influenced by the carbon chain length and the presence of additional groups including alkyls (methyl) and unsaturations. From



a chemical point of view, alkanes are considered as compounds with a very low reactivity. This low reactivity is possible since the bonds between carbon atoms on one hand and between carbon atoms and hydrogen atoms on the other hand are very strong and energetic (McQuarrie et al. [2012](#)). However, although almost inert, under the influence of various factors such as temperature and pressure, hydrocarbons can have a variable appearance and structuring.

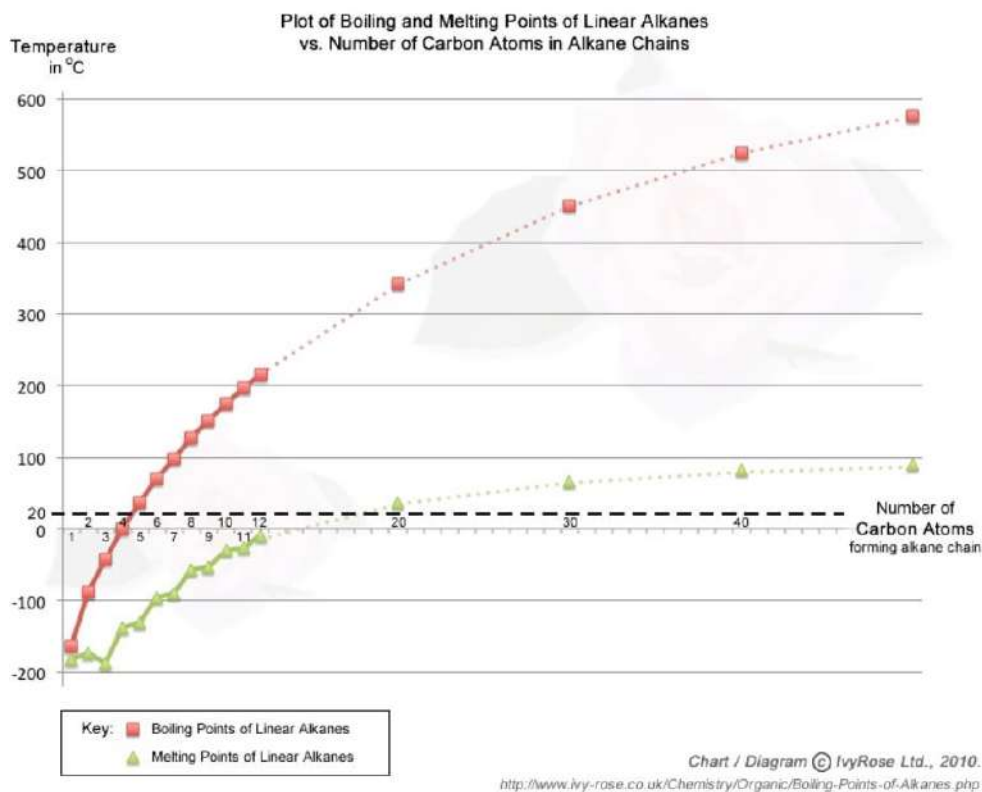
### **2.1-Melting and boiling point**

Under certain specific temperature and pressure conditions and depending on the chain length, for a liquid hydrocarbon, the vapor pressure of the liquid is equivalent to the surrounding pressure, this is called the boiling point. From a practical point of view, the boiling point defines the amount of energy required to ensure a transition state from liquid to gas. Thus, since the inter and intra molecular bonds contained in the hydrocarbon chains are very strong, the boiling point increases with the length of the carbon chain (**Figure 7**); for example, pentane and dodecane have boiling points of 36°C and 216 °C respectively. The change in physical state of hydrocarbons is possible because, when liquid hydrocarbons are brought to their boiling temperature, the kinetic energy of the particles is strong enough to break the intermolecular bonds allowing each individual molecule to break free and occupy more space in the gas phase. The kinetic energy in relation is defined by the formula  $E=1/2mv^2$ , m and v representing the molecular mass and the velocity of the molecule respectively.

The melting point is defined as the temperature at which a molecule changes state from solid to liquid. For hydrocarbons, the melting point follows the same trend as the boiling point. As a rule, for linear hydrocarbons, the higher the carbon chain length, the higher the melting point. (**Figure 7**).

Volatility of hydrocarbons is also important and is one of the parameters that leads to the possibility to use them for different applications. For instance, higher volatility hydrocarbons may be preferred for use as fuels because they evaporate more easily while lower volatility hydrocarbons may be preferred for use as lubricants or solvents because they remain longer at liquid state. This property is closely related to the boiling point since it defines the tendency of hydrocarbons to evaporate or vaporize at a given temperature and pressure.





**Figure 7: Physical properties of alkanes (C1 to C50).** Color codes are indicated on the figure.

Source: <https://www.ivyroses.com/Chemistry/Organic/Boiling-Points-of-Alkanes.php>

## 2.2-Vapor density

Vapor density is a measure of the density of a gas or vapor compared to the density of air at a given temperature and pressure. It is a useful parameter for characterizing the behavior of hydrocarbons, particularly in relation to their potential to form flammable or explosive mixtures with air (flammability is explained in the paragraph below). The principal factor that determines the vapor density is the molecular weight. Generally, hydrocarbons with higher molecular weights have higher vapor densities, meaning that their vapors are denser than air. In contrast, hydrocarbons with lower molecular weights have lower vapor densities and their vapors are less dense than air. The vapor density of a hydrocarbon can be determined experimentally by measuring the weight of a known volume of the gas or vapor and comparing it to the weight of an equal volume of dry air at the same temperature and pressure. The resulting ratio of the weights is the vapor density of the hydrocarbon. For example, hexane molecular weight is 86 (g/mol) and the molecular weight of air is 29 (considering the proportion of gases). The quotient of 86/29 is 2.9. Hexane vapors density, then, is 2.9.

### **2.3-Flammability**

Hydrocarbons are generally flammable substances due to their high carbon and hydrogen content. Flammability can be defined as the measure of a substance's ability to ignite and sustain combustion in the presence of oxygen. The flammability of a hydrocarbon depends on several factors, including its molecular structure, boiling point/vapor density (described above) and concentration in air (Affens and McLaren [1972](#), Albahri [2003](#))

In general, hydrocarbons with lower boiling points and lower molecular weights are more flammable than those with higher boiling points and higher molecular weights. The lower boiling point hydrocarbons tend to evaporate more easily this increase their concentration in air and makes them more likely to ignite. Additionally, the simpler, more linear hydrocarbons are typically more flammable than their more complex counterparts. For example, when hydrocarbons are mixed with air at or above their lower explosive limit (LEL), a spark or other ignition source can ignite the mixture, resulting in a fire or explosion. Similarly, hydrocarbons with high vapor pressure may release flammable vapors even at low temperatures, which can create a fire hazard. The flammability of hydrocarbons is an important consideration in many applications, especially in fuel production, transportation, and storage, as well as in petrochemical processing and environmental monitoring. It is critical to understand the flammability characteristics of different hydrocarbons to ensure safe handling and storage, as well as to develop appropriate fire prevention and mitigation strategies.

## **3-Occurrence in living organisms and biological roles**

As previously mentioned, the biosynthetic pathways of fatty acid-derived hydrocarbons are present in various living organisms such as bacteria, plants, algae, and insects (Zuo et al. [2019](#)). The biological roles of these hydrocarbons have been mostly studied in plants and insects and is fairly well known in these organisms. In contrast, the role of these compounds is still poorly understood in unicellular organisms.

### **3.1-In insects**

In insects, a variety of C<sub>11</sub>-C<sub>43</sub> hydrocarbons can be produced by some secretory cells called oenocytes and are mostly localized in the cuticle (Hackman, [1984](#), Lockey [1988](#)). These hydrocarbons are implicated in the adaptation to the desiccating atmosphere of terrestrial life by protecting insects from water loss (Hadley [1984](#)). Apart from this protective role, some

hydrocarbons are pheromones implicated in recognition between species (Howard [1993](#)) or in sexual reproduction (Carlson et al. [1984](#), Scott [1986](#), Scott et al. [1988](#)).

### **3.2-In plants and algae**

The role of hydrocarbons in plants is complex and multifaceted, with important implications for plant growth and survival. Very-long-chain-alkanes C27-C35 are major components of the cuticle, an extracellular hydrophobic layer that seals the epidermis of the aerial plant parts and plays an essential role in plant adaptation and survival of plants to the desiccating terrestrial environment (Jetter and Kunst [2008](#)). Studies have shown that the amount and composition of cuticular alkanes can vary widely depending on plant species, different organs within the same plants and in response to environmental conditions.

In algae, it has been known for a long time that the green colonial microalga *Botryococcus braunii* excretes C27 *n*-alkadienes and trienes, presumably derived from VLC Fatty acids (Metzger and Largeau [2005](#)). These hydrocarbons are excreted together with other hydrocarbon terpenes and are thought to have an important role in biofilm formation. Based on this observation, a study was conducted to look for fatty acid-derived hydrocarbons in model green microalgae such as *C. reinhardtii* and *C. variabilis* (Sorigué et al. [2016](#)). This study demonstrated that C15-C17 alka(e)nes are found in the cells and as volatile compounds in the gas phase of the cultures and showed that they were indeed synthesized from fatty acids. It was later shown that a wide range of volatile organic compounds (VOCs), including alkanes, alkenes and halohydrocarbons are produced in various algal species (Zuo et al. [2019](#)). The function of hydrocarbons in algae is likely to be linked to photosynthetic membranes (Moulin et al. [2021](#)).

### **3.3-In cyanobacteria and other bacteria**

Various species of cyanobacteria and algae contain linear hydrocarbons derived from fatty acids, with heptadecane being the main compound in all species (Han et al [1968](#)). A comparative study of the metabolites of 11 cyanobacterial species (which were previously called “blue-green algae”) has shown that they produce alkanes of chain length between 15 and 18 (for 8 of the 11 species). The predominance of these algal species in bloom type environments allowed to conclude that these algae are geological source material for C15 to C18 hydrocarbons (Winters et al. [1969](#)). Because of their occurrences in diverse species, the function of hydrocarbons in cyanobacteria has been studied but it remains unclear. Hydrocarbon mutants of *Synechococcus* were shown to be impacted in growth due to a decreased efficiency

of the photosynthesis especially in cold conditions (Herman and Pflieger, [2014](#), Lea-Smith et al. [2016](#)). Another study revealed that hydrocarbons are essential for optimal cell size, division, and growth of cyanobacteria. In the same study it was also shown that hydrocarbons were located in plasma and thylakoids membranes and are responsible for membrane flexibility. Additionally, a role of hydrocarbons in salt tolerance of cyanobacteria has also been suggested (Yamamori et al. [2018](#)).

Although the presence of LC hydrocarbons in other bacteria has been demonstrated (*Micrococcus luteus*, *Xanthomonas campestris*, *Pseudomonas* and *Shewanella oneidensis*) their biological role is still unclear and faces many speculations. The only functional study concerns *Shewanella oneidensis*. In this bacterium, the depletion of hydrocarbons has been shown to result in a decrease of growth in cold conditions, suggesting a role of bacterial hydrocarbons in adaptation to temperature variations (Sukovich et al. [2010](#)).

## **4-Diverses applications, economic importance, and possible alternatives**

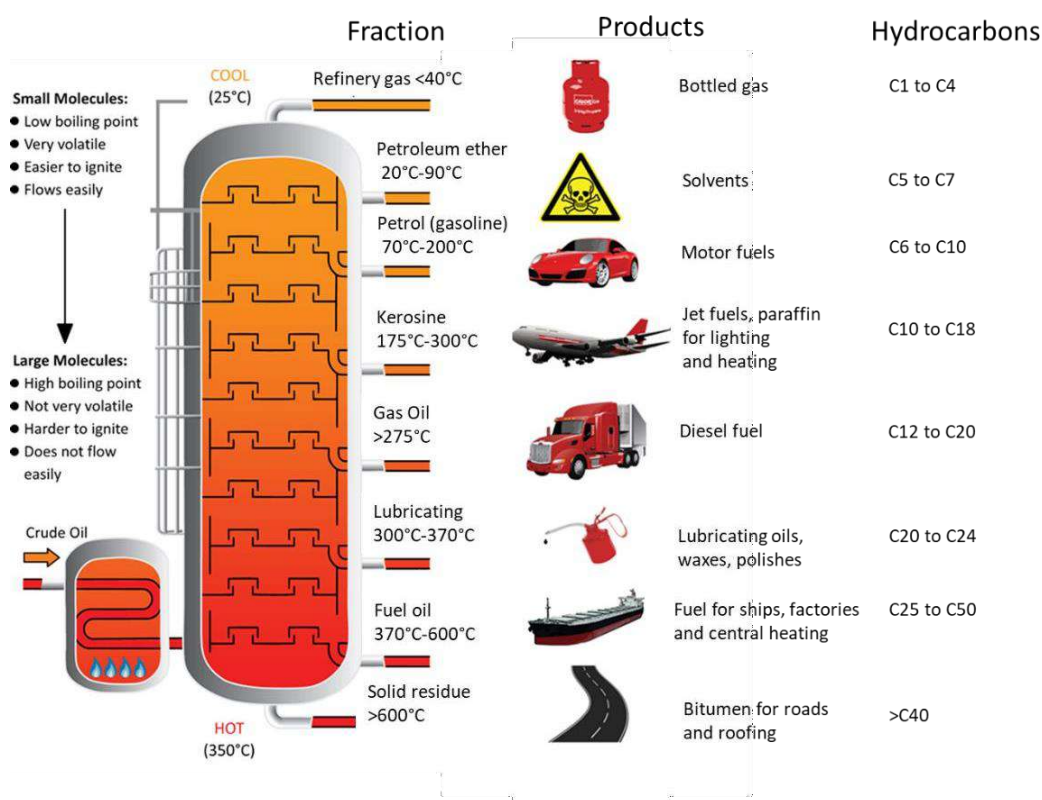
### **4.1-Diverses applications and economic importance**

Hydrocarbons are used in various applications ranging from transportation and heating to manufacturing and agriculture. They are one of the most important primary energy sources of the world and come almost exclusively from fossil resources: crude oil, natural gas, and coal (Hall et al. [2003](#)). Based on boiling point of its components, crude oil is split into different fractions by industrial distillation (**Figure 8**). Fractions harvested from distillation columns remain a mixture of hydrocarbons with various applications.

The most common use of hydrocarbons is related to the transport industry. Depending on their nature and chain length, they are used to power most vehicles (including cars, airplanes, and trains) in which they can be qualified with different designations such as gasoline, diesel, kerosene etc. Hydrocarbons are also used extensively in the heating sector. Natural gas, propane and butane are used to heat homes, buildings, and industrial processes. Another significant application of hydrocarbons is in the manufacturing sector. Hydrocarbon-based chemicals are used in the production of a wide range of products, including plastics, fertilizers, pesticides, pharmaceuticals, and synthetic fibers. Thanks to their textural, volatility and organic properties, some hydrocarbons (mostly C7-C22) can be used in cosmetics formulations to provide some specific aspect to the products or to serve as solvents for other organic compounds.

The economic importance of hydrocarbons cannot be overstated. The hydrocarbon industry is one of the largest and most lucrative in the world, generating trillions of dollars in revenue

each year. Hydrocarbons play a significant role in international trade since many countries depend heavily on hydrocarbon imports to meet their energy needs, making energy security a critical concern for many governments. The Saudi Arabia, Russia, and United states, are major exporters (because they have largest reserves) of oil and gas, generating significant revenue.



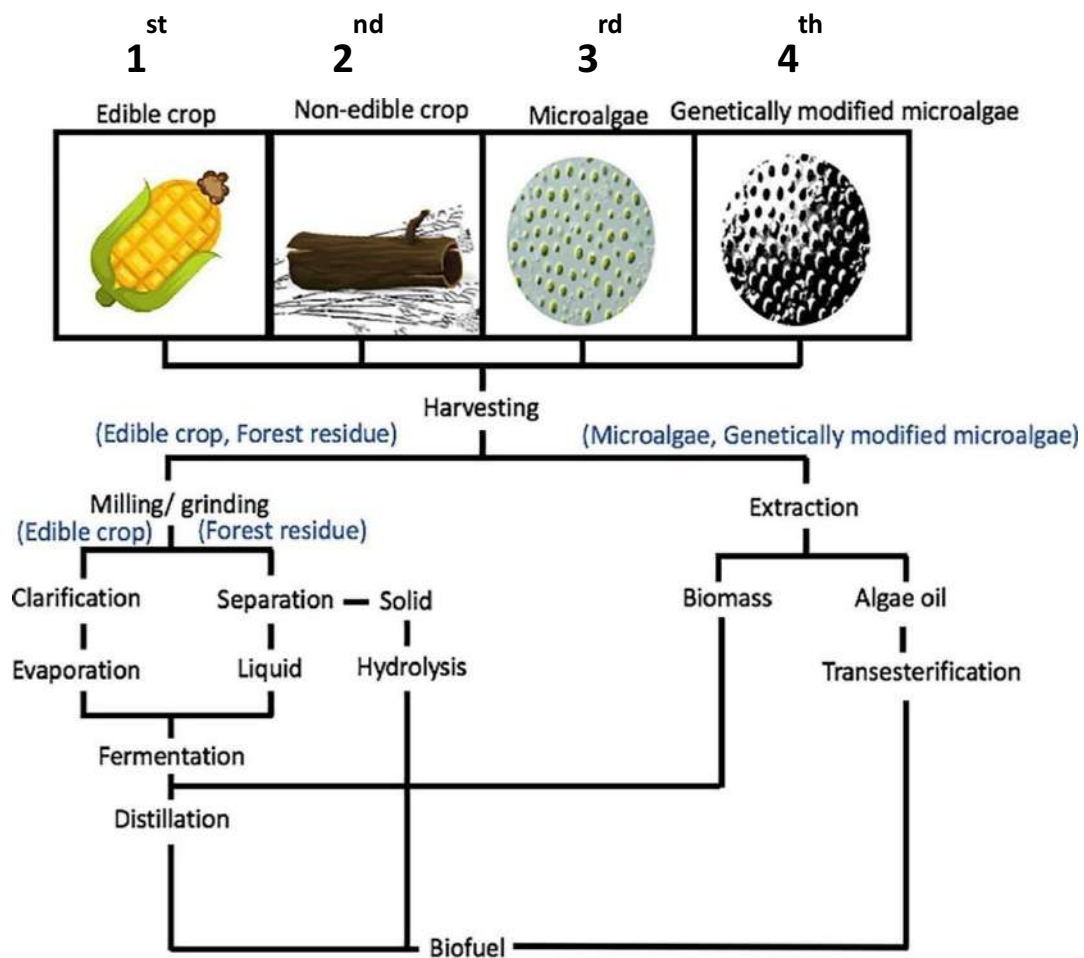
**Figure 8: Major distillation fractions from crude petroleum.** Image modified from <http://science-resources.co.uk/>. Fractions names, data on temperatures and size of hydrocarbons are from Eneh [2011](#).

## 4.2. Challenges and possible alternatives

Like developed before, hydrocarbons are important for human activity through their utility for different applications. Unfortunately, the use of hydrocarbons faces some inconveniences. Mostly from fossils fuels, their combustion releases greenhouse gases, which contribute to climate change and air pollution. The extraction and transportation of hydrocarbons can also have negative impacts on local environments and communities, including habitat destruction, water pollution, and noise pollution.

As the needs of hydrocarbons are growing all over the world, to address these concerns, there is increasing interest in developing alternative, sustainable, and renewable sources of fuels. Four generations of biofuels have then been developed (**Figure 9**).

(i)-First-generation biofuels are made from edible crops such as corn, sugarcane, and soybeans (Naik et al., [2010](#)). These crops are grown specifically for biofuel production, which can lead to competition with food production and potential land-use changes. First-generation biofuels include ethanol and biodiesel and are currently the most widely produced biofuels worldwide. Ethanol is produced through the fermentation of sugars found in corn, sugarcane, and other crops. The resulting ethanol is then blended with gasoline and used as a transportation fuel. Biodiesel are made from oil (triacylglycerols) extracted from oilseeds or oil palm fruits by transesterification with methanol (or ethanol). Therefore, from a chemical point of view, biodiesel is fatty acid methyl esters (or ethyl esters). These molecules are structurally similar to hydrocarbons of diesel but not identical: biodiesel is an oxygenated fuel. First-generation biofuels have been criticized because of their impact on food prices and the environment justifying the development of other generations of biofuel (also called “advanced biofuels”).



**Figure 9: Biofuel production from different generations.** (Adapted from Aron et al. [2020](#))



(ii)-In contrast to first generation biofuels, second-generation biofuels (Naik et al. [2010](#)) are made from non-edible biomass such as agricultural and forestry waste, and energy crops that do not compete with food production. The use of macroalgae has also started to be studied (Kumar et al. [2021](#)). Second-generation biofuels are produced through gasification, pyrolysis or fermentation. These methods can produce a wider range of biofuels, including cellulosic ethanol. Second generation biofuels also comprise fuels with hydrocarbons: renewable diesel is composed of hydrocarbons and is made from vegetable oil, like traditional biodiesel, but primarily from waste and residues. In the production process, impurities are removed from the raw materials which are then hydrotreated at a high temperature. The outcome is a fuel that has chemical composition very similar with fossil diesel. However, this alternative source of hydrocarbons would require an important quantity of raw material for the processes to be economically predominant.

(iii)-Third-generation biofuels are made from microorganisms such as microalgae (Wijffels and Barbosa [2010](#)) or bacteria and yeasts (Keasling et al. [2021](#)). Third-generation biofuels have the potential to produce yearly more biodiesel per surface area occupied than first and second-generation biofuels. The production of third-generation biofuels is still in the early stages of development, and commercialization has not yet been achieved. Algae can grow quickly, can commonly accumulate up to 50% oil (triacylglycerols) per dry weight and can be grown in various environments, including wastewater and saline water. Algae can also be harvested continuously, making it a more productive source of biofuel than crops. However, algae-based biofuels are still not economically viable, mostly due to the high cost of growing and harvesting microalgae at an industrial scale. Transgenic microalgae are sometimes called the 4<sup>th</sup> generation biofuels. Use of non-photosynthetic microbes to produce biofuels would require progress in yields through engineering metabolic pathways (especially to produce volatile compounds that can be easily recovered from the microbial cultures).

### **C-Hydrocarbon-forming enzymes that use fatty acids and derivatives.**

The important roles of hydrocarbons in the economy (fuel, chemistry etc.) have triggered many studies aiming at the identification of their biosynthetic pathways and the enzymes involved in the synthesis of these compounds have aroused great interest within the community of biotechnologists. The main questions that have been addressed in these studies are:

- a- Which species can produce hydrocarbons?
- b- What are the enzymes responsible for hydrocarbon synthesis?

- c- What are the mechanisms of hydrocarbon-producing enzymes?
- d- How can we improve the turnover of hydrocarbon-producing enzymes and modify their substrate specificity?

In the following sections, we will address some of these questions for the main hydrocarbon-forming enzymes in order of their discovery, from the least to the most recent. Only some of the studied hydrocarbon-forming enzymes will be developed here. However, a summarized study of all the enzymes known to date will be presented.

## 1-ADO

In the 90's, a study showed that the cyanobacterial mats of hot springs in USA and New Zealand contained long-chain hydrocarbons (C15-C20) (Shiea et al. [1990](#)). The recurrent detection of these same compounds in cyanobacteria cultures allowed to identify these photosynthetic microorganisms as a plausible biological source of hydrocarbons. With the advent of genome sequencing, about 50 cyanobacterial genomes were already available in the early 2000s. In 2010, Schirmer and coworkers showed the presence of heptadecane and pentadecane in the cultures of 10 species of cyanobacteria out of 11 cyanobacteria species with known genomes that were initially selected (**Table 4**). By sequence alignment (genomes of the 10 hydrocarbon-producing cyanobacteria) and subtraction (of the genome of the non-hydrocarbon-producing cyanobacteria), 17 potential candidate genes were identified, 10 of which already had assigned functions.

Cyanobacterium	Genome size (Mb)	Hydrocarbons previously reported	Alkanes observed in this study
<i>Synechococcus elongatus</i> PCC7942	2.7	Yes*	Heptadecane, pentadecane
<i>S. elongatus</i> PCC6301	2.7	Yes*	Heptadecane, pentadecane
<i>Synechocystis</i> sp. PCC6803	3.5	Not reported	Heptadecane
<i>Prochlorococcus marinus</i> CCMP1986	1.7	Not reported	Pentadecane
<i>Anabaena variabilis</i> ATCC29413	6.4	Yes†	Heptadecane, methyl-heptadecane
<i>Nostoc punctiforme</i> PCC73102	9.0	Not reported	Heptadecane
<i>Gloeobacter violaceus</i> PCC7421	4.6	Not reported	Heptadecane
<i>Nostoc</i> sp. PCC7120	6.2	Not reported	Heptadecane, methyl-heptadecane
<i>Cyanothece</i> sp. PCC7425	5.7	Not reported	Heptadecane
<i>Cyanothece</i> sp. ATCC51142	4.9	Not reported	Pentadecane
<i>Synechococcus</i> sp. PCC7002	3.0	Yes*	No alkanes observed

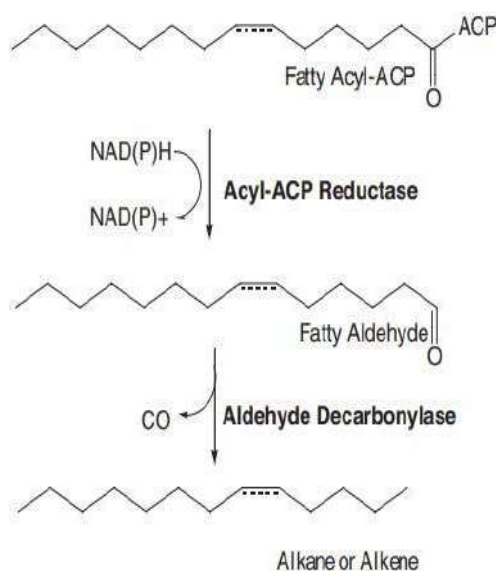
**Table 4: Cyanobacterial species selected for the subtractive genome approach used to identify ADO and AAR.** All species had fully sequenced genomes. Note that this approach was possible because the authors managed to identify one species that did not have alkanes (PCC7002) and that literature reports for this species suggested it did contain alkanes (From Schirmer et al. [2010](#)).



Knockout in cyanobacteria of the two remaining candidate genes that were in tandem and heterologous expression in *E. coli* of the two proteins encoded allowed the identification of the two proteins necessary and sufficient for synthesis of C15-C17 *n*-alkanes and *n*-alkenes in the cyanobacterium *Synechocystis sp.* PCC6803: an acyl ACP reductase (AAR) using NAD(P)H as a cofactor and an “aldehyde decarbonylase” belonging to the superfamily of ferritin-like di-iron proteins (Schirmer et al. [2010](#)) (**Figure 10**).

When the cyanobacterial aldehyde decarbonylase was initially named, it was assumed that the coproduct of the reaction was carbon monoxide (CO). However, this enzyme was later shown to produce formate- and not carbon monoxide- as a coproduct (Warui et al. [2011](#)) and it was also demonstrated to require dioxygen for the reaction (Li et al. [2012](#)). Therefore, it was renamed Aldehyde-Deformylating Oxygenase (ADO).

Heterologous production of ADO in different hosts has been done but resulted in the



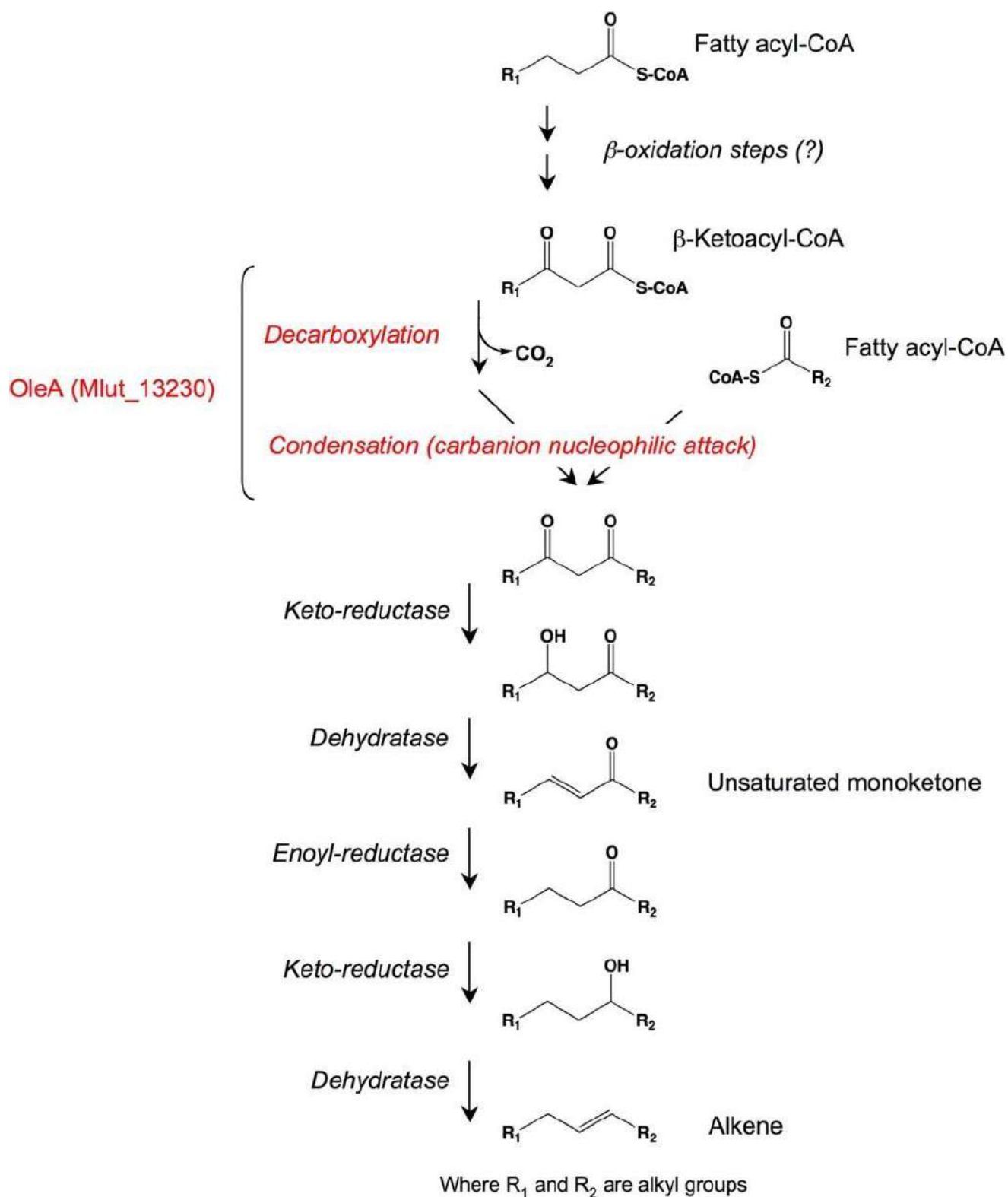
**Figure 10: Pathways of hydrocarbons synthesis originally proposed in the cyanobacterium *S. elongatus*.** A fatty acyl-ACP (C<sub>n</sub>) is first converted by an acyl-ACP reductase to the corresponding (C<sub>n-1</sub>) fatty aldehyde, which is subsequently transformed into a (C<sub>n-1</sub>) alkane or alkene by a second enzyme. A saturated fatty acyl-ACP will yield an alkane and an unsaturated one will yield an alkene. ACP: acyl carrier protein. (From Schirmer et al. [2010](#)).

production of only small amounts of hydrocarbons. The low yields observed with the AAR/ADO pathway come from different factors. It has been reported that the C-terminal domain of ADO from marine cyanobacterium *Prochlorococcus marinus* has a particular motif sequence (RMSAYGLAAA) affine for the ATP-dependent clpAP and Lon systems in *E. coli*

which disposes it to a very low accumulation and/or stability in vivo (Liu et al. [2020](#)); this sequence was also shown to be conserved in different homologous proteins. The modification of this degron motif resulted in an increase of ADO activity in vivo. In addition, biochemical studies have shown that ADO has a very low turnover ( $1 \text{ min}^{-1}$ ) (André et al. [2013](#)).

## 2- OleABCD

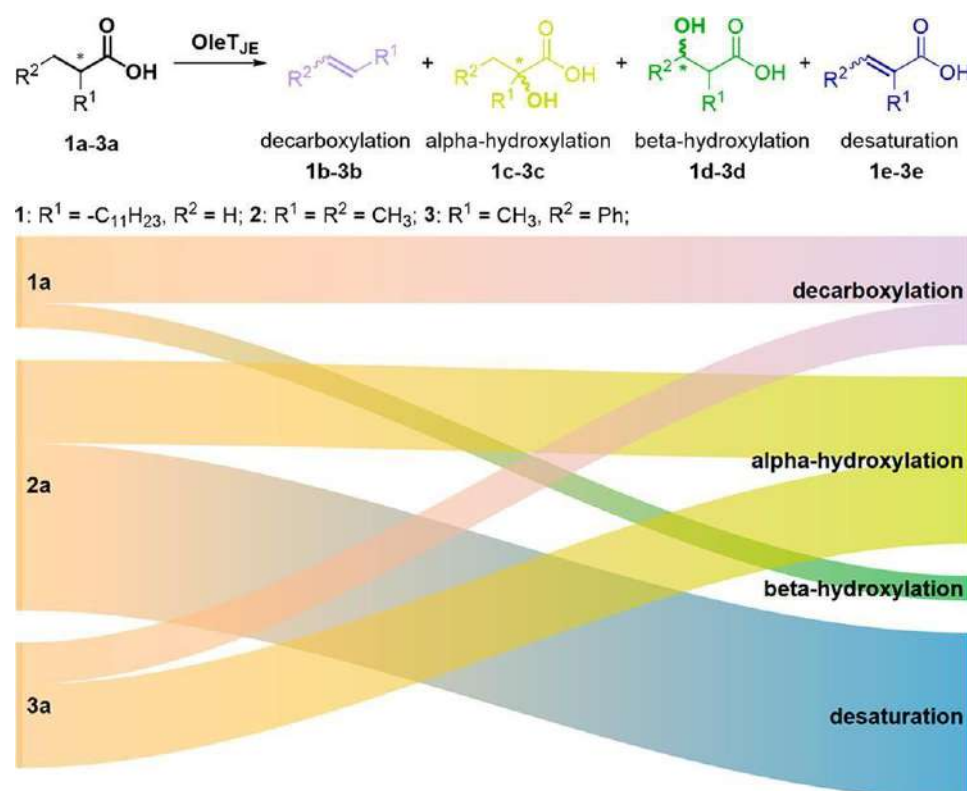
*Sarcina lutea* ATCC 533 (now *Kocuria rhizophila*) a gram-positive bacterium, was reported since the 1960s to contain very-long-chain aliphatic hydrocarbons including  $C_{22}$  to  $C_{29}$  (Tornabene [1967](#)). Since the genome of this bacterium is not sequenced, the investigation of genes implicated in its hydrocarbon biosynthetic pathway was not done. Later, the genome of *Micrococcus luteus*, a gram-positive bacteria of mammalian skin microbiota and a very close relative of *Sarcina lutea* genome, was sequenced (Young et al. [2010](#)). *Micrococcus luteus* was also reported to produce alkanes. Based on genome analysis and heterologous production in *E. coli*, a cluster of 3 genes including: *OleA*, *OleBC* and *OleD* was identified to be sufficient for alkane production. However, the individual roles of the corresponding proteins OleA, OleBC and OleD in the alkane biosynthesis remain unclear. Based on BLASTp analysis, OleA, OleBC and OleD are hypothesized to be respectively: decarboxylase (of the  $\beta$ -ketoacyl-CoA); alpha/beta hydrolase and ketoreductase. Based on these hypothetical catalytic activities, a model was proposed (**Figure 11**)(Beller et al. [2010](#)).



**Figure 11: Proposed pathway for alkene biosynthesis from condensation of fatty acids.** Compounds shown as Coenzyme A thioesters may, in fact, be acyl carrier protein (ACP) thioesters (From Beller et al. [2010](#)).

### 3-OleTje

OleTje is one of the keys enzymes implicated in hydrocarbons production in some bacterium species. First isolated from *Jeotgalicoccus spp.*, OleTje is implicated in the conversion of mainly long-chain fatty acids into terminal olefins (1-alkenes). Sequence analysis of OleTje shows that the enzyme belongs to the cyp152 family of cytochrome P450s, which includes bacterial fatty acid hydroxylases. Some cyp152 enzymes have a double activity including decarboxylation and hydroxylation of fatty acids (in  $\alpha$ - and/or  $\beta$ -position). Testing in vitro palmitic acid on the purified OleTje protein, 1-pentadecene,  $\alpha$ -hydroxy palmitic acid and  $\beta$ -hydroxy palmitic acid were produced confirming decarboxylation,  $\alpha$ - and  $\beta$ -hydroxylation activities (Rude et al. [2011](#)) (**Figure 12**). Later, a 4<sup>th</sup> catalytic activity of OleTje was demonstrated: the desaturation of fatty acids through two consecutive hydrogen atom abstraction steps (Pickl et al. [2018](#)).

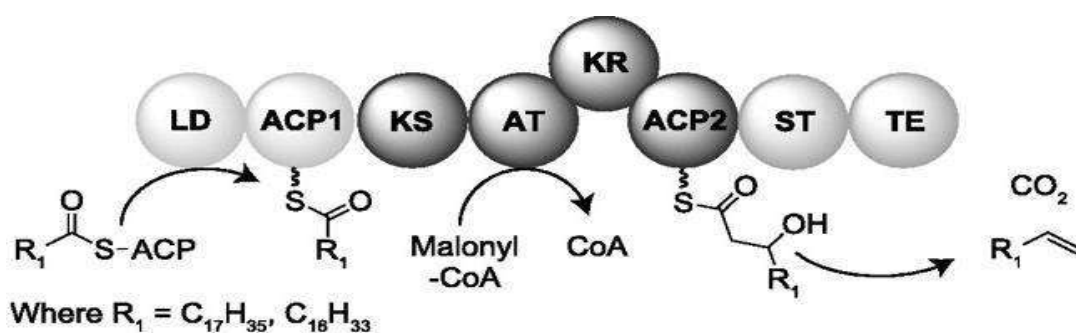


**Figure 12: Proposed mechanisms for OleTje including fatty acid decarboxylation and  $\alpha$ - or  $\beta$ -hydroxylation and desaturation routes.** OleT<sub>JE</sub>-catalyzed transformation of fatty acid substrates 1a–3a at room temperature and Sankey representation of the product distribution. The width of the lines corresponds to the relative amount of substrate that was converted (left-hand side) and the relative amount of product that was formed (right-hand side) (From Pickl et al. [2018](#)).

Since OleT<sub>JE</sub> is not very selective and can perform several reactions on fatty acids, the use of this enzyme for biotechnological application faces some limitations. Clearly, it was important to find some ways to guide the catalysis of the enzyme. Combining computational (density functional theory and QM/MM) methods, a research group was able to design in silico a double Asn242Arg/Arg245Asn mutant which performs preferentially decarboxylation (Cantú Reinhard et al. [2020](#)).

#### 4-Ols

Olefin synthase (Ols) is the second hydrocarbon pathway that was discovered in cyanobacteria (Mendez-Perez et al. [2011](#)). Contrary to the AAR/ADO enzymes, Ols is a modular multidomain protein. The domains of Ols are organized to perform a cascade of reaction: (i) elongation; (ii) sulfonation; and (iii) decarboxylation leading to the production of 1-alkenes from acyl-ACP. The proposed mechanism is shown in **Figure 13**. It was later shown that cyanobacteria have either the AAR/ADO pathway or the Ols one (Coates et al. [2014](#)).

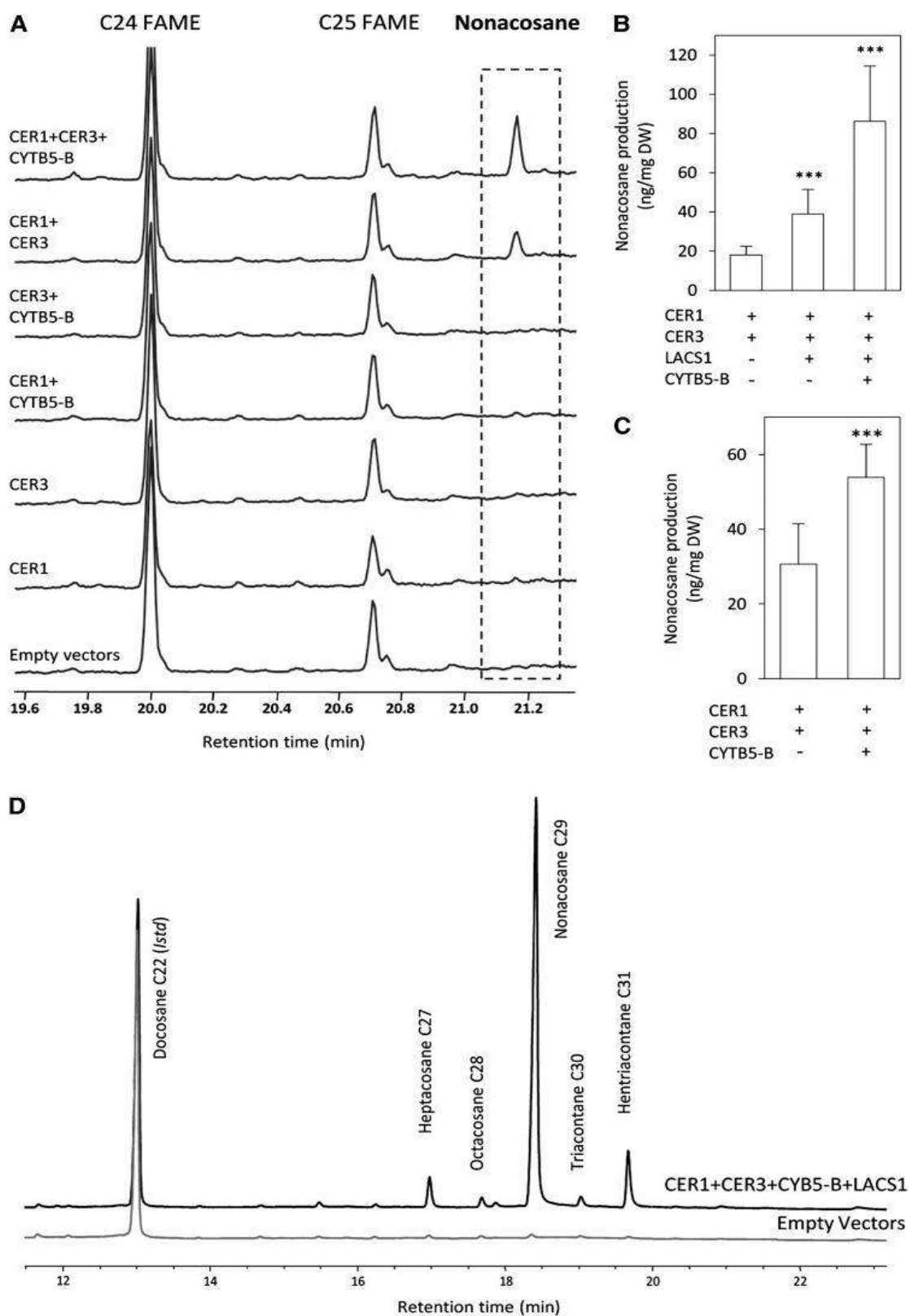


**Figure 13: Domain organization and proposed mechanism of the olefin synthase encoded by *Ols*** (From Mendez-Perez et al. [2011](#)).

#### 5-CER1/CER3

As indicated before, the synthesis in plants of VLC alkanes, is well known and the biological role of these hydrocarbons is fairly well understood. However, the biosynthetic pathways implicated in their production is still unclear. Eceriferum (*Cer*), meaning without waxes, corresponds to the name given to *Arabidopsis* mutants affected in cuticular wax deposition. Interestingly, biochemical and molecular characterization of the mutant *cer1* suggests its implication in VLC alkanes production. In fact, *cer1* mutants, in addition to the absence of VLC alkanes, can accumulate fatty aldehydes compared to the wild type of *Arabidopsis* (Jenks et al.

[1995](#)). A more recent study shows that overexpression of *cer1* in *Arabidopsis* is coupled with high content of VLC alkanes. It was thus hypothesized that CER1 is implicated in the conversion of aldehydes into hydrocarbons (Bourdenx et al. [2011](#)). However, to conclude on areal activity of CER1 protein, it would have been important to show in vitro or in vivo that CER1 is sufficient to convert aldehydes into VLC alkanes. Unfortunately, VLC aldehydes are not commercially available to test in vitro the expected aldehyde decarbonylase activity for CER1. Since the heterologous expression of CER1 failed to produce VLC alkanes in yeast, it was hypothesized that there could be another partner important for CER1 activity (Bernard et al. [2012](#)). Potential physical interactants of CER1 protein were then searched using split-ubiquitin yeast two-hybrid system. This approach showed that CER1 can interact with itself but also with its homolog CER3, a cytochrome b5 (CYTB5-B) and a long-chain acyl-CoA synthetase (LACS1). Coexpression studies in a yeast strain producing VLC fatty acids >C26 (due to overexpression of the Sur4 fatty acid elongase) showed that CER1 needed to be expressed with CER3 to yield VLC alkanes and that additional expression of CYTB5f and LACS1 each increased 2-fold the alkane production (**Figure 14**).



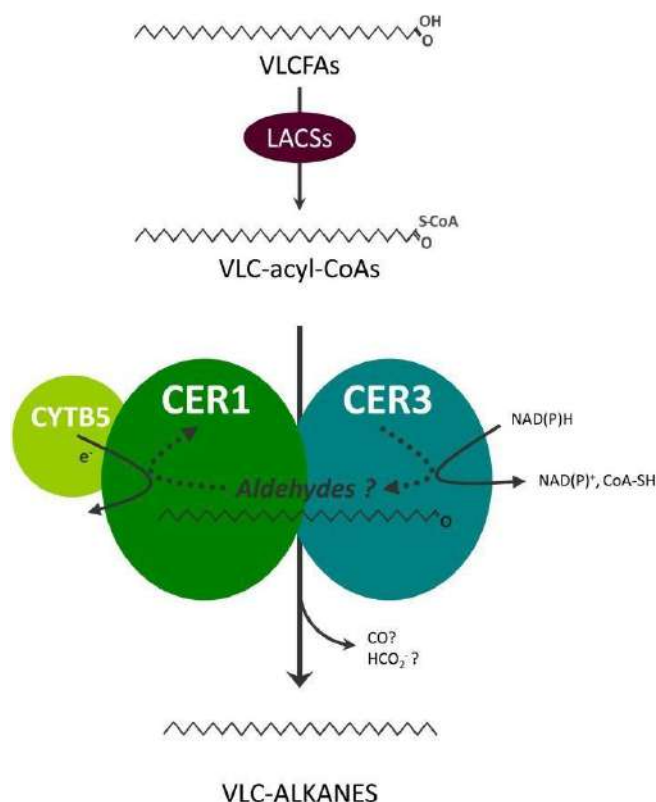
**Figure 14: Synthesis of VLC Alkanes in INVSur4# Yeast Coexpressing Various Combinations of CER1 together with CER3, CYTB5-B, and LACS1.**

(A) The INVSur4# yeast coexpressing Arabidopsis CER1 and CER3 has the ability to synthesize *n*-nonacosane. Additional coexpression of CYTB5-B enhances *n*-nonacosane production. GC-MS traces

of total fatty acyl chain analyses of INVSur4# cotransformed with denoted Arabidopsis transgenes or with corresponding empty vectors as control. **(B)** and **(C)** Quantitative analysis of *n*-nonacosane production shows that Arabidopsis LACS1 (B) and CYTB5-B ([B] and [C]) enhance alkane production when coexpressed together with CER1 and CER3 in INVSur4#. **(B)** Yeast cells were grown in stringent medium lacking His, Trp, Leu, and uracil. Mean values of nonacosane (ng/mg of dry weight [DW]) are given with sd ( $6 \leq n \leq 8$ ). Significance was assessed by Student's *t* test (\*\*\*,  $P < 0.01$ ). **(C)** Yeast cells were grown in stringent medium lacking Trp, Leu, and uracil. Mean values of *n*-nonacosane are expressed in ng per mg of dry weight with sd ( $n = 14$ ). Significance was assessed by Student's *t* test (\*\*\*,  $P < 0.01$ ). **(D)** The INVSur4# yeast coexpressing Arabidopsis CER1, CER3, CYTB5-B, and LACS1 produces VLC alkanes with chain lengths ranging from 27 to 31 carbon atoms. INVSur4# control strain (transformed with empty vectors) shows no alkane production. GC-FID traces of the hydrocarbon fractions after separation from total lipid extract by TLC. Docosane (20  $\mu$ g) was used as internal standard (From Bernard et al. [2012](#)).

Based on these results, a tentative model of the alkane-forming biosynthesis complex of plants was thus proposed (**Figure 15**). In this model, the alkane-forming complex is located on the ER membrane; CER3 is hypothesized to be a reductase converting the VLC acyl-CoAs formed by LACSs to VLC aldehydes; CER1 is the decarbonylase converting the latter to VLC alkanes; and NAD(P)H and cytochrome b5 (CYTB5) serve as electron donors for the reaction of CER1 and CER3 respectively. Further biochemical work will be needed to determine the exact reaction(s) catalyzed by CER1 and CER3.





**Figure 15: Proposed Biochemical model in which Arabidopsis CER1 and CER3 proteins Act Synergistically with cytochrome b5 for very-long-chain Alkane Synthesis.**

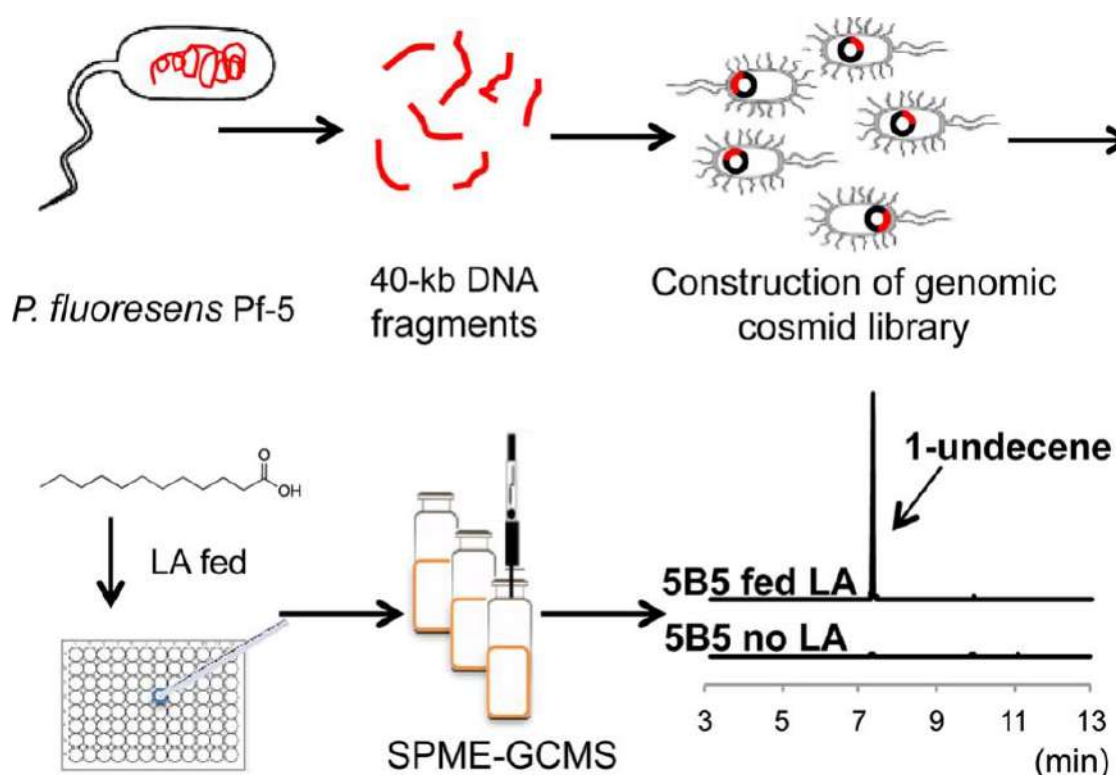
CER1 and CER3 proteins would form an enzymatic complex catalyzing the conversion of VLC acyl-CoAs to VLC alkanes. VLC fatty acids activated by the long-chain acyl CoA synthetases (LACSs) in VLC acyl-CoAs would be used as precursors of VLC alkane synthesis. A mandatory CER1/CER3 heterodimer would efficiently catalyze a two-step reaction starting with the reduction of acyl-CoA to a potential intermediate aldehyde subsequently decarbonylated to alkane with the loss of one carbon potentially in carbon monoxide or formate. CYTB5s would interact with the di-iron catalytic core of CER1, providing electron(s) required for the decarbonylation reaction (From Bernard et al. [2012](#))

## 6- UndA and UndB

UndA and UndB are nonheme mononuclear iron oxidases isolated in *Pseudomonas* species and shown to catalyze the production of 1-alkenes (Rui et al. [2014](#), Rui et al. [2015](#)). Using a sensitive SPME-GCMS analysis technique, Rui et al. (Rui et al. [2014](#)) first showed that different species of *Pseudomonas* were able to produce 1-100 ng/ml of 1-undecene. To isolate the gene responsible for such production, a heterologous screening of a genomic bank of *Pseudomonas fluorescence* was performed in *E. coli*. This screening showed that only one gene was sufficient to induce 1-undecene production; it was named *UndA* (Rui et al. [2014](#)). Interestingly, while the natural production of 1-undecene varies by two orders of magnitude in *Pseudomonas* species (1-100 ng/ml), the heterologous expression in *E. coli* of *UndA* homologous genes leads to a similar titer (3-6 mg/l). This suggests that there could be an

additional pathway for 1-undecene production in some species (mainly the species producing naturally the higher titer). The re-screen of *Pseudomonas fluorescence* Pf-5 genomic bank lead to the discovery of a second alkene-forming protein called UndB, which is a membrane-bound desaturase-like protein (Rui et al. [2015](#)) (**Figure 16**).

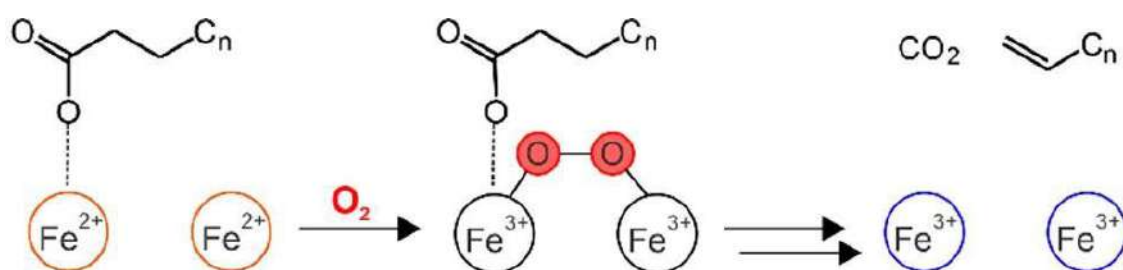
UndA and UndB are conserved only in a few bacterial species including *Acinetobacter*, *Pseudomonas Burkholderia*, and *Myxococcus*. Both UndA and UndB have a quite similar biocatalytic activity consisting of the decarboxylation of fatty acids to their corresponding 1-alkenes releasing CO<sub>2</sub> as co-product. More specifically, while UndA can catalyse the conversion of fatty acids C10-C16, UndB can convert a wider range covering C6-C16 fatty acids. It was also shown that the catalytic activity of these enzymes needed the presence of two essentials factors: Fe<sup>II</sup> as a cofactor and O<sub>2</sub> as an activator.



**Figure 16: Identification of *UndB* gene responsible of 1-undecene production by screening.**

The *Pseudomonas fluorescence* pF-5 genomic library was heterologously expressed in *E. coli* with dodecanoic acid (LA) feeding, followed by 1-undecene accumulation analysis using headspace solid phase microextraction (SPME)–gas chromatography mass spectrometry (GCMS) detection method (From Rui et al. [2015](#)).

For a better understanding of the catalytic mechanism, UndA enzyme was fed in vitro with different substrates and D-labelled dodecanoic acid derivatives including: 2-hydroxydodecanoic acid (AHDA), 3-hydroxydodecanoic acid (BHDA), [ $\alpha,\alpha$ -D2] dodecanoic acid, [D23] dodecanoic acid, and 2,3-dodecenoic acid (DEA). The analysis of the products of these substrates in addition to the need of  $\text{Fe}^{\text{II}}$  and  $\text{O}_2$  lead to the conclusion that the catalytic mechanism can be divided into 3 steps ( **Figure 17**): (I) fatty acid first binds to the mononuclear iron( $\text{Fe}^{\text{II}}$ ) center at the active site of UndA and forms a  $\text{Fe}^{\text{II}}$ -fatty acid intermediate, (II) Thanks to oxygen molecule, a Fatty acid- $\text{Fe}^{\text{(III)}}$ -superoxide complex is formed.(III) the last step consist of the abstraction of Hydrogen from  $\beta$ -carbon of the fatty acid by the oxygen molecule. Then an electron transfer occurs followed by a spontaneous scission of the carbon-carbon bond forming 1-undecene,  $\text{CO}_2$  and  $\text{H}_2\text{O}$  products (Zhang et al. [2019](#)).



**Figure 17: Schematic representation of different steps of fatty acids conversion to corresponding C-1, 1-alkene by UndA.** (From Zhang et al. [2019](#))

The most significant metabolic engineering strategy used to improve alkane production was coexpression of *UndB* with *UcFatB2*, a gene encoding a plant thioesterase. *UcFatB2* preferentially hydrolyzes dodecanoyl-ACPs thereby increasing the pool of free dodecanoic acid. Coexpression of *UndB/UcFatB2* lead to a 100-fold increase in the 1-undecene content compared to a strain overexpressing only *UndB* (Rui et al. [2015](#)).

## 7-FAP

### 7.1-Discovery

The fact that C15 to C17 alka(e)nes were found in *C.reinhardtii*, *C.variabilis* NC64A and *Nannochloropsis* sp. cultures and that no homologs of known hydrocarbon synthesis genes from plants, cyanobacteria or insects were present in the genomes of these microalgae indicated the existence of a yet unknown type of alka(e)ne-producing pathway (Sorigué et al. [2016](#)). This

prompted Sorigué and coworkers to try to isolate the protein responsible for the biosynthesis of C15 to C17 alka(e)nes in these species.

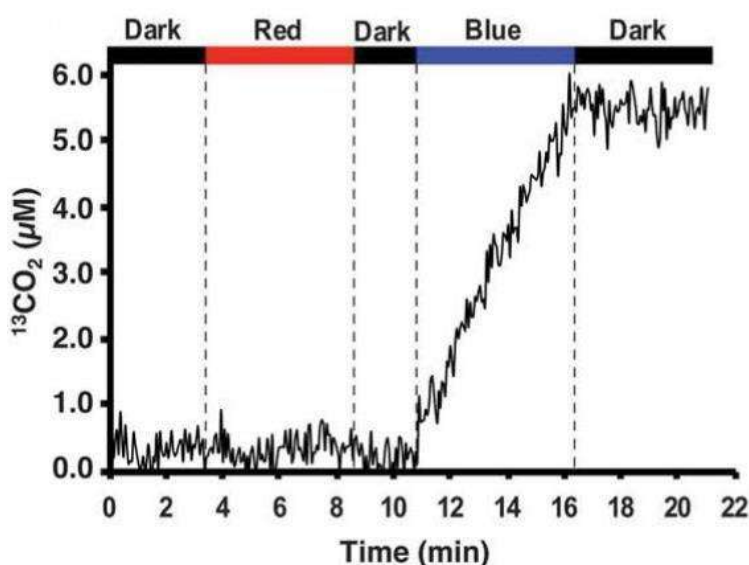
A feeding experiment had shown that following the addition of perdeuterated palmitic acid (D31) to cells, not only D29-heptadecene but also D31-16:0 (palmitic), D31-18:0 (stearic), and D29-18:1 (oleic) acids was formed (Sorigué et al. [2016](#)). This result indicated that the exogenous palmitic acid was metabolized, namely elongated, desaturated and decarboxylated (at least formally decarboxylated). Therefore, the alkene-forming enzyme was likely to be a decarboxylase converting  $C_n$  fatty acids to  $C_{n-1}$  hydrocarbons. Interestingly, it was also noted that the production of hydrocarbons in *C. variabilis* and *C. reinhardtii* was only possible in the presence of light.

To isolate the protein responsible for the hydrocarbon-forming activity, an enzymatic activity purification approach was thus developed (Sorigué et al. [2017](#)). The activity assay consisted in measuring qualitatively the formation of deuterated pentadecane (using solid phase microextraction and GC analysis) following the addition of perdeuterated palmitic to *C. variabilis* protein extracts. *C. variabilis* was chosen because of its higher content in hydrocarbons than *Chlamydomonas*. The pentadecane-forming activity was first measured in a *Chlorella* homogenate and then in protein fractions resulting from various purification steps (ion exchange columns etc.). Finally, partially purified fractions of *C. variabilis* containing the pentadecane-forming activity and coming from three independent protein purifications (which used different detergents), were sent for proteomic analysis. Only 10 proteins were in common between the three purifications. And of these 10 candidate proteins, 9 were homologs of well characterized proteins. The last candidate belonged to a family of flavoproteins, called Glucose-Methanol-Choline (GMC) oxidoreductases. To prove that this predicted GMC oxidoreductase was responsible for the production of the hydrocarbons detected in *Chlorella*, the cDNA was isolated from *Chlorella* total RNAs and expressed in *E. coli*, a bacterium devoid of alkanes and alkenes. Analysis of the total fatty acids and other hydrophobic compounds of the cultures by GC-MS analysis showed the presence of C13-17 alkane and alkenes in the strain expressing the GMC oxidoreductase. This experiment thus proved that this protein was a novel alka(e)ne-forming enzyme. This discovery was patented (Beisson et al. [2016](#)) and published (Sorigué et al. [2017](#)).

The recombinant protein was then purified and analyzed by mass spectrometry. This approach confirmed the presence of a flavin adenine dinucleotide (FAD) co-factor, which is also present in all other GMC oxidoreductases investigated before. Study of the activity of the purified recombinant GMC oxidoreductase showed that it was indeed converting  $C_n$  free fatty acids (but

not esterified fatty acids) to the corresponding C<sub>n-1</sub> alkane or alkene (without creating a terminal double bond). An activity assay conducted in the presence of 1-<sup>13</sup>C-palmitate, resulted in the production of <sup>13</sup>CO<sub>2</sub> detected by membrane inlet mass spectrometry (MIMS). This demonstrated that the enzyme was indeed a decarboxylase.

Since FAD is a well-known chromophore and the formation of hydrocarbons in the algae was known to be strictly dependent on the presence of light (Sorigué et al. [2016](#)), the authors suspected that the GMC oxidoreductase could be itself a light-dependent enzyme (a light-activated enzyme, or even a photoenzyme). To verify this hypothesis, the activity of the enzyme was monitored continuously under dark and light by quantifying the CO<sub>2</sub> co-product using MIMS. Results showed that the enzyme is inactive in the dark or under red light, decarboxylation being possible only in the presence of continuous blue light. (**Figure 18**).



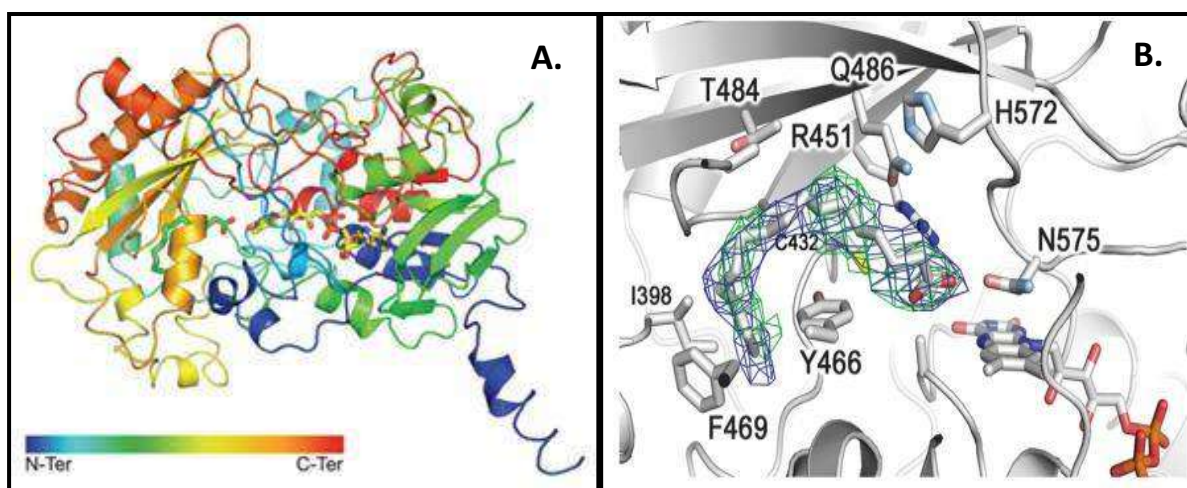
**Figure 18: Light dependency of the *C. variabilis* alkane synthase.** Activity of the purified enzyme under successive light conditions. Activity on 1-<sup>13</sup>C-palmitate was monitored via release of <sup>13</sup>CO<sub>2</sub> by using membrane inlet mass spectrometry. (Sorigué et al. [2017](#))

Since the activity of the enzyme stopped immediately after the light was turned off, it was concluded that the enzyme is not just a light-activated protein (which can continue to be active in the dark after being activated by light) but an enzyme dependent on light for each catalytic cycle (i.e., it is a photoenzyme). The *C. variabilis* GMC oxidoreductase was thus renamed Fatty Acid Photodecarboxylase (CvFAP) (Sorigué et al. [2017](#)). FAP is now classified in the Enzyme Classification as E.C. 4.1.1.106 (<https://iubmb.qmul.ac.uk/enzyme/EC4/0101b.html#106>), the group 4.1.1 being the Carboxy-Lyases.



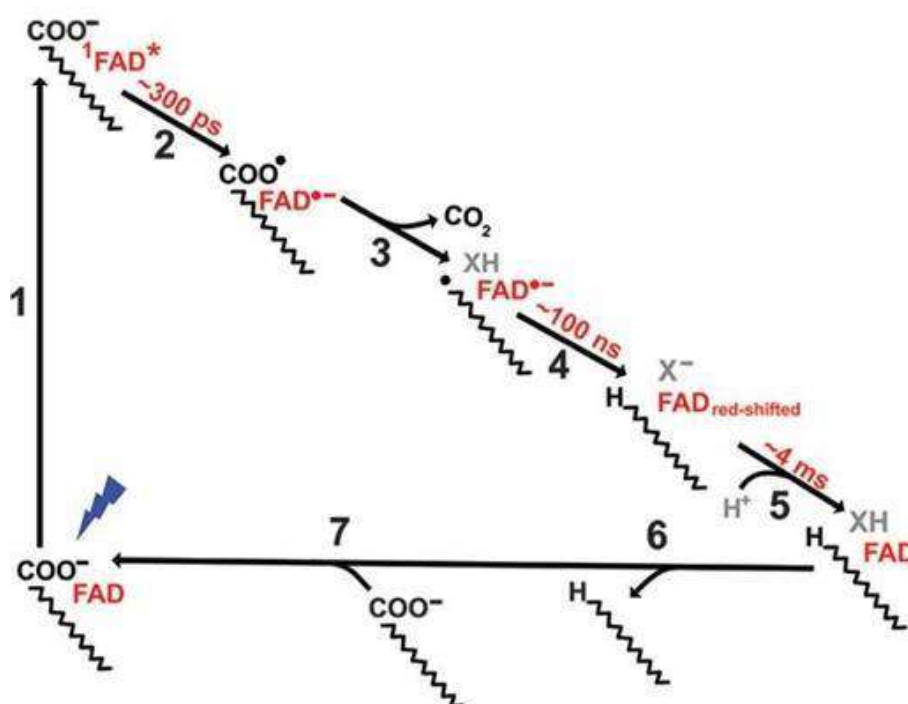
## 7.2-Structure and mechanism

Following the discovery of CvFAP in 2016, crystal structure of the recombinant CvFAP protein expressed in *E. coli* was solved at a medium resolution (3.15 Å). Note that the protein was co-purified with a fatty acid substrate which could be seen in the structure (**Figure 19A, B**). This fatty acid was present in a narrow hydrophobic tunnel framed by two domains (C and N terminal), which are present in most GMC oxidoreductases with known structure. At the end of the tunnel lied the FAD cofactor, which was close to the carboxy group of the fatty acid substrate. Some residues were thought to be important for the fatty acid substrate stabilization and/or catalytic steps (e.g. Y466, R451, C432, Q486) (**Figure 19B**).



**Figure 19: Structure of CvFAP.** **A.** Overall architecture of the enzyme in complex with FAD and a fatty acid. The structure is represented as an illustration colored from blue to red from the N to the C terminus, with the FAD and the fatty acid represented in stick. The long N-terminal helix is projected toward a non-crystallographic symmetry-related molecule that is not represented here for clarity. **B.** Details of the fatty acid-binding site with the side chains of residues within 4 Å of the substrate shown in stick. The omit map electron density associated to the fatty acid is also shown and contoured at 0.5s (2Fo-Fc; blue) and 2s (Fo-Fc; green). Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr (From Sorigué et al. [2017](#)).

Photoenzymes have the advantage to be good models for mechanistic studies because the reaction can be triggered by light and all the enzyme molecules in a sample can thus be synchronized. The capacity of the FAD cofactor to emit fluorescence, to change state and have a characteristic absorbance allowed the FAP photocycle to be elucidated based on transient absorption spectroscopy on nanosecond to millisecond time scales (**Figure 20**).



**Figure 20: First Proposed photocycle of CcFAP based on time resolved spectroscopy approach.**

Suggested model of the FAP photocycle. Directly observed intermediates and kinetics are shown in red (From Sorigué et al. [2017](#)). In the first step (1) the FAD cofactor is excited by a blue photon. (2) Fluorescence kinetics of FAD shows that in the presence of a substrate, the singlet excited state FAD ( $^1\text{FAD}^*$ ) is quenched, decreasing its lifetime from 5 ns to 300 ps. It is assumed that this 300 ps step corresponds to the abstraction by the  $^1\text{FAD}^*$  of an electron from the carboxylate of the deprotonated fatty acid. This second step thus leads to the formation of  $\text{FAD}^{\bullet-}$  and a carbonyloxy radical ( $\text{RCOO}^\bullet$ ). This radical undergoes spontaneous decarboxylation (3), yielding a radical alkyl ( $\text{R}^\bullet$ ) and  $\text{CO}_2$ . In the next step (4), the  $\text{FAD}^{\bullet-}$  is transformed to  $\text{FAD}_{\text{red-shifted}}$  in a 100 ns timescale. This step corresponds to the back electron transfer from  $\text{FAD}^{\bullet-}$  to the alkyl radical, coupled with a proton transfer from an unknown donor X (water molecule or amino acid residue). This step leads to the formation of  $\text{FAD}_{\text{red-shifted}}$  on a 100 ns timescale. (5)  $\text{FAD}_{\text{red-shifted}}$  then goes back to a normal FAD form in 4 ms, which is thought to be due to the deprotonated donor X that gets a proton back from a water molecule. In the last steps, the alkane escapes from the enzyme (6) and another deprotonated fatty acid binds to the active site (7) and a new cycle begins. (Sorigué et al., [2017](#))

## 7.3-Biochemical properties

### 7.3.1-Substrate specificity

The first characterization of the enzyme purified from *E. coli* (Sorigué et al. [2017](#)) showed that FAP can convert not only saturated fatty acids from C12 to C18 but also unsaturated fatty acids such as oleic acid (18:1 (9c)) and vaccenic acid (18:1 (11c) and 18:1 (11t)). In addition, using

equal concentrations of substrates (0.4 mM) at pH 8.5, the enzyme showed a more efficient conversion for palmitic (16:0) and margaric (17:0) acids. A conversion but at low proportion was also possible on behenic acid (22:0). Later in 2018 (Huijber et al. [2018](#)), Frank Hollman's team confirmed this substrate specificity (**Table 5**) using a cell homogenate of *E. coli* expressing CvFAP and a higher concentration of substrate (30 mM) in a buffer at pH 8.5 containing 30% DMSO. Importantly, in choosing this concentration of DMSO, they also demonstrated that the conversion of palmitic acid increased with the proportion of this solvent in solution until it reached a plateau at 30%.

Substrate	[Product] [mM]	Conversion [%]	TON (CvFAP)
C <sub>12</sub> H <sub>24</sub> O <sub>2</sub> (lauric acid)	3.0	11	500
C <sub>14</sub> H <sub>28</sub> O <sub>2</sub> (myristic acid)	6.9	25	1150
C <sub>16</sub> H <sub>32</sub> O <sub>2</sub> (palmitic acid)	27.7	96	4610
C <sub>17</sub> H <sub>34</sub> O <sub>2</sub> (margaric acid)	28.7	96	4780
C <sub>18</sub> H <sub>36</sub> O <sub>2</sub> (stearic acid)	26.1	92	4350
C <sub>18</sub> H <sub>34</sub> O <sub>2</sub> (Δ9) (oleic acid)	17.7	65	2950
C <sub>18</sub> H <sub>32</sub> O <sub>2</sub> (Δ9, 12) (linoleic acid)	14.6	49	2600
C <sub>20</sub> H <sub>40</sub> O <sub>2</sub> (arachidic acid)	25.7	90	4580

**Table 5: Substrate conversion by wt CvFAP.** Natural fatty acids including C12-C20 were tested. Conversion:  $[\text{product}]_{\text{final}} \times ([\text{product}]_{\text{final}} + [\text{substrate}]_{\text{final}})^{-1}$ . TON:  $[\text{product}]_{\text{final}} \times [\text{CvFAP}]^{-1}$  (From Huijbert et al. [2018](#)).

As previously discussed, fatty acids are very poorly soluble in aqueous solutions, so the use of an organic solvent in such a proportion allows a better solubilization of fatty acids in the bulk. In order to evaluate the capacity of shorter substrates conversion by the wild type enzyme, butyric, valeric, isobutyric, isovaleric and 2 methyl butyric acids (Amer et al. [2020](#)) were tested; unfortunately, only a very small proportion of these substrates were converted. To explain this preference of the enzyme for longer substrate chains, it was hypothesized that the positioning of the substrates in the hydrophobic pocket of the enzyme is variable depending on the chain length with optimal stability for LC fatty acids such as C16 and C17. Various 18:1 fatty acid were then tested (18:1 (9c), 18:1 (9t); 18:1 (6c) 18:1 (11t)). The correlation of the conversion



rate of these substrates with the distance between the carboxy group of these substrates and the flavin moiety of the FAD cofactor (assessed by molecular modeling) supported the initial hypothesis.

### 7.3.2-Interfacial properties

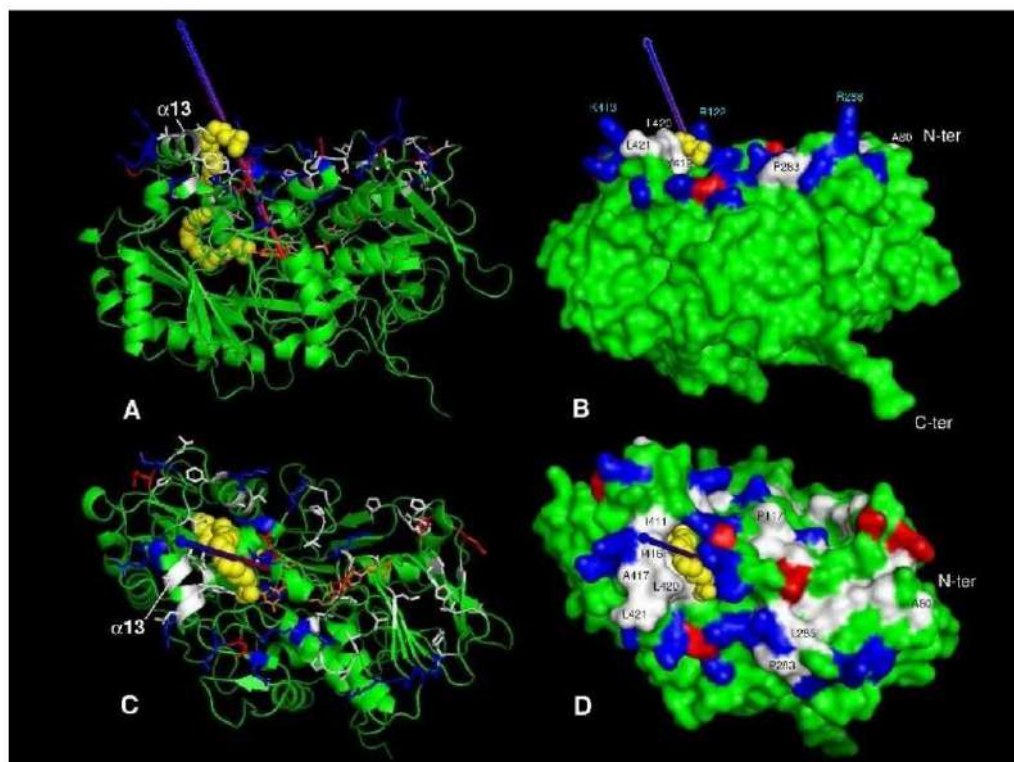
As explained in section 2.3, LC fatty acids, the natural substrates of FAP, are virtually insoluble in water and can disperse into various forms depending on concentration, pH and presence of ions. Since FAP can convert these 'insoluble' substrates in aqueous solutions, it was expected that this enzyme has certain properties allowing the binding to lipid interfaces that result from the dispersion of fatty acids. Enzymes capable of converting substrates at interfaces are called interfacial enzymes. One of the most obvious and studied models of interfacial enzymes are lipases and phospholipases (Gelb et al. 1995), which act on LC lipids, i.e., substrates that are water-insoluble by nature.

The activity of CvFAP on palmitic acid randomly dispersed in bulk, or embedded either in phospholipid liposomes, phospholipid-stabilized microemulsions or mixed micelles was compared (Aselmeyer et al. [2021](#)). Results show that FAP displays a preference for the substrate present in organized lipid assemblies, namely in liposomes and at the surface of microemulsions. The kinetics of CvFAP adsorption onto phospholipid and galactolipid monomolecular films was also studied. Data suggests that FAP has the ability to bind to and penetrate into membranes, with a higher affinity in the presence of fatty acids.

The ability to bind to interfaces such as lipid interfaces is possible thanks to a particular region of the protein called the interfacial recognition site (IRS). The IRS is characterized by the presence of amphiphilic  $\alpha$ -helices enriched with amino acids favoring interaction with lipid interfaces (e.g., mostly hydrophobic residues in lipases). The IRS can be found near the active site or in a remote location from the active, as in gastric and pancreatic lipases respectively (Aloulou et al. [2006](#)). Also, the IRS can be found at the surface of the enzyme or be covered by a so-called lid structure which opens in solution, thereby uncovering the IRS and allowing binding of the protein to the interface.

Since a high-resolution structure of CvFAP was available (Sorigué et al. [2021](#)), a search for IRS type surfaces was also performed in the study of interfacial properties of CvFAP (Aselmeyer et al. [2021](#)). There is no lid-like structure in CvFAP but a putative IRS could be found. It consisted of clusters of hydrophobic and basic residues surrounding the active site entrance (**Figure 21**): hydrophobic residues of the N-terminal end, including P78, V79 and A80 are exposed to the same side with mostly basic amino acids including R109, K125, R278

forming a plateau. This putative IRS with resulting dipolar moment suggests the orientation of FAP at negatively charged interfaces.



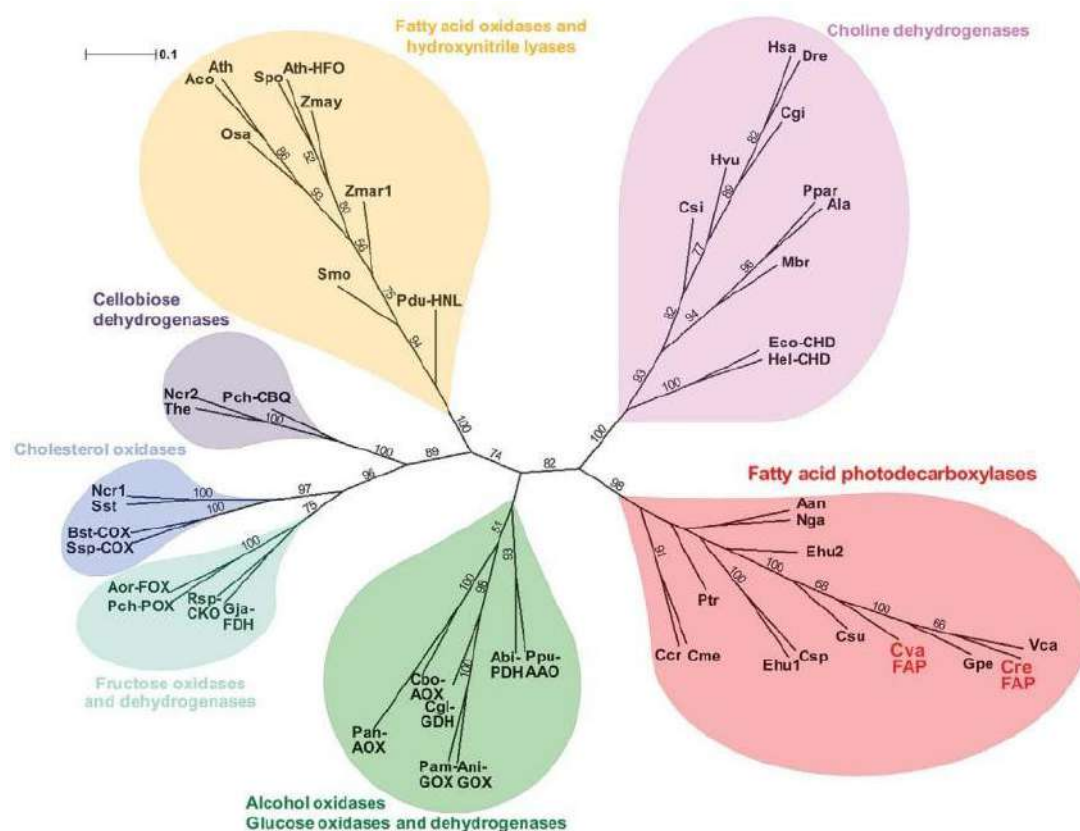
**Figure 21: Interfacial Recognition Site and dipolar moment in CvFAP.**

Side (A, B) and top (C, D) views of CvFAP are shown either as ribbon model with secondary structure elements (A, C) or molecular surface representation (B,D). Top views are aligned on a plateau found on one side of CvFAP, from which one co-crystallized C18 fatty acid molecule (shown as yellow spheres) emerges from the enzyme active site entrance and interacts with several residues from the amphipathic helix  $\alpha 13$ . A second fatty acid molecule is found deeper within the active site. All atoms of CvFAP are colored green, except those belonging to hydrophobic (white), basic (blue) and acidic (red) amino acids found in the plateau. The arrow originating from the center of gravity of the protein shows the orientation of the dipolar moment (1527 Debye) estimated for CvFAP using a Pymol script. The blue extremity indicates the positive pole. The structure of CvFAP used for this study is deposited in the Protein Databank under the accession number 6YRU (From Aselmeyer et al. [2021](#)).

## 7.4-Biodiversity

GMC oxidoreductases (Cavener [1992](#)) are found in various species such as plants (Dreveny et al. [2001](#)), insects and fungi (Sützl et al. [2019](#)). Within the GMC oxidoreductases, the main catalytic functions are those of oxygenases and dehydrogenases; there are however GMCs of plants having rather lyase activities particularly in the rupture of C-C bonds. Following the discovery that the alkane-forming enzyme of the green alga *Chlorella variabilis* belonged to the family of GMC oxidoreductases, a phylogenetic analysis of this family which included several algal homologs of the *Chlorella* FAP was performed (Sorigué et al. [2017](#)). The phylogenetic

relationships of about 50 selected GMC oxidoreductases from both eukaryotes and prokaryotes were analyzed. Surprisingly, the two algae GMC oxidoreductases initially found to bear FAP activity, those of *C. variabilis* and *C. reinhardtii*, were not found to group with the plant subgroup. Indeed, *C. variabilis* and *C. reinhardtii* were grouped in a subgroup containing algal sequences only and all GMC oxidoreductases originating from the various groups of algae clustered together (**Figure 22**).



**Figure 22: Unrooted phylogenetic tree of the GMC oxidoreductase family (including FAPs).**

The neighbor joining method was used. Species are abbreviated to three letters (one for the genus and two for the species epithet). Proteins with demonstrated biochemical activity are indicated by three additional capital letters: AAO, aryl alcohol oxidase; AOX, alcohol oxidase; FAP, fatty acid photodecarboxylase; CBQ, cellobiose dehydrogenase; CHD, choline dehydrogenase; COX, cholesterol oxidase; CKO, compound K oxidase; FDH, fructose dehydrogenase; FOX, fructose oxidase; GDH, glucose dehydrogenase; GOX, glucose oxidase; HFO, hydroxy fatty acid oxidase (genetic evidence only); HNL, hydroxynitrile lyase; and POX, pyranose oxidase (From Sorigué et al. [2017](#))

By contrast, some subgroups of the tree like choline dehydrogenases contained sequences of various origin, from *E. coli* to Mammals. The presence of an algal-specific subgroup was a striking result because algae are a very diverse group of organisms with various phylogenetic origins. This result suggested all algal GMC oxidoreductases may have conserved FAP activity

and also that these proteins had a biological function specific to algae and important for all algal species.

### **7.5-Possible biological functions**

The biological role(s) of the *n*-alka(e)ne products formed by FAP inside algal cells is still unknown but has started to be investigated in the model green microalgae *C.reinhardtii* (Moulin et al. [2021](#)). This study has shown that >90% of the only fatty acid-derived hydrocarbon produced in this species (7-heptadecene) is recovered in the thylakoid fraction of the cell. Consistent with product localization, the FAP amino acid sequence contains a putative chloroplast transit peptide and western blots performed on chloroplastic fractions show that FAP is present in the soluble fraction but also associated with the thylakoid membranes of the chloroplast. In a *C. reinhardtii* knockout mutant for FAP, 7-heptadecene was not detectable in the cells and photosynthesis activity was found to be reduced after cold acclimation when light intensity varied. Finally, phylogenetic analysis in Eucaryotes indicates that the FAP gene is always absent when photosynthesis is lost and mostly conserved during evolution of algal lineages. Taken together, these observations suggest that in algal cells FAP is related to photosynthetic functions.

### **7.6-Use of FAP for biotechnological purposes**

FAP is the latest hydrocarbon-producing enzyme to be discovered. Since hydrocarbons have applications in different fields, in particular in fuels, chemistry and cosmetics, FAP, like other hydrocarbon-producing enzymes, has aroused great interest in the community of biotechnologists. FAP is attractive for several reasons: (i) it is the only hydrocarbon-forming enzyme that is light-dependent, (ii) FAP catalysis is relatively simple since it does not require any additional molecules such as electron donors and other cofactors, (iii) FAP catalysis has a high conversion efficiency (up to 90%), (iv) FAP does not produce undesirable and/or toxic by-products or intermediates. These advantages prompted the use of CvFAP as a biocatalyst (the wild type protein or some mutants) either for in vitro reactions or in vivo (bioconversions) (Guo et al. [2023](#)).

#### **7.6.1-Use of wild type CvFAP**

Using lysates of *E. coli* expressing FAP and fatty acids dissolved in a buffer containing 30% of DMSO, Hollman and co-workers were able to achieve efficient conversion of 30 mM fatty acid on a 1-mL scale with 6  $\mu$ M FAP (Huijbers et al. [2018](#)): up to 96% conversion for oleic acid and 4610 turnover numbers (TON, i.e. final concentration of product/FAP concentration). In

the same study, they demonstrate that FAP can be used in a bioenzymatic cascade (using a lipase from *Candida rugosa*) to achieve the conversion of oil (triolein) into alkanes with an interesting conversion rate that can reach 83% (8280 TON).

In order to increase the activity of CvFAP on short-chain fatty acids, use of a decoy molecule was reported (Zhang et al. [2019](#)). The idea is that since FAP is reported to decarboxylate most efficiently LC fatty acids (C16-C18), addition of an alkane decoy to the SC fatty acid substrate would help fill the vacant space in the enzyme hydrophobic pocket, thereby stabilizing the substrate and increasing the activity. The most efficient combinations substrate/decoy were for which total carbon number (fatty acid substrate + alkane decoy) was  $16 \pm 1$ . For example, activity was increased 2.9-fold on butyric acid (C4:0) using C12 alkane as decoy. Using this approach, the reaction rate on a variety of SC fatty acids could be increased from 1.2 to 7.3 times using 150 mM substrate, 7.5 mM decoy and a *E. coli* homogenate containing CvFAP (6  $\mu$ M final). In the cyanobacterium *Synechocystis spp 6803*, it has been shown that overexpression of CvFAP combined with an increase in C16-C18 free fatty acids with the bacterial thioesterase TesA and the deletion of the *aas* gene encoding the native cyanobacterial fatty acyl-ACP synthase leads to a 19-fold increase in total alkane content (relative to  $\Delta aas$ ) (Yunus et al. [2018](#)).

Oleaginous yeasts are well known to efficiently accumulate lipids. An engineered *Yarrowia lipolytica* strain was used to perform hydrocarbons production. This strain modified (by harboring mainly *TGL4/KITGL3* intracellular lipases that can hydrolyze TAG) can accumulate a high amount of free fatty acids. Using FAP, free fatty acids were converted efficiently leading to a production of total hydrocarbons content reaching 58.7 mg/L (Bruder et al. [2019](#)).

To facilitate the recovery of the hydrocarbons produced by microorganisms expressing FAP, a possible strategy is to shorten their chain length to increase their volatility and thus their partitioning into the gas phase of the microbial cultures. There is therefore no need to collect and lyse cells to recover hydrocarbons. Using this strategy, Moulin and coworkers (Moulin et al. [2019](#)) have provided a proof of concept that FAP-expressing bacterial cultures can be turned into a continuous photoproduction system of volatile hydrocarbons. In this system FAP was coexpressed with a thioesterase of the plant *Umbellularia californica* known to act mostly on C10-C14 acyl-ACPs. Coexpression of FAP and Tes leads to the accumulation of up to 94mg/L of total hydrocarbons (corresponding to a 4-fold increase compared to a strain expressing FAP only). Interestingly, volatile hydrocarbons (C11 and C13) represented 30% of total hydrocarbons.

In addition, it should be noted that unsubstituted fatty acids are not the only substrates of FAP, CvFAP was also reported to catalyze the conversion of ricinoleic acid (18:1(9c, 12OH)) into

secondary fatty alcohol, amines, and fatty acids esters (Cha et al. [2020](#)). This work also represents a proof of concept that CvFAP can be used in multi enzymatic cascade reactions to produce other molecules than hydrocarbons. Other studies show that CvFAP can decarboxylate dicarboxylic acid (Zeng et al. [2021](#)) and epoxy fatty acids (Ge et al. [2022](#)).

A summary of the main studies on the use of CvFAP, the achievements and the principal test conditions is in **Table 6**.



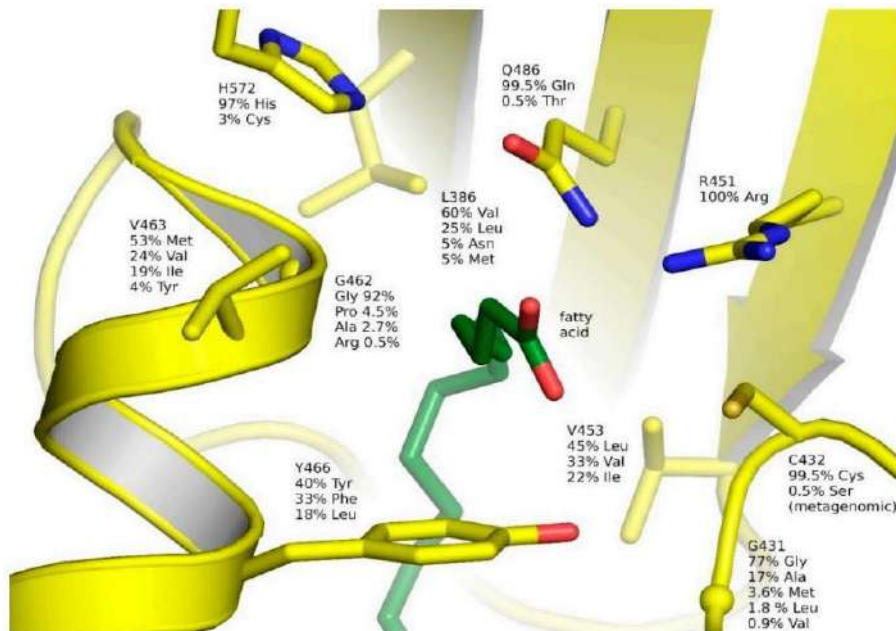
Host for recombinant expression	Biotechnological strategy	Light quality Light intensity Reaction time	Enzyme form	Major product(s) (Titer or productivity)	Ref.
<i>E. coli</i>	Use of a lipase to allow oil conversion into free fatty acids	450nm $13.7 \mu\text{mol L}^{-1} \text{ s}^{-1}$ 32h	Cell homogenate	8-Heptadecene (49.7 mM)	Huijbert et al. <a href="#">2018</a>
<i>Synechocystis</i>	Coexpression of CvFAP with a thioesterase	(-) $250 \mu\text{mol s}^{-1} \text{ m}^{-2}$ 7 days	In vivo	Pentadecane (77 mg g <sup>-1</sup> Cell Dry Weight)	Yunus et al. <a href="#">2018</a>
<i>E. coli</i>	Decoy molecule	450nm $13.8 \mu\text{mol L}^{-1} \text{ s}^{-1}$ 3h	Cell homogenate	Butane (2440 $\mu\text{M}$ )	Zhang et al. <a href="#">2019</a>
<i>Yarrowia lipolytica</i>	Expression of CvFAP	465-640 nm $200 \mu\text{mol m}^{-2} \text{ s}^{-1}$ 96h	In vivo	Heptadecane (58.7 mg L <sup>-1</sup> )	Bruder et al. <a href="#">2019</a>
<i>E. coli</i>	Expression of CvFAP with a C12-specific thioesterase	White LEDs $150 \mu\text{mol. m}^{-2} \text{ s}^{-1}$ 5 days	In vivo	Undecane (96 mg L <sup>-1</sup> )	Moulin et al. <a href="#">2019</a>
<i>E. coli</i>	Coexpression of CvFAP with an alcohol dehydrogenase and a transaminase	450nm (-) 15min/3h/12h	In vivo	(Z)-heptadec-9-en-7-ol (9.1 mM) (S,Z)- heptadec-9-en-7-amine (3.9 mM) 10- (heptanoyloxy)dec-8-ene (9.7 mM)	Cha et al. <a href="#">2020</a>
<i>Yarrowia lipolytica</i>	Optimization of glucose fermentation to maximize acyl-CoA to alka(e)ne flux	Blue LEDS (-) 17 days	In vivo	Heptadecane (1.47 g L <sup>-1</sup> )	Li et al. <a href="#">2020</a>
<i>E. coli</i>	Conversion of dicarboxylic acids	Blue light (-) 6h	Cell homogenate	Tetradecane (10.28 mM)	Zeng et al. <a href="#">2021</a>
<i>E. coli</i>	Immobilisation of CvFAP	white LEDs $62 \mu\text{mol.m}^{-2} \text{ s}^{-1}$ 20h	Immobilised enzyme in Deep Eutectic Solvent	Heptadecane (81% of product analytical yield)	Chanquia et al. <a href="#">2022</a>
<i>E. coli</i>	Immobilization of CvFAP using Eupergit C250 L	455 nm $42 \mu\text{mol.L}^{-1} \text{ s}^{-1}$ (-)	Purified protein	Pentadecane (5.7 g L <sup>-1</sup> h <sup>-1</sup> )	Simic et al. <a href="#">2022</a>
<i>E. coli</i>	Expression of CvFAP	455–460 nm (-) 10h	Cell homogenate	Heptadecane (9.5 mM)	Liu et al. <a href="#">2023</a>

**Table 6: Summary of different studies made on CvFAP wild type protein.** For some of the studies, (-) is indicated when information is not available. Studies are arranged in chronological order. Ref. indicate the corresponding article reference.

### 7.6.2-Use of CvFAP mutants

Several studies show that wild type FAP converts preferentially C16-C18 fatty acids (Sorigué et al. [2017](#); Zhang et al. [2019](#)). Successful approaches of decoy molecules showing that the length of the catalytic hydrophobic channel of the enzyme matched with a LC fatty acid also strongly supported the consensus view that CvFAP needed to be engineered to enhance his efficiency toward SC fatty acids. Based on conventional protein engineering, directed evolution and rational design approaches, a variety of CvFAP mutants was made.

The most successful CvFAP mutant was at Gly462 position. This amino acid is strongly conserved (92%) in the putative FAPs found in nature (Aleksenko et al. [2020](#)). It can be replaced in some FAPs by Pro (2.7%), Ala (4.5%) or Arg (0.5%) (**Figure 23**). Interestingly, the mutation of Gly462 into Isoleucine or Valine also leads to the enhancement of the enzyme activity on shorter chain substrates, especially butyric and pentanoic acids in *E. coli* and *Halomonas* engineered strains: *Halomonas* strain feeded with butyric acid were able to produce about 10 and 17-fold more propane with Gly462Val and Gly462Iso mutants respectively compared to the wild type CvFAP (Amer et al. [2020](#), Trisrivirat et al. [2020](#)).



**Figure 23: Variation in FAP amino acids surrounding the carboxylate moiety of the fatty acid.**

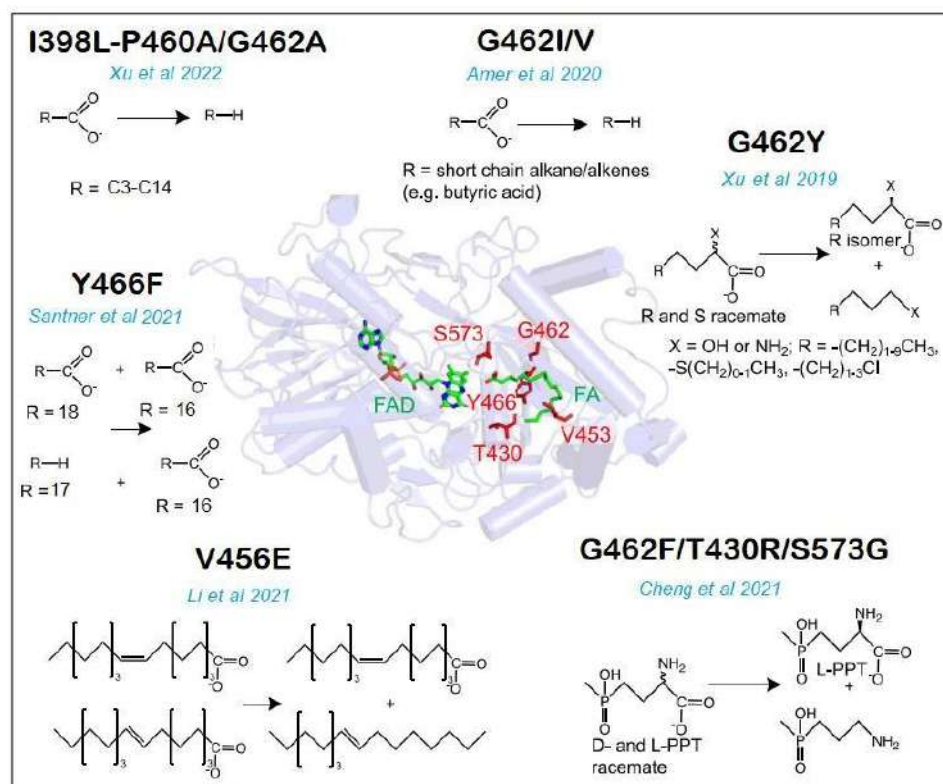
For each amino acid, frequencies of observing different amino acids at this position in FAPs are shown. Less-frequent alternatives are in some cases omitted for clarity. The structure and residue numbers correspond to CvFAP. C $\alpha$  atoms of Gly431 and Gly462 are shown as spheres (From Aleksenko et al. [2020](#)).



Xu et al. reported that I398L and P460A/G462A mutant have an increased activity on several fatty acids including C3-C14. A 29 to 552-fold increase in decarboxylation catalytic efficiency was achieved by a Focused Rational Iterative Site-specific Mutagenesis (FRISM) approach (Xu et al. [2022](#)).

Several studies reported engineered CvFAP to enhance selectivity. The G462F/T430R/S573G triple mutant converts preferentially D, L-phosphinothricin (PPT) to optically pure L-PPT since the stereoselectivity is increased (96% of L-PPT) (Cheng et al. 2020); V453E mutant convert most efficiently *trans* fatty acids (1000-fold more selective than wild type) (Li et al. [2021](#)); Rationally designed Y466F mutant decarboxylates preferentially C18 fatty acids in a mixture containing C18 and C16 fatty acids; CvFAP has a C7 to C15 alkane chemical yield of 50% but for Y466Fmutant it is 92%(Santner et al. [2021](#)).

All the reported mutants are summarized in **Figure 24** and the analytical reaction conditions are shown in **Table 7**.



**Figure 24: Designed mutants of CvFAP.** The references of the original articles are in blue (Adapted from Hedison et al. [2022](#)).

Strain	Mutation and preferred substrate	Light quality Light intensity Reaction time	Enzyme form	Major product(s), Increased fold selectivity compared to CvFAP	Reference
<i>E. coli</i>	G462Y $\alpha$ -Hydroxy acid( S isomer)	450nm - 12h	Cell homogenate	Hydroxy alkanes 4.5 Fold	Xu et al. <a href="#">2019</a>
<i>E. coli</i>	G462F/T430R/S573G D-PPT	450 nm - 12h	Cell homogenate	L-PPT E value increased from 1.2 to 193.6	Cheng et al. <a href="#">2020</a>
<i>E. coli</i>	V453E Elaidic acid (C18:1, $\Delta^9$ trans)	Blue light - 30 min	Cell homogenate	8-Heptadecene 1000-fold	Li et al. <a href="#">2021</a>
<i>E. coli</i>	Y466F C18 FA	450nm $62 \mu\text{mol.m}^{-2}.\text{s}^{-1}$ 20h	In vivo	Heptadecane 2 fold	Santner et al. <a href="#">2021</a>
<i>E. coli</i>	G624I/V Butyric acid	455-470 nm - 16-18h	In vivo	Propane 10-17-Fold	Amer et al. <a href="#">2020</a>
<i>E. coli</i>	I398L <i>n</i> -Octanoic acid P460A/G462A		Purified protein	Heptane 29-552-Fold	Xu et al. <a href="#">2022</a>

**Table 7: Resume of mutants designed on CvFAP and principal characteristics.**

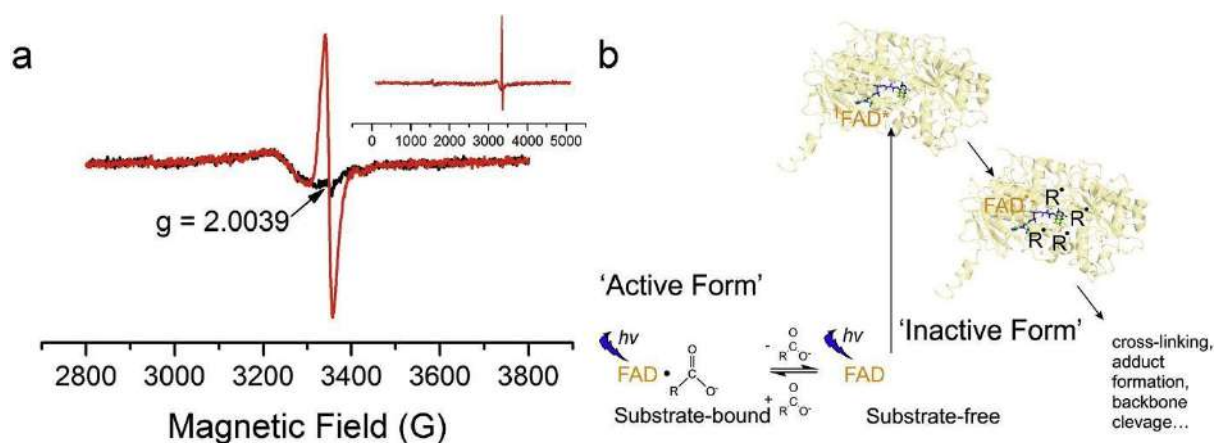
## 7.7-Fatty acid photodecarboxylase limitations

### 7.7.1-Inhibition/ Inactivation by light and thermostability

Sorigué et al. showed that CvFAP activity increases linearly with white light intensity up to  $2000 \mu\text{mol.m}^{-2}.\text{s}^{-1}$ . The first report of a possible negative effect of light intensity on FAP protein activity was reported by Moulin and coworkers (Moulin et al. [2019](#)) when studying the effect of light intensity on the hydrocarbon production in *E. coli* cells. Increasing light from 150 to  $1500 \mu\text{mol.m}^{-2}.\text{s}^{-1}$  results in a drastic decrease of hydrocarbon production and is coupled with a drop in amount of soluble FAP protein highlighted by immunodetection. It was hypothesized that FAP can be probably photodamaged in the presence of an excess of light. Later, it has been shown that purification under dim red light is more efficient to isolate active CvFAP from *E. coli* (Lakavath et al [2020](#)) since production in white light leads to a catalytic unstable protein. In the same study, using EPR spectroscopy (**Figure 25a**), they show that a large radical signal is observed on an illuminated protein without substrate (no radical was observed in presence of substrate). This suggested a light-dependent irreversible photoinactivation of FAP especially under substrate deprived condition. Based on this observation, a photoinactivation mechanism was proposed (**Figure 25b**).

It has also been demonstrated that photostability of FAP is improved in the presence of *n*-octanoic acid (Wu et al. [2021](#)). These observations definitively show that a more photostable mutant must be found and/or light intensity control must be achieved especially in continuous production system to allow the enzyme to make the maximum total turnover under reasonable reaction time. Since the catalytic activity is dependent on the presence of FAD in the protein, active mutants should probably have the same limitation regarding photoinactivation. Light intensity and reaction time should then be the most important parameters to be studied.

In addition to light intensity, another parameter that can influence FAP activity is the temperature. It has been shown that CvFAP protein activity drastically decreases beyond  $35^{\circ}\text{C}$  (Sorigué et al.[2017](#), Sun et al. [2021](#)). Interestingly, using ancestral sequence reconstruction technique, a putative ancestor protein with increased thermostability ( $15^{\circ}\text{C}$  more) was obtained (Sun et al. [2021](#)). The ancestor protein is 20-fold more active when incubated at  $50^{\circ}\text{C}$ . This study suggests that thermostability improvement can drastically increase the total turnover numbers (TON) under continuous conditions. This parameter should be integrated with light intensity and irradiation duration in continuous alkane production.



**Figure 25: Radical based deactivation of CvFAP.** (a) Continuous-wave electron paramagnetic resonance spectroscopy (CW-EPR) showing the formation of radical species when samples were illuminated with a 455 nm LED. The black and red traces correspond to the non-illuminated and 30 min illuminated CvFAP proteins, respectively. The inset shows an expanded view of the CW-EPR spectra. (b) Proposed photoinactivation mechanism of CvFAP. The FAD cofactor is photoexcited to form the photoexcited singlet ( $^1\text{FAD}^*$ ) state. This highly oxidising form of the FAD cofactor abstracts an electron from a proximal amino acid (R), forming a radical amino acid ( $\text{R}^\bullet$ ) and the FAD semiquinone (From Lakavath et al [2020](#)).

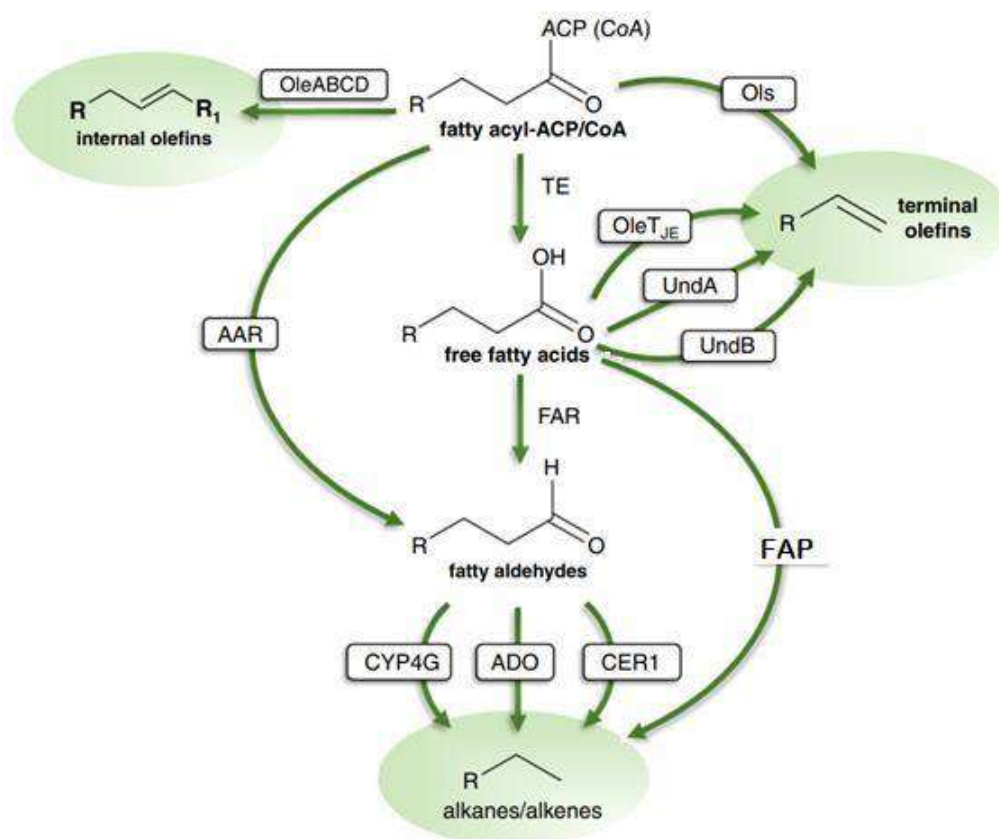
### 7.7.2-Physical state of the substrate

So far, all the strategies to improve FAP activity have been focused on enzyme engineering to improve turnovers (Huijbert et al. [2018](#), Amer et al. [2020](#), Zhang et al. [2019](#)), selectivity (Xu et al. [2018](#), Li et al. [2021](#), Santner et al. [2021](#)), thermostability (Sun et al. 2021), or on metabolic engineering strategies to enhance precursors of the substrates (Moulin et al. [2019](#), Yunus et al. [2018](#), Bruder et al. [2019](#), Li et al. [2020](#)). However, the way the substrates are structured at the interface may also be important for FAP activity (Aselmeyer et al. [2021](#)) and should also be considered as a potential limiting factor. Therefore, before engineering a FAP or when assessing the performance of an engineered FAP, parameters that influence fatty acid assembly in solution should be thoroughly investigated (pH, substrate concentration, buffer composition, etc.).

## 8-Summary of the different engineering techniques used to boost alkane production

Alkanes forming enzymes represent a promising alternative to fossil fuels. However, the catalytic pathways and enzymes are very diverse (**Figure 26**) and, in view of industrial applications, each has some limitations including: low conversion efficiency, the need for expensive cofactors, the presence of toxic and/or undesirable side products intermediates etc...

To overcome these limitations, engineering techniques are continuously being developed. These techniques can be classified into two main groups: attempts to increase the biodisavailability of hydrocarbon precursors and approaches to improve the enzymes at the catalytic level. The main techniques developed are summarized in **Figure 27**.



**Figure 26: Fatty acid-derived hydrocarbon biosynthesis pathways.** Abbreviations: AAR : acyl-ACP reductase; ACP: acyl-acyl carrier protein; ACR : acyl-CoA reductase; FAP: Fatty acid photodecarboxylase, FAR: fatty acid reductase; OleA,B,C and D : 4 sequential enzymes from the bacterium *Xanthomonas campestris* catalyzing the head-to-head non-decarboxylative thiolytic condensation of two acyl-CoAs to a LC internal alkene; OleT : cytochrome P450 from the bacterium *Jeotgalicoccus sp.*; UndA and B: nonheme iron(II)-dependent oxidase from the bacterium *Pseudomonas sp.*; ADO, aldehyde deformylating oxygenase from cyanobacteria; CYP4G, Cytochrome P450 4G from insects; CER1, ECERIFERUM 1 from plants; Ols, type-I polyketide synthase from cyanobacteria. (Adapted from Herman and Zhang [2016](#)).

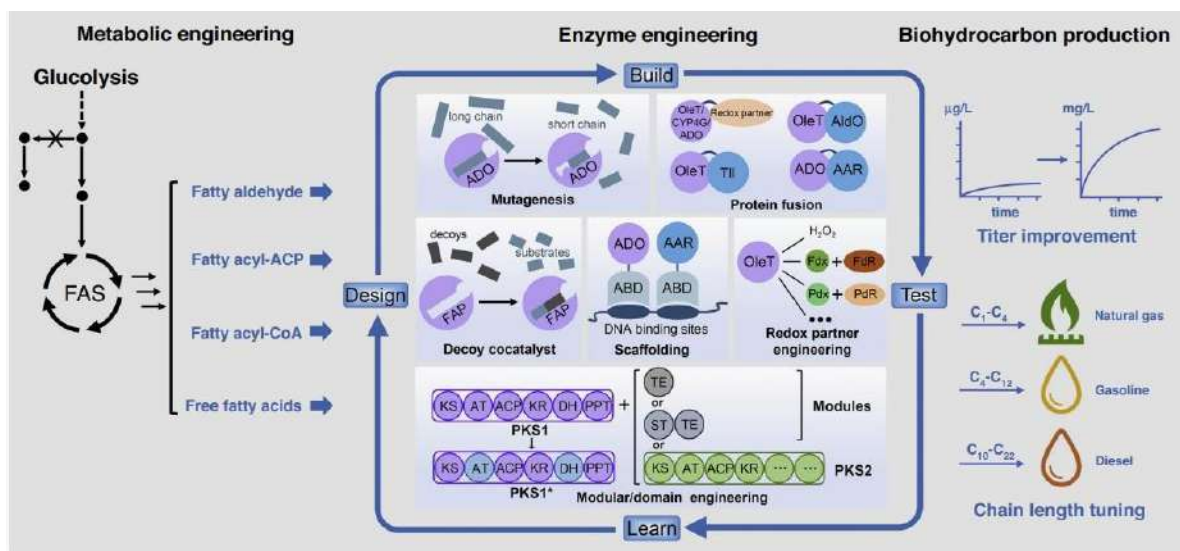


Figure 27: Schematic representation of the various metabolic and enzyme engineering techniques used to boost hydrocarbon production. (From Liu and Li, [2020](#))

## References

- Affens, W. A., & McLaren, G. W. (1972). Flammability properties of hydrocarbon solutions in air. *Journal of Chemical and Engineering Data*, 17(4), 482-488.
- Albahri, T. A. (2003). Flammability characteristics of pure hydrocarbons. *Chemical Engineering Science*, 58(16), 3629-3641.
- Aleksenko, V. A., Anand, D., Remeeva, A., Nazarenko, V. V., Gordeliy, V., Jaeger, K. E., ... & Gushchin, I. (2020). Phylogeny and structure of fatty acid photodecarboxylases and glucose-methanol-choline oxidoreductases. *Catalysts*, 10(9), 1072.
- Aloulou, A., Rodriguez, J. A., Fernandez, S., van Oosterhout, D., Puccinelli, D., & Carrière, F. (2006). Exploring the specific features of interfacial enzymology based on lipase studies. *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids*, 1761(9), 995-1013.
- Amer, M., Wojcik, E. Z., Sun, C., Hoeven, R., Hughes, J. M., Faulkner, M., ... & Scrutton, N. S. (2020). Low carbon strategies for sustainable bio-alkane gas production and renewable energy. *Energy & Environmental Science*, 13(6), 1818-1831.
- Andre, C., Kim, S. W., Yu, X. H., & Shanklin, J. (2013). Fusing catalase to an alkane-producing enzyme maintains enzymatic activity by converting the inhibitory byproduct H<sub>2</sub>O<sub>2</sub> to the cosubstrate O<sub>2</sub>. *Proceedings of the National Academy of Sciences*, 110(8), 3191-3196.
- Aselmeyer, C., Légeret, B., Bénarouche, A., Sorigué, D., Parsiegla, G., Beisson, F., & Carrière, F. (2021). Fatty Acid Photodecarboxylase Is an Interfacial Enzyme That Binds to Lipid–Water Interfaces to Access Its Insoluble Substrate. *Biochemistry*, 60(42), 3200-3212.
- Ayala-Bribiesca, E., Turgeon, S. L., & Britten, M. (2017). Effect of calcium on fatty acid bioaccessibility during in vitro digestion of Cheddar-type cheeses prepared with different milk fat fractions. *Journal of dairy science*, 100(4), 2454-2470.
- Beisson, F., Sorigué, D. & Légeret, B. (2016) New fatty acid decarboxylase and its uses. *PATENT APPLICATION EP3246401A1*.
- Beller, H. R., Goh, E. B., & Keasling, J. D. (2010). Genes involved in long-chain alkene biosynthesis in *Micrococcus luteus*. *Applied and environmental microbiology*, 76(4), 1212-1223.
- Bernard, A., Domergue, F., Pascal, S., Jetter, R., Renne, C., Faure, J. D., ... & Joubès, J. (2012). Reconstitution of plant alkane biosynthesis in yeast demonstrates that Arabidopsis ECERIFERUM1 and ECERIFERUM3 are core components of a very-long-chain alkane synthesis complex. *The Plant Cell*, 24(7), 3106-3118.
- Bourdenx, B., Bernard, A., Domergue, F., Pascal, S., Léger, A., Roby, D., & Joubès, J. (2011). Overexpression of Arabidopsis ECERIFERUM1 promotes wax very-long-chain alkane biosynthesis and influences plant response to biotic and abiotic stresses. *Plant Physiology*, 156(1), 29-45.
- Bruder, S., Moldenhauer, E. J., Lemke, R. D., Ledesma-Amaro, R., & Kabisch, J. (2019). Drop-in biofuel production using fatty acid photodecarboxylase from *Chlorella variabilis* in the oleaginous yeast *Yarrowia lipolytica*. *Biotechnology for biofuels*, 12(1), 1-13.
- Cantú Reinhard, F. G., Lin, Y. T., Stańczak, A., & de Visser, S. P. (2020). Bioengineering of cytochrome P450 OleTJE: How does substrate positioning affect the product distributions? *Molecules*, 25(11), 2675.



- Carlson, D. A., & Bolten, A. B. (1984). Identification of Africanized and European Honey Bees, Using Extracted Hydrocarbons. *Bulletin of the ESA*, 30(2), 32-35.
- Cavener, D. R. (1992). GMC oxidoreductases: a newly defined family of homologous proteins with diverse catalytic activities. *Journal of molecular biology*, 223(3), 811-814.
- Cha, H. J., Hwang, S. Y., Lee, D. S., Kumar, A. R., Kwon, Y. U., Voß, M., ... & Park, J. B. (2020). Whole-cell photoenzymatic cascades to synthesize long-chain aliphatic amines and esters from renewable fatty acids. *Angewandte Chemie*, 132(18), 7090-7094.
- Chanquia, S. N., Benfeldt, F. V., Petrovai, N., Santner, P., Hollmann, F., Eser, B. E., & Kara, S. (2022). Immobilization and Application of Fatty Acid Photodecarboxylase in Deep Eutectic Solvents. *ChemBioChem*, e202200482.
- Cheng, F., Li, H., Wu, D. Y., Li, J. M., Fan, Y., Xue, Y. P., & Zheng, Y. G. (2020). Light-driven deracemization of phosphinothricin by engineered fatty acid photodecarboxylase on a gram scale. *Green Chemistry*, 22(20), 6815-6818.
- Cistola, D. P., Hamilton, J. A., Jackson, D., & Small, D. M. (1988). Ionization and phase behavior of fatty acids in water: application of the Gibbs phase rule. *Biochemistry*, 27(6), 1881-1888.
- Coates, R. C., Podell, S., Korobeynikov, A., Lapidus, A., Pevzner, P., Sherman, D. H., & Gerwick, W. H. (2014). Characterization of cyanobacterial hydrocarbon composition and distribution of biosynthetic pathways. *PloS one*, 9(1), e85140.
- Dreveny, I., Gruber, K., Glieder, A., Thompson, A., & Kratky, C. (2001). The hydroxynitrile lyase from almond: a lyase that looks like an oxidoreductase. *Structure*, 9(9), 803-815.
- Eneh, O. C. (2011). A review on petroleum: source, uses, processing, products, and the environment. *Journal of Applied Sciences*, 11(12), 2084-2091.
- Ge, R., Zhang, P., Dong, X., Li, Y., Sun, Z., Zeng, Y., Bichuang, C., & Zhang, W. (2022). Photobiocatalytic decarboxylation for the synthesis of fatty epoxides from renewable fatty acids. *ChemSusChem*, 15(20), e202201275.
- Gelb, M. H., M. K. Jain, Hanel, H. M., Berg, O., (1995). Interfacial enzymology of glycerolipid hydrolases: lessons from secreted phospholipases A(2). *Annu Rev Biochem* 64: 653-688.
- Guo, X., Xia, A., Li, F., Huang, Y., Zhu, X., Zhang, W., ... & Liao, Q. (2022). Photoenzymatic decarboxylation to produce renewable hydrocarbon fuels: A comparison between whole-cell and broken-cell biocatalysts. *Energy Conversion and Management*, 255, 115311.
- Hackman, R. H. (1984). Cuticle: biochemistry. In *Biology of the Integument: Invertebrates*, Springer-Links, Jürgen Bereiter-Hahn, A. Gedeon Matoltsy, K. Sylvia Richards eds., 583-610.
- Hadley, N. F. (1984). Cuticle: ecological significance. *Biology of the integument: Invertebrates*, Springer-Links, Jürgen Bereiter-Hahn, A. Gedeon Matoltsy, K. Sylvia Richards eds., 685-693.
- Hall, C., Tharakan, P., Hallock, J., Cleveland, C., & Jefferson, M. (2003). Hydrocarbons and the evolution of human culture. *Nature*, 426(6964), 318-322.



- Han, J., McCarthy, E. D., Hoeven, W. V., Calvin, M., & Bradley, W. H. (1968). Organic geochemical studies, II. A preliminary report on the distribution of aliphatic hydrocarbons in algae, in bacteria, and in a recent lake sediment. *Proceedings of the National Academy of Sciences*, 59(1), 29-33.
- Hedison, T. M., Heyes, D. J., & Scrutton, N. S. (2022). Making molecules with photodecarboxylases: A great start or a false dawn?. *Current Research in Chemical Biology*, 2, 100017.
- Herman, N. A., & Zhang, W. (2016). Enzymes for fatty acid-based hydrocarbon biosynthesis. *Current opinion in chemical biology*, 35, 22-28.
- Howard, D. J., & Harrison, R. G. (1993). Reinforcement: origin, dynamics, and fate of an evolutionary hypothesis. In *Hybrid zones and the evolutionary process*, 46-69.
- Huijbers, M. M., Zhang, W., Tonin, F., & Hollmann, F. (2018). Light-driven enzymatic decarboxylation of fatty acids. *Angewandte Chemie International Edition*, 57(41), 13648-13651.
- Jenks, M. A., Tuttle, H. A., Eigenbrode, S. D., & Feldmann, K. A. (1995). Leaf epicuticular waxes of the eceriferum mutants in Arabidopsis. *Plant physiology*, 108(1), 369-377.
- Jetter, R., & Kunst, L. (2008). Plant surface lipid biosynthetic pathways and their utility for metabolic engineering of waxes and hydrocarbon biofuels. *The Plant Journal*, 54(4), 670-683.
- John B. Ohlrogge, Jan G. Jaworski, Post-Beittenmiller D. (1993) Lipid Metabolism in Plants. eBook ISBN9781351074070
- Kang, M. K., & Nielsen, J. (2017). Biobased production of alkanes and alkenes through metabolic engineering of microorganisms. *Journal of Industrial Microbiology and Biotechnology*, 44(4-5), 613-622.
- Kanicky, J. R., & Shah, D. O. (2002). Effect of degree, type, and position of unsaturation on the pKa of long-chain fatty acids. *Journal of colloid and interface science*, 256(1), 201-207.
- Kanicky, J. R., & Shah, D. O. (2003). Effect of premicellar aggregation on the pKa of fatty acid soap solutions. *Langmuir*, 19(6), 2034-2038.
- Keasling, J., Garcia Martin, H., Lee, T. S., Mukhopadhyay, A., Singer, S. W., & Sundstrom, E. (2021). Microbial production of advanced biofuels. *Nature Reviews Microbiology*, 19(11), 701-715.
- Klevens, H. B. (1948). Critical micelle concentrations as determined by refraction. *The Journal of Physical Chemistry*, 52(1), 130-148.
- Kumar, D., Pugazhendhi, A., Bajhaiya, A. K., & Gugulothu, P. (2021). Biofuel production from Macroalgae: present scenario and future scope. *Bioengineered*, 12(2), 9216.
- Lakavath, B., Hedison, T. M., Heyes, D. J., Shanmugam, M., Sakuma, M., Hoeven, R., ... & Scrutton, N. S. (2020). Radical-based photoinactivation of fatty acid photodecarboxylases. *Analytical Biochemistry*, 600, 113749.
- Lea-Smith, D. J., Ortiz-Suarez, M. L., Lenn, T., Nürnberg, D. J., Baers, L. L., Davey, M. P., ... & Howe, C. J. (2016). Hydrocarbons are essential for optimal cell size, division, and growth of cyanobacteria. *Plant Physiology*, 172(3), 1928-1940.
- Li, D., Han, T., Xue, J., Xu, W., Xu, J., & Wu, Q. (2021). Engineering fatty acid photodecarboxylase to enable highly selective decarboxylation of trans fatty acids. *Angewandte Chemie*, 133(38), 20863-20867.

- Li, F., Wu, X., Lam, P., Bird, D., Zheng, H., Samuels, L., ... & Kunst, L. (2008). Identification of the wax ester synthase/acyl-coenzyme A: diacylglycerol acyltransferase WSD1 required for stem wax ester biosynthesis in *Arabidopsis*. *Plant Physiology*, 148(1), 97-107.
- Li, J., Ma, Y., Liu, N., Eser, B. E., Guo, Z., Jensen, P. R., & Stephanopoulos, G. (2020). Synthesis of high-titer alkanes in *Yarrowia lipolytica* is enabled by a discovered mechanism. *Nature Communications*, 11(1), 6198.
- Li, N., Chang, W. C., Warui, D. M., Booker, S. J., Krebs, C., & Bollinger Jr, J. M. (2012). Evidence for only oxygenative cleavage of aldehydes to alkanes and formate by cyanobacterial aldehyde decarbonylases. *Biochemistry*, 51(40), 7908-7916.
- Liu B., & Benning C. (2013) Lipid metabolism in microalgae distinguishes itself. *Curr Opin Biotechnol*. 24(2):300-9.
- Liu, K., & Li, S. (2020). Biosynthesis of fatty acid-derived hydrocarbons: perspectives on enzymology and enzyme engineering. *Current opinion in biotechnology*, 62, 7-14.
- Liu, Y., Chen, J., Khusnutdinova, A. N., Correia, K., Diep, P., Batyrova, K. A., ... & Mahadevan, R. (2020). A novel C-terminal degron identified in bacterial aldehyde decarbonylases using directed evolution. *Biotechnology for biofuels*, 13, 1-11.
- Liu, Y., Liu, W. Q., Huang, S., Xu, H., Lu, H., Wu, C., & Li, J. (2023). Cell-free metabolic engineering enables selective biotransformation of fatty acids to value-added chemicals. *Metabolic Engineering Communications*, 16, e00217.
- Lockey, K. H. (1988). Lipids of the insect cuticle: origin, composition, and function. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry*, 89(4), 595-645.
- Mat Aron, N. S., Khoo, K. S., Chew, K. W., Show, P. L., Chen, W. H., & Nguyen, T. H. P. (2020). Sustainability of the four generations of biofuels—a review. *International Journal of Energy Research*, 44(12), 9266-9282.
- McQuarrie, D. A et al. (2012). General chemistry. 4th edition, University Science Books.
- Mendez-Perez, D., Begemann, M. B., & Pfleger, B. F. (2011). Modular synthase-encoding gene involved in  $\alpha$ -olefin biosynthesis in *Synechococcus* sp. strain PCC 7002. *Applied and environmental microbiology*, 77(12), 4264-4267.
- Mendez-Perez, D., Herman, N. A., & Pfleger, B. F. (2014). A desaturase gene involved in the formation of 1, 14-nonadecadiene in *Synechococcus* sp. strain PCC 7002. *Applied and environmental microbiology*, 80(19), 6073-6079.
- Metzger, P., & Largeau, C. (2005). *Botryococcus braunii*: a rich source for hydrocarbons and related ether lipids. *Applied microbiology and biotechnology*, 66, 486-496.
- Moulin, S., Légeret, B., Blangy, S., Sorigué, D., Burlacot, A., Auroy, P., ... & Beisson, F. (2019). Continuous photoproduction of hydrocarbon drop-in fuel by microbial cell factories. *Scientific Reports*, 9(1), 13713.
- Naik, S. N., Goud, V. V., Rout, P. K., & Dalai, A. K. (2010). Production of first- and second-generation biofuels: a comprehensive review. *Renewable and sustainable energy reviews*, 14(2), 578-597.
- Oldfield, E., & Lin, F. Y. (2012). Terpene biosynthesis: modularity rules. *Angewandte Chemie International Edition*, 51(5), 1124-1137.

- Pickl, M., Kurakin, S., Cantú Reinhard, F. G., Schmid, P., Pöcheim, A., Winkler, C. K., ... & Faber, K. (2018). Mechanistic studies of fatty acid activation by CYP152 peroxygenases reveal unexpected desaturase activity. *ACS catalysis*, 9(1), 565-577.
- Ralston, A. W., & Hoerr, C. W. (1942). The solubilities of the normal saturated fatty acids. *The Journal of Organic Chemistry*, 7(6), 546-555.
- Romero, C. M., Jiménez, E., & Suárez, F. (2009). Effect of temperature on the behavior of surface properties of alcohols in aqueous solution. *The Journal of Chemical Thermodynamics*, 41(4), 513-516.
- Rude, M. A., Baron, T. S., Brubaker, S., Alibhai, M., Del Cardayre, S. B., & Schirmer, A. (2011). Terminal olefin (1-alkene) biosynthesis by a novel P450 fatty acid decarboxylase from *Jeotgalicoccus* species. *Applied and environmental microbiology*, 77(5), 1718-1727.
- Rui, Z., Harris, N. C., Zhu, X., Huang, W., & Zhang, W. (2015). Discovery of a family of desaturase-like enzymes for 1-alkene biosynthesis. *Acs Catalysis*, 5(12), 7091-7094.
- Rui, Z., Li, X., Zhu, X., Liu, J., Domigan, B., Barr, I., ... & Zhang, W. (2014). Microbial biosynthesis of medium-chain 1-alkenes by a nonheme iron oxidase. *Proceedings of the National Academy of Sciences*, 111(51), 18237-18242.
- Santner, P., Szabó, L. K., Chanquia, S. N., Merrild, A. H., Hollmann, F., Kara, S., & Eser, B. E. (2021). Optimization and engineering of fatty acid photodecarboxylase for substrate specificity. *ChemCatChem*, 13(18), 4038-4046.
- Schirmer, A., Rude, M. A., Li, X., Popova, E., & Del Cardayre, S. B. (2010). Microbial biosynthesis of alkanes. *Science*, 329(5991), 559-562.
- Scott, D. (1986). Sexual mimicry regulates the attractiveness of mated *Drosophila melanogaster* females. *Proceedings of the National Academy of Sciences*, 83(21), 8429-8433.
- Scott, D., R. C. Richmond, and D. A. Carlson. (1988.) Pheromones exchanged during mating: A mechanism for mate assessment in *Drosophila*. *Anim. Behav.* 36:1164-1173
- Shiea, J., Brassell, S. C., & Ward, D. M. (1990). Mid-chain branched mono-and dimethyl alkanes in hot spring cyanobacterial mats: a direct biogenic source for branched alkanes in ancient sediments? *Organic Geochemistry*, 15(3), 223-231.
- Simić, S., Jakstaite, M., Huck, W. T., Winkler, C. K., & Kroutil, W. (2022). Strategies for transferring photobiocatalysis to continuous flow exemplified by photodecarboxylation of fatty acids. *ACS catalysis*, 12(22), 14040-14049.
- Small, D. M. (1968). A classification of biologic lipids based upon their interaction in aqueous systems. *Journal of the American Oil Chemists Society*, 45(3), 108-119.
- Sorigué, D., Légeret, B., Cuiné, S., Blangy, S., Moulin, S., Billon, E., ... & Beisson, F. (2017). An algal photoenzyme converts fatty acids to hydrocarbons. *Science*, 357(6354), 903-907.
- Sorigué, D., Légeret, B., Cuiné, S., Morales, P., Mirabella, B., Guédeney, G., ... & Beisson, F. (2016). Microalgae synthesize hydrocarbons from long-chain fatty acids via a light-dependent pathway. *Plant Physiology*, 171(4), 2393-2405.
- Sukovich, D. J., Seffernick, J. L., Richman, J. E., Gralnick, J. A., & Wackett, L. P. (2010). Widespread head-to-head hydrocarbon biosynthesis in bacteria and role of OleA. *Applied and environmental microbiology*, 76(12), 3850-3862.

- Sun, Y., Calderini, E., & Kourist, R. (2021). A reconstructed common ancestor of the fatty acid photo-decarboxylase clade shows photo-decarboxylation activity and increased thermostability. *ChemBioChem*, 22(10), 1833-1840.
- Sützl, L., Foley, G., Gillam, E. M., Bodén, M., & Haltrich, D. (2019). The GMC superfamily of oxidoreductases revisited: analysis and evolution of fungal GMC oxidoreductases. *Biotechnology for Biofuels*, 12, 1-18.
- Tornabene, T. G., Gelpi, E., & Oro, J. (1967). Identification of fatty acids and aliphatic hydrocarbons in *Sarcina lutea* by gas chromatography and combined gas chromatography-mass spectrometry. *Journal of Bacteriology*, 94(2), 333-343.
- Trisrivirat, D., Hughes, J. M., Hoeven, R., Faulkner, M., Toogood, H., Chaiken, P., & Scrutton, N. S. (2020). Promoter engineering for microbial bio-alkane gas production. *Synthetic Biology*, 5(1), ysaa022.
- Warui, D. M., Li, N., Nørgaard, H., Krebs, C., Bollinger Jr, J. M., & Booker, S. J. (2011). Detection of formate, rather than carbon monoxide, as the stoichiometric coproduct in conversion of fatty aldehydes to alkanes by a cyanobacterial aldehyde decarbonylase. *Journal of the American Chemical Society*, 133(10), 3316-3319.
- Wijffels, R. H., & Barbosa, M. J. (2010). An outlook on microalgal biofuels. *Science*, 329(5993), 796-799.
- Winters, K., Parker, P. L., & Van Baalen, C. (1969). Hydrocarbons of blue-green algae: geochemical significance. *Science*, 163(3866), 467-468.
- Wu, Y., Paul, C. E., & Hollmann, F. (2021). Stabilisation of the fatty acid decarboxylase from *Chlorella variabilis* by caprylic acid. *ChemBioChem*, 22(14), 2420-2423.
- Xu, J., Hu, Y., Fan, J., Arkin, M., Li, D., Peng, Y., ... & Wu, Q. (2019). Light-driven kinetic resolution of  $\alpha$ -functionalized carboxylic acids enabled by an engineered fatty acid photodecarboxylase. *Angewandte Chemie International Edition*, 58(25), 8474-8478.
- Xu, W., Chen, Y., Li, D., Wang, Z., Xu, J., & Wu, Q. (2022). Rational design of fatty acid photodecarboxylase enables the efficient decarboxylation of medium-and short-chain fatty acids for the production of gasoline bio-alkanes. *Molecular Catalysis*, 524, 112261.
- Yamamori, T., Kageyama H., Tanaka Y, Takabe T., (2018). Requirement of alkanes for salt tolerance of Cyanobacteria: characterization of alkane synthesis genes from salt-sensitive *Synechococcus elongatus* PCC7942 and salt-tolerant *Aphanothece halophytica*. *Letters in Applied Microbiology* 67(3), 299–305.
- Young, M., Artsatbanov, V., Beller, H. R., Chandra, G., Chater, K. F., Dover, L. G., ... & Greenblatt, C. L. (2010). Genome sequence of the Fleming strain of *Micrococcus luteus*, a simple free-living actinobacterium. *Journal of bacteriology*, 192(3), 841-860.
- Yunus, I. S., Wichmann, J., Wördenweber, R., Lauersen, K. J., Kruse, O., & Jones, P. R. (2018). Synthetic metabolic pathways for photobiological conversion of CO<sub>2</sub> into hydrocarbon fuel. *Metabolic Engineering*, 49, 201-211.
- Zeng, Y. Y., Liu, L., Chen, B. S., & Zhang, W. (2021). Light-Driven Enzymatic Decarboxylation of Dicarboxylic Acids. *Chemistry Open*, 10(5), 553-559.

Zhang, B., Rajakovich, L. J., Van Cura, D., Blaes, E. J., Mitchell, A. J., Tysoe, C. R., ... & Bollinger Jr, J. M. (2019). Substrate-triggered formation of a peroxo-Fe<sup>2</sup> (III/III) intermediate during fatty acid decarboxylation by UndA. *Journal of the American Chemical Society*, 141(37), 14510-14514.

Zhang, W., Ma, M., Huijbers, M. M., Filonenko, G. A., Pidko, E. A., van Schie, M., ... & Hollmann, F. (2019). Hydrocarbon synthesis via photoenzymatic decarboxylation of carboxylic acids. *Journal of the American Chemical Society*, 141(7), 3116-3120.

Ziogas, A., Pennemann, H., & Kolb, G. (2020). Electrochemical synthesis of tailor-made hydrocarbons from organic solvent free aqueous fatty acid mixtures in a micro flow reactor. *Electrocatalysis*, 11, 432-442.

Zuo, Z. (2019). Why algae release volatile organic compounds—the emission and roles. *Frontiers in microbiology*, 10, 491.



# Objectives of the thesis.

## A-Explore the biodiversity of FAPs

When I started my thesis in September 2018, the FAP enzyme had already been discovered and characterized in the model green microalga *Chlorella variabilis*, a medium-resolution crystal structure was known and a photocycle had been proposed. In addition, a phylogenetic analysis had shown that putative homologs of *Chlorella variabilis* FAP (CvFAP) were present in several algae and a few FAP cDNAs from algae of various origin were available in the laboratory. One of the most important remaining questions was to determine whether the fatty acid photodecarboxylase activity was conserved in other algae, especially beyond green algae (red algae, diatoms etc.). However, a major issue to answer this question was that the *E. coli* expression conditions used to produce CvFAP were not appropriate for any of the homologs tested. My first objective was therefore to produce some soluble form of FAPs from other green algae and also from other groups (at least one or two FAPs) and test the enzyme activity to determine if it was a fatty acid decarboxylase and if it was light-dependent as CvFAP.

**B-Contribute to a multidisciplinary study on the mechanism of FAP** Following the discovery of FAP in 2016, based on transient absorption spectroscopy, a mechanism of the photocycle was proposed in the paper revealing the discovery of the enzyme. A crystal structure was also published in the same article (The first mechanism and crystal structure are described in the bibliographic introduction.). However, the structure was a medium-resolution one and the mechanism was incomplete. Some parts of the structure were not resolved and some steps of the photocycle were still hypothetical. Further spectroscopic and crystallographic studies with a more exhaustive palette of techniques, as well as the use of mutants of active-site residues were thus necessary. The second objective of my thesis was therefore to contribute to the large efforts of an international consortium of laboratories to perform detailed crystallographic and spectroscopic studies on CvFAP and variants. In this context, I was able to perform protein purification and biochemical and structural characterization of CvFAP (wild type or mutated) in my laboratory with Dr. Damien Sorigué, as well as transient absorption spectroscopy experiments in the laboratory of Dr. Pavel Müller and Dr. Klaus Brettel in CEA Saclay.

### **C-Characterize the substrate specificity of FAP and optimize the photoconversion of SC and MC fatty acids substrates.**

The first study of CvFAP showed that this biocatalyst can convert a wide range of free fatty acids from lauric acid (12:0) to behenic acid (22:0) but had preference for C16-C18 fatty acids. Several studies confirmed this preference for LC fatty acids. However, most of the biocatalysis experiments on SC and MC fatty acids reported in the literature were done using crude extracts from *E. coli* and looked only at the chemical yield of the FAP (by quantifying hydrocarbon production). In order to reinvestigate the activity of CvFAP on SC and MC fatty acids, we decided to do a more thorough substrate specificity study of the enzyme by combining transient absorption spectroscopy and optimization of in vitro activity. The third objective of my thesis was thus dual: (i) understand at the photochemical level if, using different fatty acids, there were any differences in the spectroscopic characteristics of FAP, (ii) investigate or reinvestigate more systematically the influence of various parameters on the activity of purified FAP (fatty acid chain length, fatty acid concentration, pH, light intensity, etc.).



## **Chapter 2: Exploring the biodiversity of fatty acid photodecarboxylase**



## Objectives, personal contribution, and main results

Molecular phylogeny analysis of GMC oxidoreductases (Sorigué et al. [2017](#)) has shown that a homolog of CvFAP is present in many algal species (macroalgae or microalgae) from all kinds of phylogenetic groups (green, algae, red algae, secondary endosymbiotic algae). The first main objective of this study was therefore to identify more putative FAPs in algal genomes and determine whether the fatty acid decarboxylase activity was conserved in some selected putative FAPs from various algal groups and whether it was also light-dependent. Since we were also interested in improving the continuous production of volatile hydrocarbons in a microbial system (Moulin et al. [2019](#)), the idea of this first objective was also to be able to have a first idea of whether all FAPs are likely to have a substrate specificity profile similar to CvFAP or, if among FAP biodiversity, we are likely to find natural variants with preference for other chain lengths, especially shorter than C16-C18. For this purpose, we selected a limited set of less than 10 putative FAPs of microalgae or macroalgae from a diversity of algal groups. The second main objective was about FAP biological role(s) which were still elusive. If FAP activity was conserved in algae, it was logical to imagine that hydrocarbons generated by FAP would play a structural, metabolic or regulatory role in relation to light. Through phenotypic characterization of a FAP knockout mutant in the green microalga *Chlamydomonas reinhardtii*, we therefore aimed to start the investigate the possible biological role(s) of FAP.

**Personal contribution:** To determine whether CvFAP homologs also have FAP activity or not, I performed heterologous expression in *E. coli*. The *E. coli* strain commonly used for CvFAP expression was BL21 pRil; unfortunately, we found that several CvFAP homologs could not be produced correctly in this strain. My first task was thus to identify another strain of *E. coli* able to produce other FAPs in soluble form and at a level sufficient for activity assay. To do this, I performed the expression of the selected FAPs in a dozen of *E. coli* protein expression strains under various conditions and estimated levels of soluble and total FAP using western blots. After identification of the best strain and conditions, I carried out the experiments of heterologous production on a larger scale and quantified the hydrocarbons and fatty acids produced by the bacterial cultures under light or in the dark.

**Main results:** Heterologous expression of a few selected homologs of CvFAP showed that FAP activity was indeed likely to be conserved in algae. In addition, the selected FAPs could produce C13-C17 hydrocarbons with a variety of profiles. Regarding the role of *Chlamydomonas* FAP,

the subcellular fractionation data show that FAP and its products are found in the thylakoid membrane fraction; the study of FAP knockouts showed that their photosynthesis was affected during light intensity variations under cold conditions. Taken together, the results suggest that in *Chlamydomonas* the light-dependent hydrocarbon-forming activity of FAP is likely to be related to the normal functioning of photosynthetic membranes, and that this role is probably conserved in many algae. Another major information coming from this work is that FAPs do not all have the same fatty acid specificity.

This work was published in the journal Plant Physiology in 2021. Publication link:

<https://academic.oup.com/plphys/article/186/3/1455/6226522>

## **Article 1:**

**Fatty acid photodecarboxylase is an ancient  
photoenzyme that forms hydrocarbons in the  
thylakoids of algae**





# Fatty acid photodecarboxylase is an ancient photoenzyme that forms hydrocarbons in the thylakoids of algae

Solène L.Y. Moulin,<sup>1,†</sup> Audrey Beyly-Adriano,<sup>1</sup> Stéphan Cuiné ,<sup>1</sup> Stéphanie Blangy ,<sup>1</sup> Bertrand Légeret ,<sup>1</sup> Magali Floriani ,<sup>2</sup> Adrien Burlacot ,<sup>1,‡</sup> Damien Sorigué ,<sup>1</sup> Poutoum-Palakiyem Samire ,<sup>1</sup> Yonghua Li-Beisson ,<sup>1</sup> Gilles Peltier <sup>1</sup> and Fred Beisson <sup>1,\*,\$</sup>

1 CEA, CNRS, Aix-Marseille University, Institute of Biosciences and Biotechnologies of Aix-Marseille (BIAM), UMR7265, CEA Cadarache, 13108 Saint-Paul-lez-Durance, France

2 Institut de Radioprotection et de Sécurité Nucléaire (IRSN), PRP-ENV/SRTE/LECO, Cadarache, 13108 Saint-Paul-Lez-Durance, France

\*Author for communication: frederic.beisson@cea.fr

†Present address: Stanford University, 279 Campus Dr, Stanford, CA 94305

‡Present address: Howard Hughes Medical Institute, Department of Plant and Microbial Biology, 111 Koshland Hall, University of California, Berkeley, CA 94720-3102, USA.

\$Senior author.

F.B. conceived the original research project; F.B., S.L.Y.M., and G.P. designed the experiments and analyzed the data; S.L.Y.M., A.B.-A., S.C., S.B., D.S., B.L., A.B., P.P.S. performed experiments; M.F. performed the transmission electronmicroscopic study; S.L.Y.M. performed phylogenetic analysis; F.B. and S.M. wrote the article with contributions from Y.L.-B. and G.P.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (<https://academic.oup.com/plphys/pages/general-instructions>) is: Fred Beisson (frederic.beisson@cea.fr).

## Abstract

Fatty acid photodecarboxylase (FAP) is one of the few enzymes that require light for their catalytic cycle (photoenzymes). FAP was first identified in the microalga *Chlorella variabilis* NC64A, and belongs to an algae-specific subgroup of the glucose–methanol–choline oxidoreductase family. While the FAP from *C. variabilis* and its *Chlamydomonas reinhardtii* homolog CrFAP have demonstrated in vitro activities, their activities and physiological functions have not been studied in vivo. Furthermore, the conservation of FAP activity beyond green microalgae remains hypothetical. Here, using a *C. reinhardtii* FAP knockout line (*fap*), we showed that CrFAP is responsible for the formation of 7-heptadecene, the only hydrocarbon of this alga. We further showed that CrFAP was predominantly membrane-associated and that >90% of 7-heptadecene was recovered in the thylakoid fraction. In the *fap* mutant, photosynthetic activity was not affected under standard growth conditions, but was reduced after cold acclimation when light intensity varied. A phylogenetic analysis that included sequences from Tara Ocean identified almost 200 putative FAPs and indicated that FAP was acquired early after primary endosymbiosis. Within Bikonta, FAP was retained in secondary photosynthetic endosymbiosis lineages but absent from those that lost the plastid. Characterization of recombinant FAPs from various algal genera (*Nannochloropsis*, *Ectocarpus*, *Galdieria*, *Chondrus*) provided experimental evidence that FAP photochemical activity was present in red and brown algae, and was not limited to unicellular species. These results thus indicate that FAP was conserved during the evolution of most algal lineages where photosynthesis was retained, and suggest that its function is linked to photosynthetic membranes.

Received August 17, 2020. Accepted March 07, 2021. Advance access publication April 15, 2021

© American Society of Plant Biologists 2021. All rights reserved. For permissions, please email: journals.permissions@oup.com

## Introduction

Most organisms have the ability to synthesize highly hydrophobic compounds made only of carbon and hydrogen called hydrocarbons (HCs). Many HCs are isoprenoids but others like *n*-alkanes and their unsaturated analogues (*n*-alkenes) derive from fatty acids (Herman and Zhang, 2016). In plants, C<sub>29</sub>–C<sub>35</sub> *n*-alkanes are synthesized in the epidermis from very-long-chain fatty acids, and secreted onto the surface of aerial organs (Lee and Suh, 2013). Plant *n*-alkanes are important for adaptation to the terrestrial environment because they constitute a major part of the wax layer of the extracellular cuticle that prevents the loss of internal water. Occurrence of *n*-alka(e)nes has also been reported in microorganisms such as microalgae (Sorigué et al., 2016) and cyanobacteria (Schirmer et al., 2010). The magnitude and biogeochemical importance of HC production by marine cyanobacteria (and possibly microalgae) is an intriguing question (Lea-Smith et al., 2015; Valentine and Reddy, 2015). Microbial HCs are presumably mostly located in membranes, but their physiological function has not been studied much. While in cyanobacteria *n*-alka(e)nes have been proposed to play roles in cell growth, cell division, photosynthesis, and salt tolerance (Berla et al., 2015; Lea-Smith et al., 2016; Yamamori et al., 2018; Knoop and Pakrasi, 2019), in microalgae the biological function of *n*-alka(e)nes is still completely unknown. Besides the elucidation of their biological roles, *n*-alkanes and *n*-alkenes have also attracted attention because of their interest as fuels, cosmetics, lubricants, and as synthons in organic chemistry (Jetter and Kunst, 2008; Jiménez-Díaz et al., 2017). Bio-based alka(e)ne production would be highly desirable to replace part of petroleum-derived HCs, and is currently the focus of intense research efforts (Zhou et al., 2018; Liu and Li, 2020).

A number of *n*-alka(e)ne-forming enzymes have been identified and characterized in the last decade, and it is now clear that conversion of fatty acids to HCs occurs through a variety of reactions and proteins (Herman and Zhang, 2016). Besides, the biosynthetic enzymes involved in the same types of reactions are not conserved across phylogenetic groups. For instance, it has been shown in bacteria that synthesis of terminal olefins (1-alkenes) occurs through decarboxylation of a saturated long-chain fatty acid, and that this reaction is catalyzed by a cytochrome P450 in *Jeotgalicoccus* spp. (Rude et al., 2011) and by a nonheme iron oxidase in *Pseudomonas* (Rui et al., 2014). In the bacterium *Micrococcus luteus*, yet another pathway has been discovered, which consists of the head-to-head condensation of fatty acids to form very-long-chain *n*-alkenes with internal double bonds (Beller et al., 2010). Cyanobacterial *n*-alka(e)nes are produced by two distinct pathways (Schirmer et al., 2010; Mendez-Perez et al., 2011), which have been shown to be mutually exclusive (Coates et al., 2014). In insects, a cytochrome P450 oxidative decarbonylase acts on an aldehyde to produce cuticular alka(e)nes (Qiu et al., 2012). The plant pathway producing the very-long-

chain *n*-alkanes of the cuticular waxes may also involve an aldehyde intermediate, and is known to require the ECERIFERUM (CER) proteins: CER1 and CER3 (Bernard et al., 2012), which are homologous to the oxygen-dependent membrane class of diiron fatty acid desaturases.

In microalgae, we have shown that C<sub>15</sub>–C<sub>17</sub> *n*-alka(e)nes occur in *Chlorella variabilis* NC64A (named *C. variabilis* from here on), and that they are synthesized through decarboxylation of long-chain fatty acids (Sorigué et al., 2016). A *C. variabilis* protein with a fatty acid decarboxylase activity was then identified as a photoenzyme (Sorigué et al., 2017), a rare type of enzyme that requires photons at each catalytic cycle (Björn, 2015). The *C. variabilis* protein was thus named fatty acid photodecarboxylase (FAP, E.C. 4.1.1.106). It is one of the few photoenzymes discovered so far, the others being the reactional centers of the photosystems, DNA photolyases, and light-dependent protochlorophyllide oxidoreductase. FAP belongs to a family of flavoproteins (Sorigué et al., 2017), the glucose–methanol–choline (GMC) oxidoreductases, which includes a large variety of enzymes present in prokaryotic and eukaryotic organisms (Zámocký et al., 2004). FAP activity thus represents an additional type of chemistry in the GMC oxidoreductase family (Sorigué et al., 2017). Molecular phylogenetic analysis has shown that *C. variabilis* FAP and CrFAP belong to an algal branch of GMC oxidoreductases. However, whether FAP activity is conserved in other algal lineages beyond green algae and whether FAP is indeed responsible for *n*-alka(e)ne formation in vivo remain to be demonstrated. Besides, the subcellular location and role of FAP in algal cells have not yet been investigated.

In this work, we isolate and characterize in *C. reinhardtii* an insertional mutant deficient in FAP (*fap* mutant strain). We show that FAP is indeed responsible for the formation of 7-heptadecene, the only fatty acid-derived HC present in this alga. In addition, we provide evidence for a thylakoid localization of *C. reinhardtii* FAP and its alkene product. We also show that growth and photosynthesis are not affected in the knockout under laboratory conditions, but photosynthetic efficiency is impacted under cold conditions when light intensity varies. Finally, we build a large molecular phylogeny of GMC oxidoreductases based on TARA Ocean data and identify almost 200 uncharacterized putative FAP sequences across algal lineages. Experimental evidence demonstrates that FAP photochemical activity is conserved in red and brown algae and, since it is also present in macroalgae, is not limited to unicellular species.

## Results

### FAP is responsible for alkene synthesis in *C. reinhardtii*

*Chlamydomonas reinhardtii* (hereafter named *C. reinhardtii*) has been previously shown to produce 7-heptadecene (C<sub>17</sub>:1-alkene) from cis-vaccenic acid (Sorigué et al., 2016) and to have a FAP homolog that can also perform photodecarboxylation of



fatty acids in vitro (Sorigué et al., 2017). Although it seemed likely that FAP proteins are indeed responsible for the synthesis of alka(e)nes produced by *C. reinhardtii* and *C. variabilis*, the possibility that the alkenes are formed in vivo by another enzyme could not be ruled out. In order to address this issue and investigate the biological role of FAP, a *C. reinhardtii* strain mutated for FAP was isolated from the *Chlamydomonas* Library Project (CLiP; Li et al., 2016). This strain showed undetectable levels of FAP protein (Figure 1) and was named *fap-1*. The only fatty acid-derived HC in *C. reinhardtii*, i.e. 7-heptadecene, could not be detected in *fap-1* knockout or in the two other *fap* mutants isolated (Figure 1). After performing nuclear complementation of *fap-1* using the genomic FAP gene expressed under the promoter *PsaD*, four independent transformants with different expression levels of *C. reinhardtii* FAP (CrFAP) were isolated (named Cp-1–4). In these complemented strains, production of 7-heptadecene was clearly related to FAP amount (Figure 1). These results thus demonstrate that CrFAP is responsible for alkene formation in vivo in *C. reinhardtii*. Cp-4 was one of the three

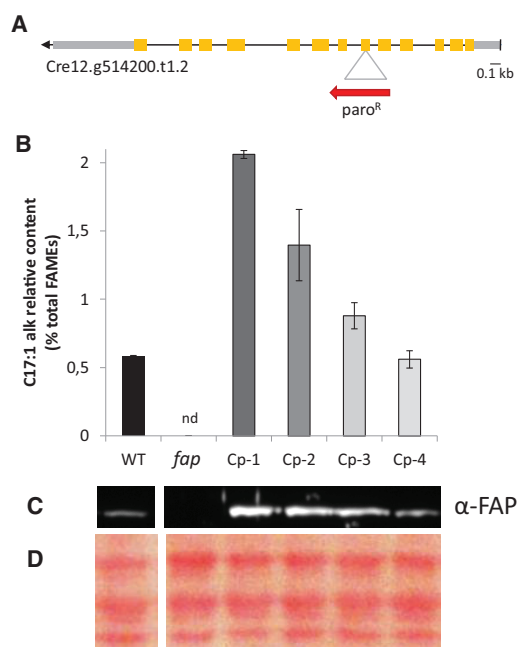
complemented strains with 7-heptadecene levels similar to the wild-type (WT) strains used for physiological studies (see “Materials and methods” for the isolation of the WT, *fap*, and Cp strains).

### FAP activity is conserved beyond green microalgae

Molecular phylogeny of GMC oxidoreductases has previously shown that CrFAP and the FAP from *C. variabilis* (CvFAP) are present in an evolutionary branch containing only sequences from algae (Sorigué et al., 2017). The term “algae” is used here in the classical sense of photosynthetic organisms that have chlorophyll *a* as their primary photosynthetic pigment and lack a sterile covering of cells around the reproductive cells (Lee, 2008). To investigate whether FAP activity has been conserved in algal groups other than green algae, genes encoding putative FAPs from selected algal lineages were cloned and expressed in *Escherichia coli*, and the bacterial HC content was analyzed. Considering the basal position of red algae, we decided to explore FAP activity in Rhodophytes, selecting the microalga *Galdieria sulphuraria* and the macroalga *Chondrus crispus*. For algae deriving from secondary endosymbiosis, we also chose the microalga *Nannochloropsis gaditana* and the macroalga *Ectocarpus siliculosus*. When cultivated under light, all *E. coli* strains expressing the various FAPs produced a range of *n*-alkanes and *n*-alkenes with different chain lengths (C15–C17) in various proportions (Figure 2A). The small amount of HCs detected in *E. coli* cells cultivated under dark was not observed in the nontransformed strain or culture medium, and is thus likely to be due to brief illumination required in the cell treatment process. Interestingly, among the FAPs investigated, we observed some profiles of conversion of *E. coli* fatty acids to HCs that are different from that of CvFAP (Figure 2B). Taken together, these results therefore show that FAP photoenzymatic activity (1) is present in red algae, (2) has been conserved in algae with secondary plastids, (3) is not limited to unicellular algae, and (4) is likely to be diverse with respect to fatty acid specificity.

### Identification of a reservoir of putative FAPs

To provide a wider picture of the occurrence and evolution of putative FAP photoenzymes within algal groups and to increase the reservoir of FAPs for future biotechnological purposes, a large phylogenetic analysis of GMC oxidoreductases sequences was conducted. We used GMC oxidoreductases retrieved from public databases, from sequenced algal genomes (Blaby-Haas and Merchant, 2019) and from the Tara Ocean project (de Vargas et al., 2015). Tara data gave a unique opportunity to enlarge the FAP dataset with marine algal species that may not be easy to grow under laboratory conditions and whose genomes have not been sequenced. Protein sequences sharing between 50% and 33% of homology with the sequence of *C. variabilis* FAP were retrieved using Basic Local Alignment Search Tool (BLAST) searches in algal genomes and in the TARA dataset (Supplemental Tables S1 and S2). Over 500 putative GMC oxidoreductases were thus identified. Additional putative GMC



**Figure 1** FAP level corresponds to the amount of 7-heptadecene in *C. reinhardtii* strains mutated in FAP. A, CrFAP gene structure and site of cassette insertion in *fap-1* knockout strain. Yellow boxes: exons; gray boxes: untranscribed regions; red arrow: antibiotic resistance cassette. B, Relative content of total fatty-acid-derived HCs measured on whole cells; the only fatty acid-derived HC of *C. reinhardtii* is 7-heptadecene (abbreviated as C17:1 alk). Values are mean  $\pm$  SD of  $n = 3$  independent experiments for each strain. nd: not detected. C, Immunoblot of total protein extract probed with anti-FAP antibody. D, Loading control of the immunoblot (Ponceau staining). B, C, and D, WT: wild type strains; *fap*: FAP knockout strains; Cp-1–4: strains complemented with FAP gene (*fap-1* background). In C and D, each track corresponds to the genotype situated just above in (B).

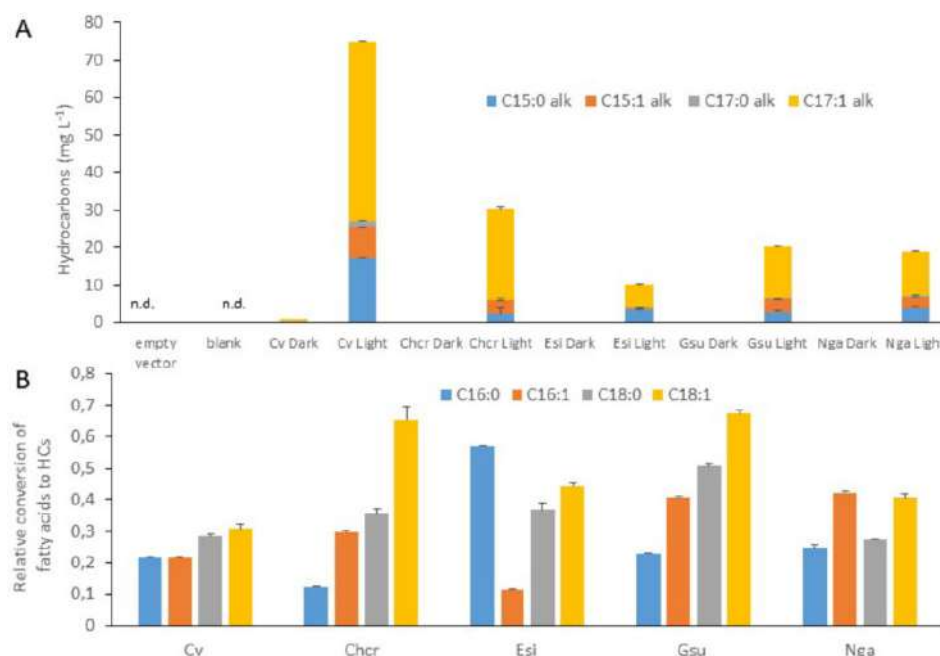
oxidoreductases from various taxa of the three domains of life were selected from public databases in order to have a representative diversity of the whole GMC oxidoreductase superfamily.

Molecular phylogenetic analysis identified almost 200 sequences of putative GMC oxidoreductases that grouped with the five FAPs that have demonstrated biochemical activity (CvFAP, CrFAP, and the four additional FAPs characterized in this study). These sequences constitute a reservoir of putative FAPs. Sequences in this FAP group all belonged to algal species (Figure 3; Supplemental Figure S1). Logo sequences built with the putative FAPs exhibited highly conserved patterns, some of which were absent in the other GMC oxidoreductases (Supplemental Figure S2). Interestingly, the residues C432 and R451 of CvFAP active site (Sorigué et al., 2017; Heyes et al., 2020) were strictly conserved, and this feature was specific to the FAP group. It is thus likely that all proteins of this group are genuine FAPs.

Sequences from plants as well as other streptophytes (including charophytes) did not group with algal FAPs (Figure 3; Supplemental Figure S1). Absence of FAP in charophytes indicated early loss of FAP function in streptophytes. No putative FAP sequences could be found in cyanobacteria, although this group is highly represented in TARA data (de Vargas et al., 2015). Phylogeny within the FAP branch

indicated that red algae (rhodophytes) sequences were the most basal. Interestingly, FAP sequences from secondary endosymbiosis-derived species appeared to be more closely related to FAPs of green algae (chlorophytes) than red algae.

Overall, the putative FAPs that could be identified in algae were present in a variety of algal groups, including stramenopiles (heterokonts), haptophytes, and dinophytes. Most eukaryotic algae harbored one putative FAP and no other GMC oxidoreductase, but a few algae contained no FAP and/or several non-FAP GMC oxidoreductases (Figure 4). Indeed, no putative FAP could be found in the sequenced algal genomes of the glaucocystophyte *Cyanophora paradoxa*; Mamiellophyceae species: *Ostreococcus* spp., *Micromonas* spp., and *Bathycoccus* spp.; and the diatom *Thalassiosira pseudonana*. Conversely, only a few algal sequences could be found in other branches of the GMC oxidoreductase family. Existence in the diatom *Thalassiosira pseudonana* of a GMC oxidoreductase grouping with bacterial choline dehydrogenase was supported by one sequence from the sequenced genome (Tps-GMC) and one sequence from Tara (48230190). A Tara sequence annotated as a *Pelagomonas* protein (5166790) also turned out not to be located on the FAP clade. The cryptophyte *Guillardia theta* contained three different GMC oxidoreductases in three different branches but none of them grouped with FAPs. *Ulva*



**Figure 2** Hydrocarbons produced in *E. coli* strains expressing various microalgal FAPs. HC content was analyzed by GC–MS after transmethylation of whole cells. Data are means  $\pm$  SD of  $n = 3$  independent cultures. Cv, *C. variabilis* NC64A; Chcr, *Chondrus crispus*; Esi, *Ectocarpus siliculosus*; Gsu, *Galdieria sulphuraria*; Nga, *Nannochloropsis gaditana*; nd, not detected. A, Amount of hydrocarbon produced in *E. coli* strains cultivated under light or dark. Empty vector strain was cultivated under light. Blank corresponds to culture medium alone. B, Relative conversion of *E. coli* major fatty acids to corresponding HCs. Data represent the amount of a HC produced under light in a FAP-expressing strain divided by the amount of the fatty acid precursor and by the total amount of HCs.

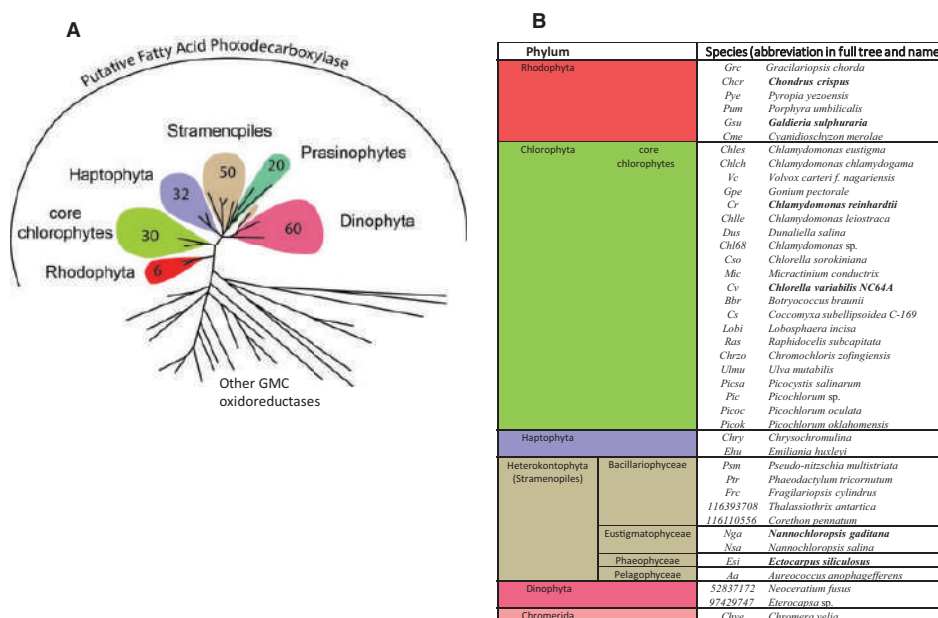
*mutabilis* had 11 predicted GMCs, but only one was in the FAP clade. The 10 other members of this multigene family of *Ulva* appeared to form a group close to plant GMC oxidoreductases. Although exceptions similar to these probably exist in algal diversity, the general picture is that most algae have only one GMC oxidoreductase and that it groups with the characterized FAP clade in the phylogenetic tree.

### *Chlamydomonas reinhardtii* FAP binds membranes, and most of its products are found in the thylakoid fraction

CrFAP is predicted by PredAlgo (Tardif et al., 2012), a software dedicated to the analysis of subcellular targeting sequences in green algae, to be localized to the chloroplast. Consistently, CrFAP is found in a set of 996 proteins proposed to be chloroplastic in *C. reinhardtii* (Terashima et al., 2011). A broader study of putative FAP targeting peptides, also using PredAlgo, indicated that FAPs from various green and red algae were largely predicted to be chloroplastic (Supplemental Figure S3). In algae with secondary plastids (i.e. containing three or four membranes), the presence of a signal peptide was consistent with targeting to the ER or chloroplast ER (CER) membrane. Further analysis performed using ASAFind (Gruber et al., 2015), a prediction tool designed to recognize CER targeting motifs in signal peptides, indicated that such a motif was present in *Ectocarpus*

*siliculosus* and *N. gaditana*. Taken together, these results suggest that FAP homologs are very likely to be localized to chloroplasts in green algae, in red algae, and also in at least some of the algae that acquired plastids through secondary endosymbiosis.

Since CrFAP is a soluble protein (Sorigué et al., 2017) and is predicted to be plastidial, we then sought to determine whether it is a stromal protein with no affinity for membranes or if it has some ability to bind thylakoid membranes. Subcellular fractionation of *C. reinhardtii* cells was therefore performed (Figure 5A). Most of the FAP protein was found in the total membrane fraction. Thylakoid-enriched membranes were then isolated from whole cells using a sucrose-gradient purification procedure (Figure 5A). Co-purification with thylakoids was performed using the D2 protein (PsbD), a thylakoid membrane protein from the photosystem II (PSII) core complex. The fact that the phosphoribulokinase, used as a control for stroma, could barely be detected in our thylakoid fraction, indicated the presence of little intact chloroplasts or cells. After the thylakoid purification procedure, the amount of CrFAP still bound to this fraction varied between purifications, but the protein was always detectable. A salt wash experiment confirmed that the CrFAP remaining in the thylakoid fraction was tightly bound to the thylakoids, and thus not likely to be a contaminating stromal protein. When analyzing the percentage of 7-



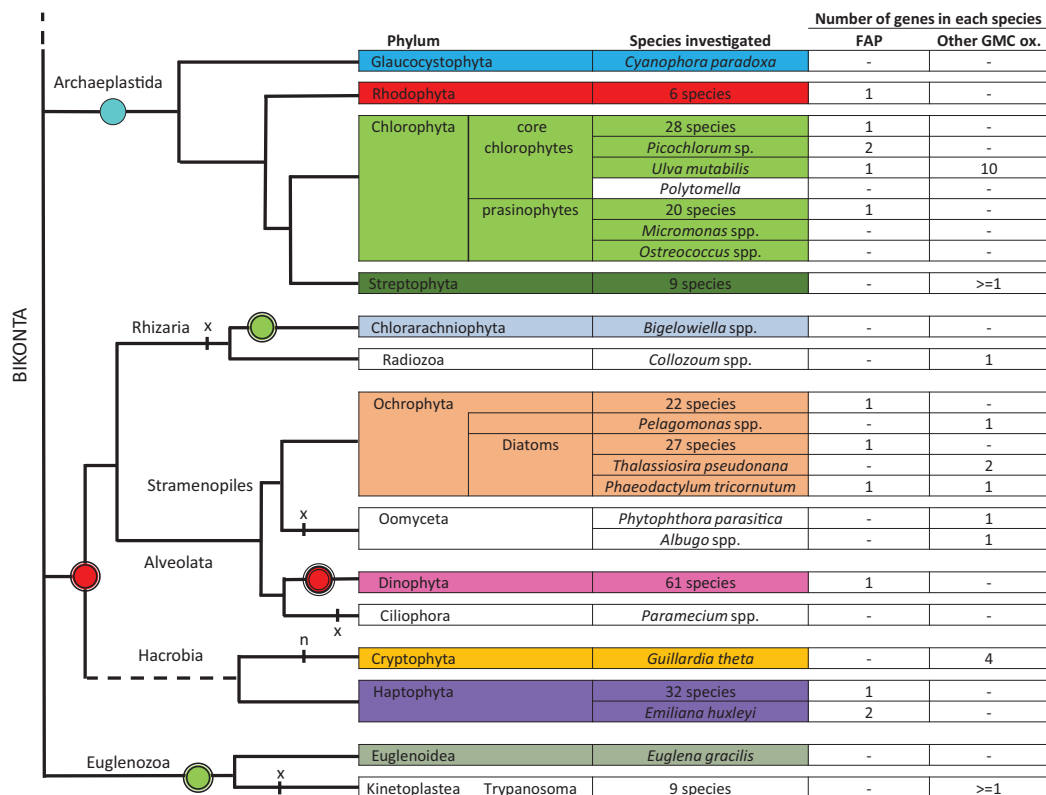
**Figure 3** Identification of a set of putative FAPs across algal groups. A, Simplified circular tree of GMC oxidoreductases showing the number of putative FAP sequences found in each group of algae. All 198 identified putative FAPs belong to algae and have been found in TARA data (161 FAPs) or in sequenced algal genomes (37 FAPs). The tree was built using maximum likelihood algorithm using GMC oxidoreductases from various kingdoms. Annotations are focused on the branch of putative FAPs; other branches are other GMC oxidoreductases. Branches have been collapsed; full tree is available in Supplemental Figure S1. B, Names of algae species with at least one putative FAP. For most algal groups, the number of species listed in B is lower than that indicated in A because many species from TARA data have no annotation down to the species level. When the biochemical activity of the FAP homolog is demonstrated (Sorigué et al., 2017 or this study), species names are indicated in bold.

heptadecene in total fatty acids in whole cells versus thylakoid-enriched membranes, a slight but significant enrichment in alkene was found in the thylakoid fraction (Figure 5B). Most importantly, by using the fatty acid C16:1(3t) as a marker of the thylakoid lipids, it could be estimated that the enrichment in 7-heptadecene corresponds, in fact, to the localization of >90% of this compound to the thylakoid fraction (Figure 5C). Taken together, these results indicated that CrFAP is a chloroplast-targeted soluble protein that has the ability to bind membranes (including the thylakoid fraction) and whose *n*-alka(e)nes products are mostly associated to the thylakoid fraction.

### 7-Heptadecene content varies with cell cycle in *C. reinhardtii*

The lack of FAP and HCs in chloroplasts of *C. reinhardtii* did not result in any obvious differences in the overall organization of cells or chloroplasts as seen by transmission electron microscopy (TEM; Supplemental Figure S4). To try to gather

clues on FAP function, publicly available FAP transcriptomic data were mined. Transcriptomic data (Zones et al., 2015) show that FAP has a similar expression pattern as those genes encoding proteins of the photosynthetic apparatus (Supplemental Figure S5). In order to determine whether FAP product varied with time, we monitored total fatty acids and 7-heptadecene content during a day–night cycles in synchronized *C. reinhardtii* cells. While a constant level of 7-heptadecene representing 0.45% of total fatty acids was found most of the time, a substantial peak (0.7%) was observed before the night (after 12 h of light), and a steep decrease in 7-heptadecene content relative to total fatty acids occurred at the beginning of the night (Figure 6A). This decrease in relative 7-heptadecene content was concomitant with cell division, as evidenced by the drop in total fatty acid content per cell that occurred at that time (Figure 6B). Interestingly, the peak in 7-heptadecene (Figure 6A) was not associated with an increase in total FAP amount between 8 h and 16 h (Figure 6C).

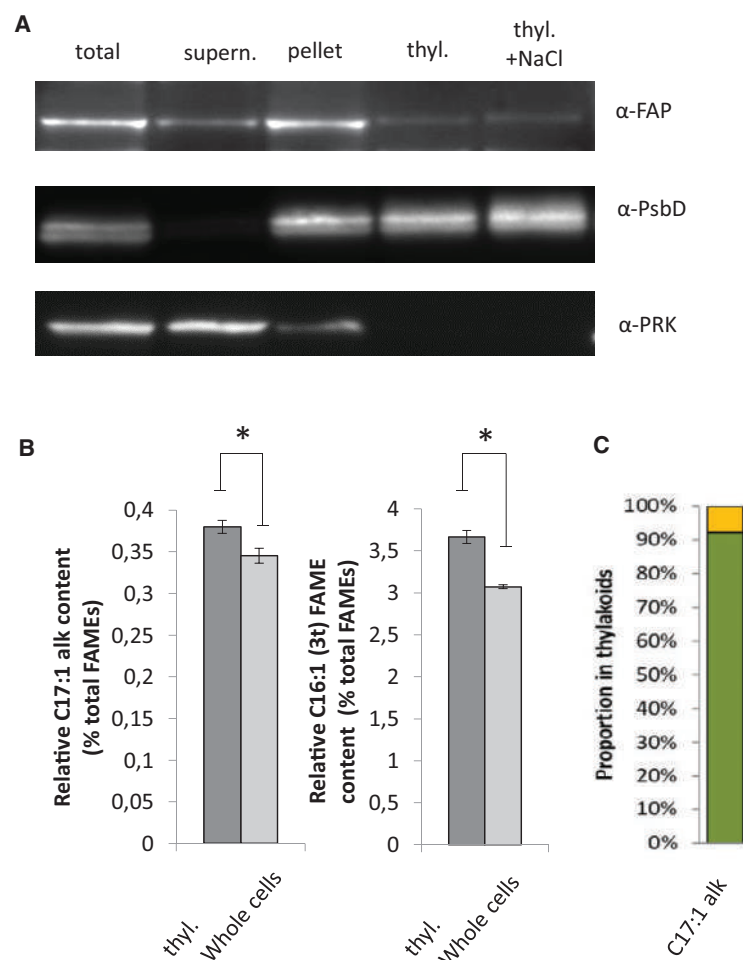


**Figure 4** Overview of the number of FAP homologs and other GMC oxidoreductases identified in eukaryotic algae and other bikonts. In most groups, there is one FAP and no other GMC oxidoreductase. Remarkable species whose number of FAP or other GMC oxidoreductases depart from this rule are listed individually. Hyphens indicate that no protein could be identified by BLAST searches. A widely accepted phylogeny of the Bikonta is used. The dashed line connects the tree of the Bikonta to the rest of the phylogenetic tree of the Eukaryotes. Photosynthetic groups or species are colored. Rounds correspond to endosymbiosis: blue round with one black circle, primary endosymbiosis; green and red rounds with two black circles, secondary endosymbiosis; red round with three black circles is for tertiary endosymbiosis in some (not all) Dinophyta; red or green rounds indicate red or green plastid origin, respectively; n: nucleomorph; x: secondary plastid loss.

### Fatty acid and membrane lipid compositions differ in the *fap* mutants

Since 7-heptadecene content varied during cell cycle and may thus play a role in cell division, growth of the WT and *fap* strains were analyzed. Growth at 25°C was compared in photoautotrophic conditions (mineral medium [MM]) and in mixotrophic conditions (Tris–acetate–phosphate medium [TAP]). No difference between WT and *fap* could be detected, neither in growth rates nor in cell volumes under these conditions (Supplemental Figure S6). In addition, no difference between WT and *fap* strains could be observed when cells were grown under various concentrations of sodium chloride (Supplemental Figure S7). Although the lack

of HCs had no effect on growth, fatty acid profiling showed some differences in C16:1(9), C18:1(9), and C18:3(9,12,15) (Supplemental Figure S8). Synchronized cells also exhibited no growth differences between WT and *fap* strains, but exhibited differences in the dynamics of some fatty acid species (Supplemental Figure S9). Changes in fatty acid profiles prompted us to perform a lipidomic analysis by ultra-performance liquid chromatography coupled with tandem mass spectrometry (UPLC-MS/MS). Interestingly, this analysis revealed that a limited set of lipid molecular species were significantly different between WT and *fap* and that they were all plastidial lipids belonging to the galactolipid classes digalactosyldiacylglycerol (DGDG) and monogalactosyldiacylglycerol



**Figure 5** FAP and its 7-heptadecene product are present in a thylakoid membrane-enriched fraction of *C. reinhardtii*. A, Western blot of protein extracts from whole WT cells before (total) and after centrifugation (supernatant and pellet), and from a thylakoid membrane-enriched fraction without (thyl.) or with (thyl. + NaCl) an additional wash with 0.5-M NaCl. Photosystem II D2 protein (PsbD) and phosphoribulokinase (PRK) are thylakoid and stromal proteins, respectively. Proteins were loaded on a constant chlorophyll basis. B, Relative content in 7-heptadecene and C16:1(3t) fatty acid in whole cells and in the thylakoid fraction (thyl.). The C16:1(3t) fatty acid is almost exclusively present in thylakoids and shown for comparison. Values are mean  $\pm$  SD of  $n = 4$  independent experiments. (\*) denote  $P < 0.05$  in two-sided Student's  $t$  test. C, Estimation of the proportion of total 7-heptadecene present in thylakoids (in green) and elsewhere in the cell (in yellow). Proportion of 7-heptadecene in thylakoids was estimated using C16:1(3t) fatty acid as a thylakoid marker (see "Materials and methods" for calculation).



(MGDG; Figure 7; Supplemental Figure S10). The decrease in the relative content of these galactolipid species appeared to be fully restored by complementation in the case of DGDG but not of MGDG (Figure 7). Taken together, these results indicate that the lack of 7-heptadecene in the *fap* mutant causes a change in thylakoid lipid composition, which is evidenced by the decrease in the relative content of at least three galactolipid species belonging to the DGDG class.

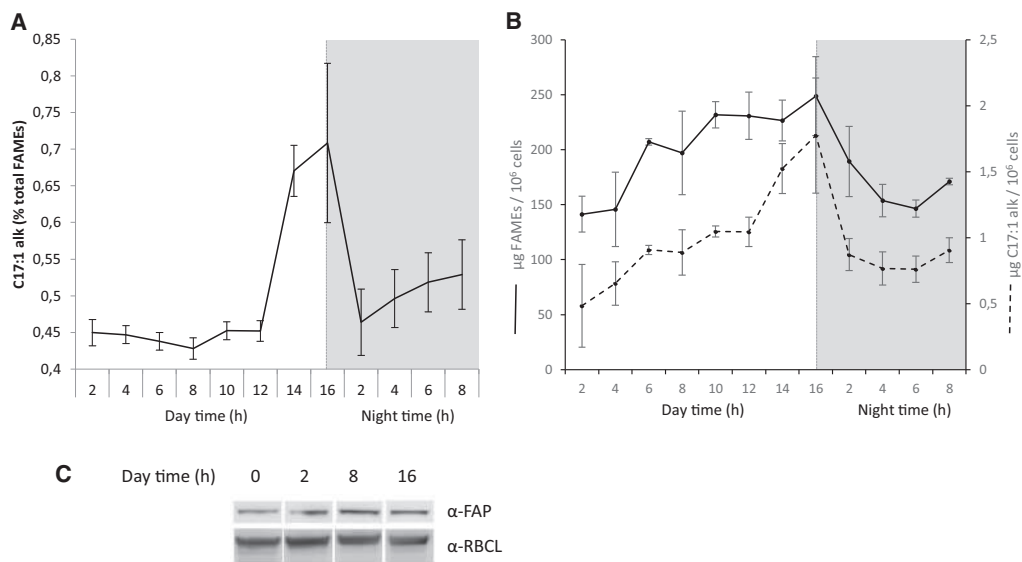
### FAP is not strongly associated to photosynthetic complexes, and lack of HCs has no effect on their organization

In cyanobacteria, a role of HCs in photosynthesis has been suggested (Berla et al., 2015) but is controversial (Lea-Smith et al., 2016). In *C. reinhardtii*, there was no difference in the 77K chlorophyll fluorescence spectrum between WT, complementation and *fap* mutant strains, which indicated no major changes in antenna distribution around photosystems (Supplemental Figure S11A). No difference in photosynthetic efficiency could be detected among WT, complementation, and *fap* strains grown under standard laboratory conditions (Supplemental Figure S11B). Membrane inlet mass spectrometry (MIMS) experiments conducted to quantify  $O_2$  exchange showed no difference in respiration or photosynthesis rates between the two genotypes (Supplemental Figure S11C). Native electrophoresis of proteins from purified thylakoids and FAP immunodetection revealed that FAP could only be detected at an apparent molecular size of the

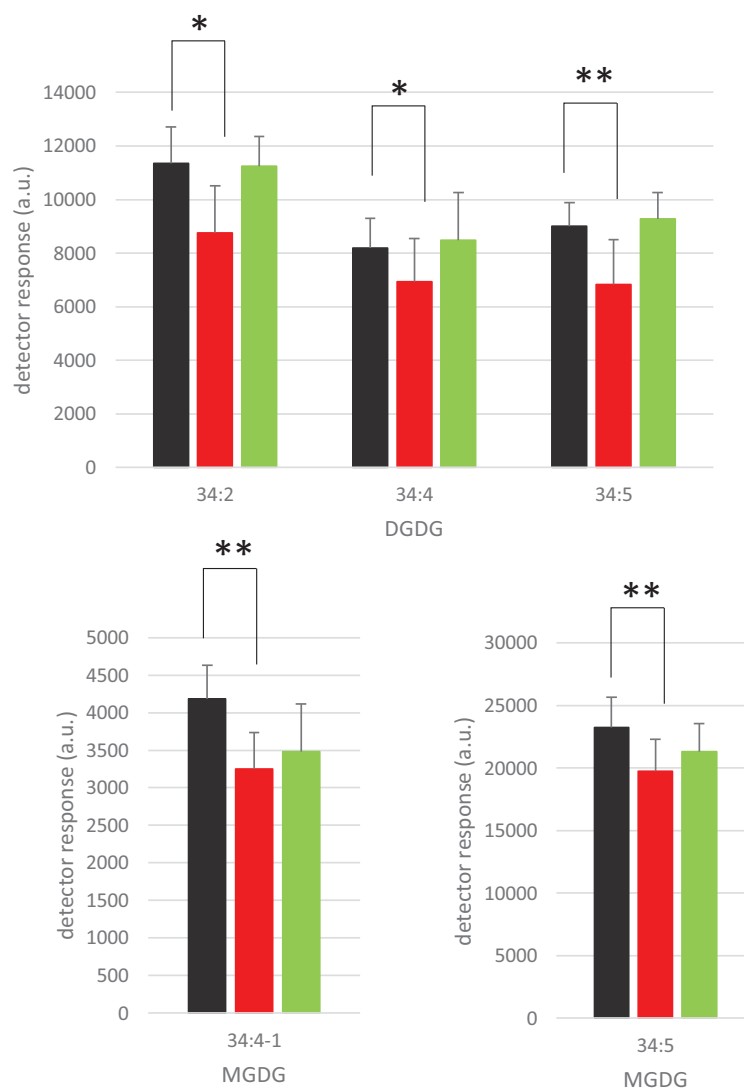
monomeric FAP (Figure 8), indicating no strong association with proteins of photosynthetic complexes. Besides, no difference in organization of photosynthetic complexes between WT and *fap* could be seen on the native protein electrophoresis.

### Photosynthesis is affected under light and cold stress in the *fap* mutants

Lack of HCs in the *fap* strain did not cause changes in the photosynthesis activity under standard growth conditions. However, since significant modifications in the composition of membrane lipids could be detected, we explored harsher conditions to further challenge photosynthetic membranes. We chose to investigate chilling temperatures because cold is well-known to affect both membrane physical properties and photosynthesis (Los et al., 2013). Using a multicultivator in turbidostat mode, we first stabilized cultures at 25°C under medium light ( $200 \mu\text{mol photons m}^{-2} \text{s}^{-1}$ ), finding the apparent electron transfer rate (ETR) showed no difference under this condition (Figure 9A). When cooling down the culture to 15°C and after 3 d of acclimation, both ETR and 77 K chlorophyll fluorescence spectra still showed no differences (Figure 9B; Supplemental Figure S11A). However, after 1 d at a lower light intensity ( $50 \mu\text{mol photons m}^{-2} \text{s}^{-1}$ , 15°C), while the maximal PSII yield was equal for all the strains (Figure 9C), the ETR was lower for the mutant when measured at high light intensities (Figure 9D). Interestingly, longer acclimation to this condition (3 d) led to the disappearance of this phenotype. This transient phenotype of the



**Figure 6** Variation of 7-heptadecene compared to total fatty acids during *C. reinhardtii* cell cycle. A, 7-Heptadecene content of cells expressed as a percent of total FAMES. Values are mean  $\pm$  SD ( $n = 3$  biological replicates). B, Total fatty acid and 7-heptadecene content per million cells during cell cycle. Total fatty acids were analyzed as FAMES by GC-MS. A and B, Data are mean  $\pm$  SD of  $n = 3$  independent cultures. The orange dashed line indicates the end of day time and the beginning of night time (gray shading). C, Immunodetection of FAP during day time. The large Rubisco subunit (RBCL) was used as a loading control.



**Figure 7** Identification of lipid molecular species significantly different between *C. reinhardtii* WT and *fap* strains. Relative abundance of glycerolipids was measured by LC–MS/MS analysis of total lipid extracts of whole cells. Only glycerolipid molecular species showing significant differences between WT and *fap* strains are shown here (see Supplemental Figure S12 for complete results). WT, *fap*, and Cp: wild-type, FAP knockout and complementation strains, respectively. Lipid extracts from the three strains were loaded on a constant total fatty acid basis. Data are mean  $\pm$  SD of  $n = 9$  independent cultures (using three different strains for each of the three genotypes). MGDG34:4-1 is one of the two species of MGDG 34:1 species (which differ by C18:3 fatty acid isomers). Asterisks indicate significant differences according to a Mann–Whitney *U*-test at \* $P < 0.05$  or \*\* $P < 0.01$ . Cells were grown in TAP medium, under  $80 \mu\text{mol photons m}^{-2} \text{s}^{-1}$  in Erlenmeyers.

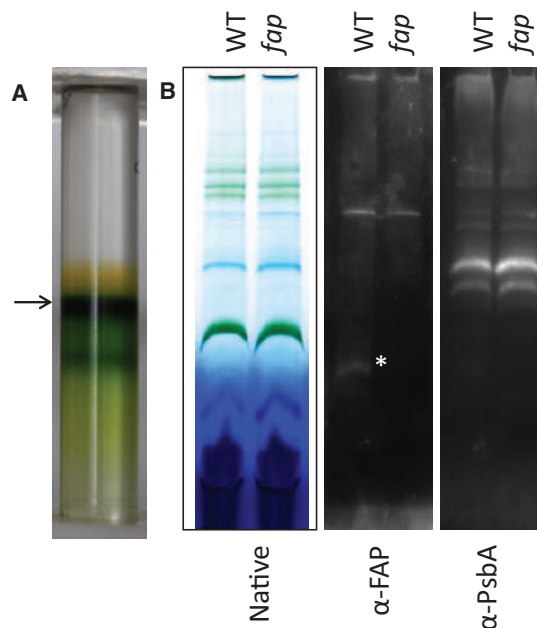
*fap* KO was clearly linked to low temperature because shifting cultures at 25°C from medium light to low light did not result in any difference in maximal PSII yield or ETR among the strains (Figure 9E).

In order to provide support for a possible link between HCs and cold acclimation, 7-heptadecene content was quantified under various growth temperatures. Relative HC content in WT cells clearly increased under cold conditions (Figure 10). Enrichment of WT membranes in HC was due to a slower decrease in HC synthesis compared to fatty acid

synthesis (Figure 10B). As expected, an increase in the relative content of polyunsaturated species occurred upon cold treatment in both WT and *fap* strains (Supplemental Figure S12), but no difference in the dynamics of fatty acid remodeling was observed between the two strains.

## Discussion

Here, we report the isolation and characterization of an insertional *C. reinhardtii* mutant deficient in FAP, and we perform phylogenetic and functional analyses of algal



**Figure 8** Thylakoid purification and immunoblot analysis. A, Purification of *C. reinhardtii* thylakoids using a sucrose density gradient. The thylakoid fraction collected is indicated by an arrow. B, Blue native polyacrylamide gel of solubilized proteins (0.5% digitonine, 0.5%  $\alpha$ -DM) and corresponding immunodetection. WT: wild-type strain, *fap*: FAP knockout strain. PsbA: Photosystem II D1 protein. Asterisk indicates FAP band.

homologs. We show that FAP and the vast majority of its 7-heptadecene product are associated to thylakoid membranes. It is also shown that a gene encoding for a putative FAP is present in most algal lineages and indeed encodes a functional FAP in some species of red algae, and of secondary algae, including some macroalgae. By studying a FAP knockout *C. reinhardtii* mutant, we provide evidence that lack of HCs is associated correlated with small changes in galactolipid composition, but has no impact on photosynthesis and growth in *C. reinhardtii* under standard culture conditions. However, in the absence of HCs generated by FAP, the photosynthetic activity is transiently affected during cold acclimation when light intensity varies. The possible significance of these results for algal physiology, as well as FAP function and evolution, is discussed below.

#### CrFAP localization and regulation in algal cells

Based on the characterization of a *fap* mutant, we first show that CrFAP is responsible for the synthesis of all fatty acid-derived HCs found in *C. reinhardtii* cells (Figure 1). Our result clearly demonstrates that the FAP activity measured in vitro for CrFAP (Sorigué et al., 2017) is not a promiscuous secondary activity and indeed corresponds to a genuine biological activity, namely the light-driven synthesis of 7-heptadecene from cis-vaccenic acid (Figure 11). Also, the results from the *fap* knockout line demonstrate that no other

enzyme is able to synthesize 7-heptadecene in *C. reinhardtii*. Furthermore, the HC production was found to correspond to the quantity of CrFAP present in complementation lines. Thus, FAP is a limiting factor for 7-heptadecene production in vivo, but the lipase(s) that may act upstream of FAP to generate the free cis-vaccenic acid does not limit the pathway. This conclusion is consistent with a previous observation that 7-heptadecene is increased eight-fold in *C. reinhardtii* when CvFAP is overexpressed (Yunus et al., 2018).

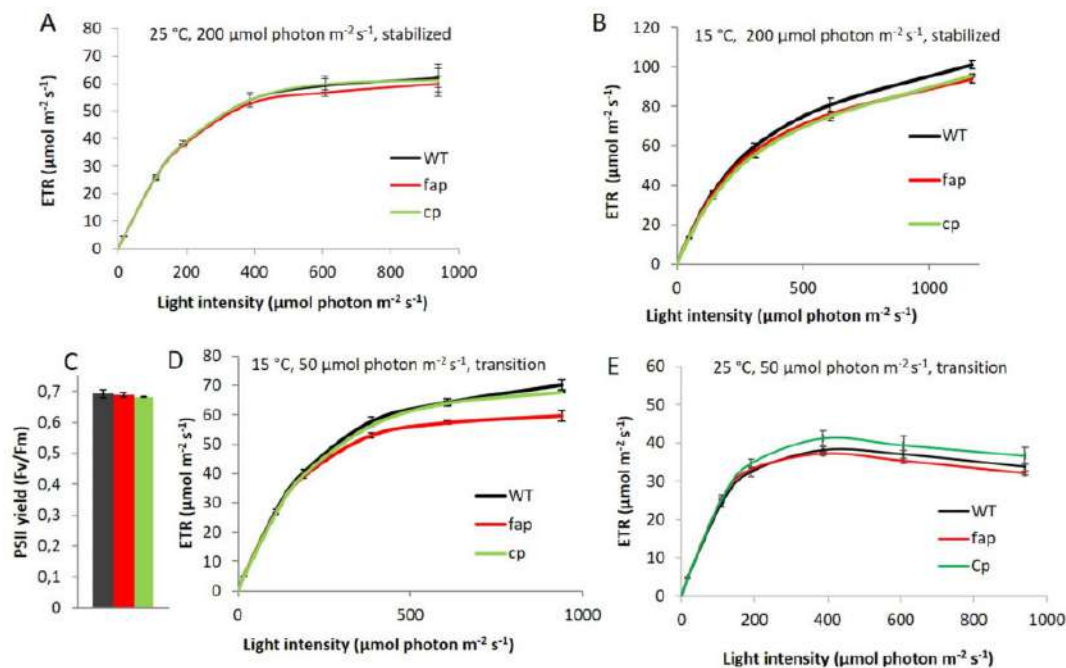
CrFAP is a soluble protein (Sorigué et al., 2017). Using sub-cellular fractionation of *C. reinhardtii* cells and anti-CrFAP antibodies, we show here that CrFAP is predominantly found in the total membrane fraction (Figure 5A). Based on the predicted plastid localization of CrFAP (Supplemental Figure S3) and the fact that the thylakoid fraction harbors >90% of the 7-heptadecene product (Figure 5C), we propose that in vivo CrFAP is a soluble chloroplastic protein that can bind to and unbind from thylakoid membranes. The capacity of CrFAP to bind to thylakoids is supported by the observation that after the thylakoid purification procedure, a fraction of CrFAP is still tightly bound to this fraction (Figure 5A). Plastidial localization may be a general feature of FAPs because presence of plastid transit peptides in FAPs seems to be a general rule in green and red algae (for primary plastids) and is also predicted for some secondary endosymbiosis-derived algae (Supplemental Figure S3).

Concerning the regulation of FAP activity in membranes, it is clear that the peak in 7-heptadecene during the cell cycle (Figure 6A) is not due to an increase in the total amount of FAP protein (Figure 6C), and another mode of regulation of FAP activity must exist. Availability of free fatty acid substrates provided by a lipase acting upstream of FAP is a possibility (Figure 11). The peak in CrFAP gene expression during the day (Supplemental Figure S5) may correspond to the necessity of a high turnover of the CrFAP protein due to photodamage to the protein. This would be consistent with the idea that *C. variabilis* FAP has a radical-based mechanism (Sorigué et al., 2017), which may be the cause of photoinactivation of the protein (Lakavath et al., 2020) and degradation under excessive light (Moulin et al., 2019).

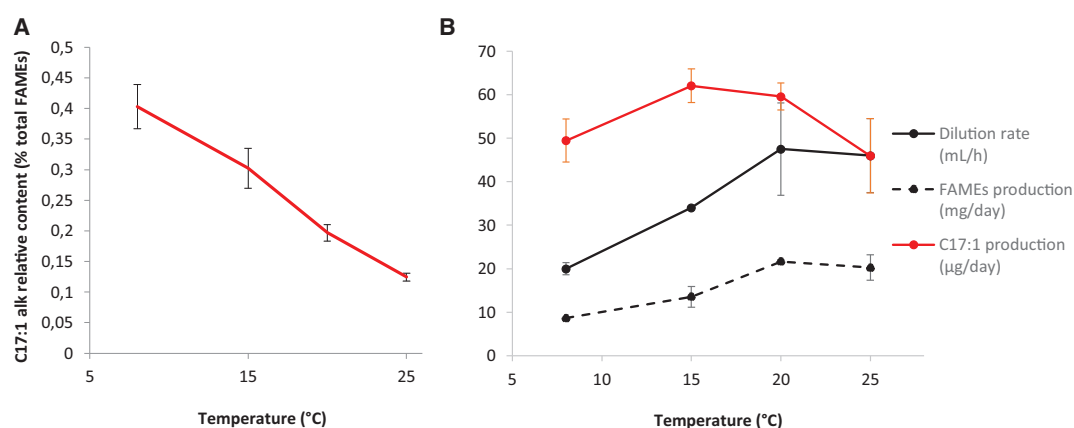
#### Role of CrFAP products in membranes

In a work reporting cyanobacterial mutants devoid of fatty acid-derived HCs, it has been suggested that HCs are located in membranes and may play a role in cell division (Lea-Smith et al., 2016). In the proposed cyanobacterial model, integration of HCs into the lipid bilayer would be responsible for membrane flexibility and curvature. HCs may play a similar role in thylakoid membranes of green algae. In light of this connection, it is interesting to note that in *C. reinhardtii*, HC production follows lipid production during cell growth except just before mitosis (Figure 6, A and B). The ratio of HCs to fatty acid methyl esters (FAMES) decreased at the beginning of night, when cells are dividing, indicating that the extra amount of HCs synthesized before the night must be somehow lost or metabolized during cell division.





**Figure 9** Photosynthetic acclimation to cold conditions in *C. reinhardtii* WT and *fap* strains. Apparent ETR at various light intensities for cells grown in photoautotrophic conditions at 25°C (A) and 15°C after 3 d acclimation (B). PSII yield (C) and ETR (D) at 15°C after exposure to lower light for one day (50  $\mu\text{mol photon m}^{-2} \text{s}^{-1}$ ). E, ETR at 25°C after transition to lower light for one day (50  $\mu\text{mol photon m}^{-2} \text{s}^{-1}$ ). WT, *fap*, and Cp: wild-type, FAP knockout, and complementation strains, respectively. Data are mean  $\pm$  SD of  $n = 3$  independent cultures (corresponding to three different strains for each genotype). Note: Figure Replacement Requested.

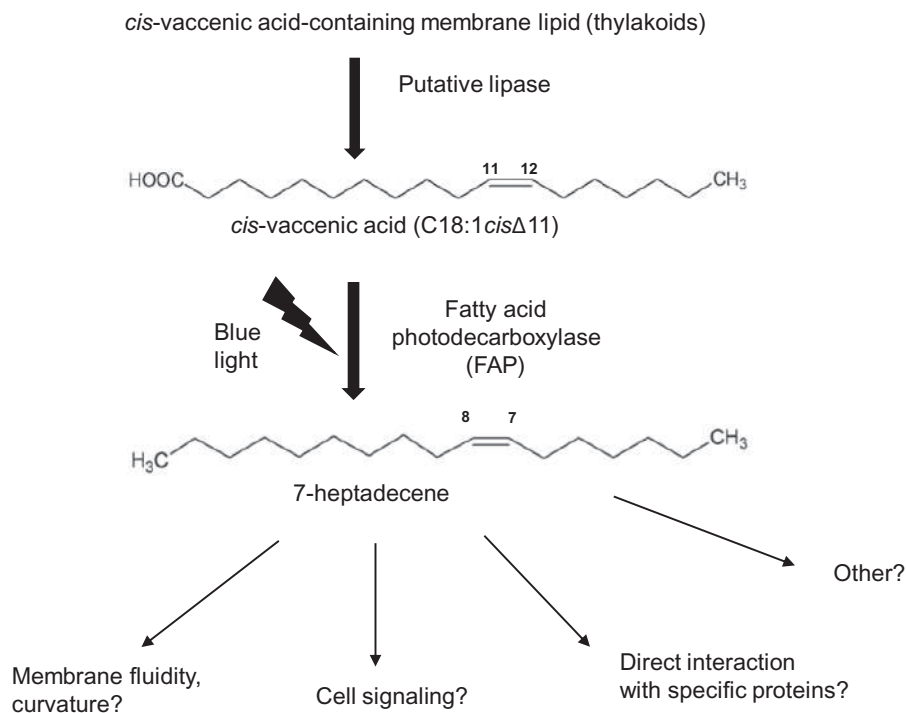


**Figure 10** Hydrocarbon production in *C. reinhardtii* WT cells cultivated at low temperatures. A, HC content in cells as percentage of total FAMES. B, Dilution rate of culture medium and production of FAMES and HCs. For A and B, Cells were grown in TAP medium in photobioreactors in turbidostat mode under 50  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$ . HC and fatty acid content were analyzed by GC-MS after transmethylation. Data are mean  $\pm$  SD of  $n = 3$  independent cultures.

Simple mechanisms that could explain HC loss during cell division involve enrichment in HCs at breaking points of plastidial membranes before cell division, exclusion from these membranes during division, and/or loss to the gas

phase of the culture due to HC volatility (Supplemental Figure 5B).

HCs might thus impact local flexibility of algal plastidial membranes and participate in lipid membrane remodeling



**Figure 11** Proposed pathway for hydrocarbon formation from fatty acids in *C. reinhardtii* and putative roles. The only fatty acid-derived HC in *C. reinhardtii* (7-heptadecene) is generated from *cis*-vaccenic acid by FAP only when cells are exposed to blue light. The fatty acid precursor must be released from a thylakoid lipid by an unknown lipase. The 7-heptadecene FAP product may play several roles in thylakoid membranes depending on temperature and light conditions.

during cell division. However, under standard culture conditions, the presence of HCs is apparently not critical for chloroplast structure (Supplemental Figure 4), cell size and cell division rate (Supplemental Figure 6). A role of cyanobacterial HCs in tolerance to salt stress has also been suggested (Yamamori et al., 2018). In *C. reinhardtii*, contrary to what has been shown in cyanobacteria, no difference could be detected in growth under increasing salt concentrations (Supplemental Figure 7). One could thus hypothesize that even if HCs are produced in chloroplasts and accumulate in thylakoids, their function might be different from that in cyanobacteria. It is also possible that laboratory culture conditions used for *C. reinhardtii* (this study) are far from natural growth conditions where HCs may be necessary. Alternatively, a compensation mechanism for HC loss may operate differently in *C. reinhardtii* and in cyanobacteria. In *C. reinhardtii*, part of this mechanism may involve changes in membrane lipid composition. Interestingly, lipidomic analysis under standard growth conditions unraveled specific changes in DGDG molecular species (Figure 7) but no other significant differences in other class of lipids (Supplemental Figure S8). Taken together, these results suggest that in *C. reinhardtii*, HCs play no crucial role in cell division and growth under standard conditions. Cells may adapt to a lack of HCs by some changes in the composition of membranes,

which could specifically involve some DGDG galactolipids. Alternatively, or in addition to this proposed effect on properties of the membrane lipid phase, it cannot be ruled out that 7-heptadecene may act locally to disrupt or enhance some specific protein–protein interactions, or may play a yet to be defined role, such as acting as a signaling molecule or its precursor (Figure 11).

The fact that FAP gene expression follows that of photosynthesis genes in day–night cycles, the likely localization of FAP in plastids of green and red algae (as well as in some secondary algae), and the localization of part of FAP and almost all its alkene product in *C. reinhardtii* thylakoids, point toward a role of FAPs in the photosynthetic function of algal cells. This idea is strongly reinforced by the conservation of the FAP-encoding gene in many eukaryotic algae but not in nonphotosynthetic protists (Figures 3 and 4) and in *Polytomella*, an algae that has kept some of its plastidial function but lost photosynthesis (Smith and Lee, 2014). As standard culture conditions did not reveal any photosynthesis phenotype in *C. reinhardtii* *fap* mutant (Supplemental Figure S11), more challenging conditions involving colder temperatures and variations in light intensity were tested. These experiments have revealed a difference between WT and *fap* mutant in the photosynthesis activity during cold acclimation after light intensity varied from medium to low

(Figure 9). This difference could be highlighted under the high-light ETR measurement conditions. Interestingly, colder temperatures were associated with membrane enrichment in HCs (Figure 10) and fatty acid profiles followed similar variations with temperature in both WT and *fap* strains (Supplemental Figure S12). Taken together, these observations indicate that adaptations in membrane lipid composition compensate partly for the loss of HCs in standard growth conditions, but not in harsher conditions, such as light intensity variations under cold temperatures.

At temperatures lower than 20°C, it is clear that cell division and FAME production slow down (Figure 10B). Interestingly, HC production does not follow the same pattern, and increases with decreasing temperatures (except at 8°C). Enrichment of membranes in HCs with decreasing temperatures is consistent with our proposal that HCs are lost during cell division (Supplemental Figure S5): lower temperature causes reduced division and would hence result in lower HC losses. Whether the sustained production of HC at low temperature is an adaptation mechanism to low temperature or not, remains to be elucidated.

### Conservation of FAP in algae

According to molecular phylogeny (Figures 3 and 4), FAP proteins appear to be specific to algae and highly conserved in many algae species. A noticeable exception is the Mamiellophyceae class of the green algae. Algae are a common denomination that covers photosynthetic eukaryotes which mainly live in aquatic environments. This polyphyletic group includes organisms derived from a first endosymbiosis, as well as organisms derived from a secondary, or even tertiary, endosymbiosis. However, a functional FAP can be found in Chlorophyta (green algae), Rhodophyta (*Chondrus* and *Galdieria*), and Stramenopiles (in the Phaeophyceae *Ectocarpus* and the Eustigmatophyceae *Nannochloropsis*), as demonstrated by heterologous expression in *E. coli* of the corresponding identified FAPs (Figure 2). FAP activity was therefore conserved during secondary endosymbiotic event(s) that gave rise to the red algae lineage. Moreover, FAP activity is not specific to the unicellular state, as FAPs were also functional in the pluricellular algae (macroalgae) *E. siliculosus* and *C. crispus*. Considering homology of sequences, FAP function is thus expected to be present in most algal phyla, including Haptophyta and Dinophyta. Importantly, some amino acid residues that are likely to be involved in fatty acid substrate stabilization or photocatalysis, such as CvFAP Arg451 or Cys432 (Sorigué et al., 2017; Heyes et al., 2020), are strictly conserved in the 198 putative FAPs (Supplemental Figure S2). This observation reinforces the idea that all the putative FAPs identified in this work have the ability to photo-produce HCs from fatty acids.

FAP neofunctionalization from GMC oxidoreductases may have occurred early during evolution of algae, almost concomitantly with the very first endosymbiosis shared by green and red algae. No GMC could be found in

Glaucocystophyta (Figure 4), which may indicate that this neofunctionalization event has occurred after separation of this group from red and green algae. However, it should be noted that so far only one complete Glaucocystophyta genome is available. Absence of FAP in charophytes (Supplemental Figure 1) indicates early loss of FAP function in Streptophyta. Phylogenetic analysis indicates that FAPs from secondary endosymbiosis lineages are more closely related to core Chlorophyta than to Rhodophyta. FAP could thus be one of the genes that was inherited from green algae by horizontal gene transfer (Moustafa et al., 2009).

Concerning the conservation of FAP activity, it should be noted that some of the FAPs selected for heterologous expression had fatty acid preference profiles different from that of CvFAP (Figure 2B). For example, while CvFAP converted *E. coli* fatty acids into HCs with a slight preference for C18 over C16 fatty acids, *Chondrus* and *Galdieria* FAPs showed higher preference for C16 over C18, and saturated over unsaturated fatty acids. In contrast, *Ectocarpus* FAP had a strong preference for C16:0 over C16:1 fatty acid, and *N. gaditana* FAP preferred unsaturated over saturated chain for both C16 and C18 fatty acids. This indicates that the algal biodiversity is likely to contain FAPs with fatty acid specificities different from the FAPs of *C. variabilis* and *C. reinhardtii*. FAPs with different properties may be useful for a biobased production of specific HCs, for instance shorter chain semi-volatile HCs for liquid fuels (Moulin et al., 2019).

In conclusion, the results presented here show that FAP activity is conserved beyond green microalgae, and identify a big reservoir of FAPs that may be useful for biotechnological applications. It also provides some important clues for future studies aiming at unravelling the exact role of the FAP photoenzyme in eukaryotic algae.

## Materials and methods

### Strains and culture conditions

The *fap* mutant and its corresponding WT strain of *C. reinhardtii* were ordered from the CLiP library (Li et al., 2016). Upon receipt, strains were plated on TAP medium and streaked to allow formation of single colonies. For each strain, after 1-week growth in the dark, three single-clone-derived colonies were randomly chosen for characterization. WT strains are CC-4533 cw15 mt<sup>−</sup> for mating type minus and CC-5155 cw15 mt<sup>+</sup> (Jonikas CMJ030 F5 backcross strain) [isolate E8] for mating type plus. Mutant LMJ.RY0402.226794 was used in this study, which is predicted to harbor a first insertional cassette in the coding sequence of Cre12.g514200 which encodes FAP. A second insertion in the line LMJ.RY0402.226794 was predicted in Cre14.g628702. To remove this side mutation, we backcrossed the mutant strain to CC-5155. Analysis of one full tetrad showed two progeny strains resistant to paromomycin and which were mutated in *FAP* gene. The region of Cre14.g628702 was amplified by polymerase chain reaction (PCR) and sequenced for the four progeny strains of the

tetrad. No insertion was actually found, therefore a potential insertion at this locus was ruled out. Work on mutant strains was conducted on one isolated parental strain with the mutation from LMJ.RY0402.226794, and on two mutants from the backcross with CC5155 (full tetrad). These three strains are thereafter named *fap-1*, *fap-2*, and *fap-3*. WT strains were parental strain CC5155 (WT-1) and single colony-derived lines of background strain CC4533 (WT-2, WT-3). The average of the three WT strains and the average of the three knockout strains are labeled as WT and *fap*, respectively. For liquid culture experiments, cells were grown in 24 deep well plates of 25 mL culture, under 100  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$  with constant shaking at 25°C. Cells were grown in TAP or MM (Harris, 1989) for mixotrophic and autotrophic conditions, respectively. Cell growth was followed using a cell counter Multisizer (Beckman Coulter). For the day–night cycle experiment, cells were cultivated autotrophically in 1L-photobioreactors in turbidostat mode ( $\text{OD}_{880\text{nm}}$  at 0.4) under 16 h of light (40  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$ ) and 8 h of dark at 25°C. For photosynthesis analysis, cells were grown autotrophically in 80 mL photobioreactors (multicultivator, Photon Systems Instruments) in turbidostat mode ( $\text{OD}_{680\text{nm}}$  at 0.8). Conditions were 25°C, with medium light intensity (200  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$ ) or low light intensity (50  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$ ); or 15°C, with medium or low light intensity. All cultures were performed under ambient air.

### Complementation of the *fap* mutants

The construct for complementation of the knockout strain for *FAP* gene was performed using pSL-Hyg vector containing an *AphVII* cassette conferring hygromycin resistance. This vector allowing nuclear transformation was kindly provided by Pr. Steven Ball (University of Lille, France). A WT copy of the *FAP* gene was obtained by PCR of WT genomic DNA. It was cloned into TOPO-XL vector. pSL-Hyg vector and *FAP* gene were digested with *EcoRV* and *SpeI* and ligated. Then, the vector was linearized with *PvuI* and was electroporated into the *fap* strains. Level of complementation was verified by immunoblot (to assess quantity of protein) and by transmethylation of whole cells (to assess quantity of HCs). The two *fap* mutants from a full tetrad were complemented. Three complemented strains with HC levels similar to WT were kept for physiological studies (e.g. Cp-4 shown in Figure 1).

### SDS–PAGE and immunodetection

Cells (10–15 mL) were harvested by centrifugation at 3,000g for 2 min. Pellets were then frozen in liquid nitrogen and stored at -80°C until use. Pellets were resuspended in 400 mL 1% (w/v) sodium dodecyl sulfate (SDS) and then 1.6 mL acetone precooled to -20°C was added. After overnight incubation at -20°C, samples were centrifuged (20,800 g, 10 min, 4°C). Supernatant was removed and used for chlorophyll quantification using a UVmc spectrophotometer (SAFAS). Pellets were resuspended to 1-mg chlorophyll  $\text{mL}^{-1}$  in Lithium dodecyl sulfate buffer in the presence of

NuPAGE reducing agent (ThermoFischer), and loaded on 10% (w/v) polyacrylamide gel electrophoresis (PAGE) Bis–Tris SDS gel. To load equal protein amounts for immunoblot analysis, protein contents were estimated by Coomassie Brilliant Blue staining of the gel using an Odyssey IR Imager (LICOR). After gel electrophoresis, proteins were transferred to nitrocellulose membranes for 75 min at 10 V using a semi-dry set up. Membranes were blocked in Tris-buffered saline Tween 20 buffer with milk 5% (w/v) overnight at 4°C and then incubated at room temperature in the presence of the following antibodies: anti-Cyt f, anti-AtpB, anti-PsaD, anti-PsbA, anti-LHCSR3 (Agrisera), or anti-FAP (see below). After 2-h incubation, primary antibody was removed by rinsing three times in TBST, and a peroxidase-coupled secondary antibody was added for at least 1 h. Luminescence was detected with a Gbox imaging system (Syngene).

### Production of anti-CrFAP antibodies

A codon-optimized synthetic gene encoding *C. reinhardtii* FAP (Sorigué et al., 2017) was cloned into the pLIC7 expression vector, allowing the production of a recombinant FAP fused to TEV-cleavable His-tagged *E. coli* thioredoxin. Production was performed in the *E. coli* BL21 Star (DE3) strain initially grown at 37°C in TB medium. Induction was initiated at an  $\text{OD}_{600\text{nm}}$  of 0.8 by adding 0.5-mM isopropyl  $\beta$ -D-thiogalactoside (IPTG), and cultures were then grown at 20°C. Following overnight incubation, cells were centrifuged and protein was purified as described previously (Sorigué et al., 2017). Purity of the purified protein was controlled on SDS–PAGE, and it was brought to a final concentration of 2  $\text{mg mL}^{-1}$  using an Amicon-Ultra device (Millipore). Polyclonal antibodies against FAP were raised in rabbits (ProteoGenix, Schiltigheim, France).

### Analysis of HCs and fatty acids

For quantification of HCs, approximately one hundred million cells were pelleted by centrifugation in glass tubes. Transmethylation was conducted by adding 2 mL of methanol containing 5% (v/v) sulfuric acid to the cell pellet. Internal standards (10  $\mu\text{g}$  of hexadecane and 20  $\mu\text{g}$  of triheptadecanoylglycerol) were added for quantification. Reaction was carried out for 90 min at 85°C in sealed glass tubes. After cooling down, 1 mL of 0.9% (w/v) NaCl and 500  $\mu\text{L}$  of hexane were added to the samples to allow phase separation and recovery of FAMES and HCs in the hexane phase. Samples were mixed and then centrifuged to allow phase separation. Two microliter of the hexane phase was injected in the apparatus of gas chromatography coupled to mass spectrometry and flame ionization detection (GC–MS/FID). Analyses were carried out on an Agilent 7890A gas chromatograph coupled to an Agilent 5975C mass spectrometer (simple quadrupole). A Zebron 7HG-G007-11 (Phenomenex) polar capillary column (length 30 m, internal diameter 0.25 mm, and film thickness 0.25  $\mu\text{m}$ ) was used. Helium carrier gas was at 1  $\text{mL min}^{-1}$ . Oven temperature was programmed with an initial 2-min hold time at 35°C, a first ramp from 35°C to 150°C at 15°C  $\text{min}^{-1}$ , followed by a

1-min hold time at 170°C then a second ramp from 170°C to 240°C at 5°C min<sup>-1</sup> and a final 2-min hold time at 240°C. The MS was run in full scan over 40 to 350 amu (electron impact ionization at 70 eV), and peaks of FAMES and HCs were quantified based on the FID signal using the internal standards C17:0 FAME and hexadecane, respectively.

### Chlorophyll fluorescence measurements and MIMS analysis

Chlorophyll fluorescence measurements were performed using a pulse amplitude-modulated fluorimeter (Dual-PAM 100) upon 15-min dark-adaptation under continuous stirring. Detection pulses (10 mmol photons m<sup>-2</sup> s<sup>-1</sup> blue light) were supplied at a 100-Hz frequency. Basal fluorescence ( $F_0$ ) was measured in the dark prior to the first saturating flash. Red saturating flashes (6,000 mmol photons m<sup>-2</sup> s<sup>-1</sup>, 600 ms) were delivered to measure  $F_m$  (in the dark) and  $F_m'$  (in the light). PSII maximum yields were calculated as  $(F_m - F_0)/F_m'$ , and PSII yield for each light intensity was measured with a saturating flash after 2–3 min of illumination and was calculated from  $(F_m' - F)/F_m'$ . The apparent ETR was calculated as the product of light intensity and PSII yield. MIMS was used to measure gas exchange as described previously (Burlacot et al., 2018).

### Transmission electron microscopy

Cells were grown photoautotrophically in photobioreactors under 40 μmol photons m<sup>-2</sup> s<sup>-1</sup> in turbidostat (OD<sub>880nm</sub> at 0.4). The algal cells were collected by centrifugation (1,500g, 1 min) and were immediately fixed with 2.5% (v/v) glutaraldehyde in 0.1 M, pH 7.4 sodium cacodylate buffer for 2 d at 4°C. They were then washed by resuspending 5 min in the same buffer for three times. Samples were post-osmicated with 1% (w/v) osmium tetroxide in cacodylate buffer for 1 h, dehydrated through a graded ethanol series, and finally embedded in monomeric resin Epon 812. All chemicals used for histological preparation were purchased from Electron Microscopy Sciences (Hatfield, USA). Ultra-thin sections for TEM (90 nm) were obtained by an ultramicrotome UCT (Leica Microsystems GmbH, Wetzlar, Germany) and mounted on copper grids and examined in a Tecnai G2 Biotwin Electron Microscope (ThermoFisher Scientific FEI, Eindhoven, the Netherlands) using an accelerating voltage of 100 kV and equipped with a CCD camera Megaview III (Olympus Soft imaging Solutions GmbH, Münster, Germany).

### Isolation of thylakoids and native PAGE

Thylakoids were isolated according to a protocol described previously (Chua and Bennoun, 1975). All steps were performed on ice or at 4°C with as little light as possible. Briefly, cells were pelleted and resuspended in 8-mL 25-mM HEPES, 5-mM MgCl<sub>2</sub>, 0.3-M sucrose, pH 7.5 with a protease inhibitor cocktail for plant cell and tissue extracts (Sigma P9599). Cells were disrupted with French press at a pressure of 6,000 psi. Total membranes were collected by

centrifugation (1,000g, 10 min) and washed first in 5-mM HEPES, 10-mM EDTA, 0.3-M sucrose, pH 7.5 and then in 5-mM HEPES, 10-mM EDTA, 1.8-M sucrose, pH 7.5. Sucrose gradient was 0.5-M sucrose (5 mL), 1.3-M sucrose (2 mL), and 1.8-M sucrose, initially containing thylakoids (5 mL). After ultracentrifugation (274,000g, 1h), thylakoids were collected at the interface between 0.5 and 1.3 M sucrose. They were washed with 5-mM HEPES, 10-mM EDTA, pH 7.5 (with or without 0.5-M NaCl), and resuspended at 1-mg mL<sup>-1</sup> chlorophyll for subsequent SDS-PAGE analysis. For non-denaturing conditions, thylakoids were resuspended in NativePAGE sample buffer (Life technologies) at 1-mg mL<sup>-1</sup> chlorophyll, then thylakoids were solubilized for 30 min on ice in the same volume of 1% (w/v) n-dodecyl-α-D-maltoside, 1% (w/v) digitonine (final concentrations: 0.5-mg mL<sup>-1</sup> chlorophyll, 0.5% (w/v) n-dodecyl-α-D-maltoside, and 0.5% (w/v) digitonine). For each sample, 20 μL were then loaded with 2 μL of G-250 sample additive (Life technologies) on 4–16% (w/v) NativePAGE gels (Life technologies). Cathode running buffer (Life technologies) was supplemented with 0.02% (w/v) G-250 for two-thirds of the migration, and with 0.002% (w/v) G-250 for the remaining third. Annotation of observed bands was done according to a previous publication (Pagliano et al., 2012). For immunoblot analysis, native gel was incubated in Tris Glycine SDS buffer, 10% (v/v) ethanol for 15 min and transferred onto polyvinylidene fluoride membrane using XCell II Blot module (25v, 1 h). Immunodetection was performed as described above.

Based on C16:1(3t) FAME, we determined a factor of enrichment expected for a compound that would be located exclusively within thylakoids [ratio EF=C16:1 (3t) FAME<sub>whole cells</sub>/C16:1 (3t) FAME<sub>thylakoids</sub>]. Considering the amount of C17:1 alkene found in thylakoids, we calculated the expected content in whole cells, which equals C17:1 alk<sub>thylakoids</sub>\*ratio EF. This calculated value for thylakoids was divided by the value for whole cells that had been determined experimentally, which gives the proportion of C17:1 alkene that is present in thylakoids.

### Lipidomic analysis by UPLC-MS/MS

Lipid molecular species analysis was done by UPLC-MS/MS. Lipids were first extracted with a modified hot isopropanol method. Briefly, *C. reinhardtii* cells were harvested by centrifugation at 1,000g for 5 min in glass tubes. Pellets were immediately resuspended in 1 mL of hot isopropanol (85°C) containing 0.01% (w/v) butylated hydroxytoluene. Sealed tubes were heated at 85°C for 10 min to inactivate lipases. Internal standards were added. Lipids were then extracted in 3-mL methyl tert-butyl ether (MTBE) with a phase separation with 1 mL of water. Organic phase was collected and aqueous phase was washed with an additional milliliter of MTBE. Organic phases were evaporated under a gentle nitrogen stream and resuspended in 500 μL of a mixture of acetonitrile/isopropanol/ammonium acetate 10 mM (65:30:5, v/v/v). Lipid molecular species were analyzed on an ultimate RS 3000 UPLC system (ThermoFisher, Waltham,



USA) connected to a quadrupole-time-of-flight 5,600 mass spectrometer (AB Sciex, Framingham, MA, USA) equipped with a duo-spray ion source operating in positive mode. Lipid extracts were first separated on a Kinetex™ (Kinetex, Atlanta, USA) C182.1 × 150 mm 1.7 μm column (Phenomenex, Torrance, USA). Two solvent mixtures, acetonitrile–water (60:40, v/v) and isopropanol–acetonitrile (90:10, v/v), both containing 10-mM ammonium formate at pH 3.8, were used as eluents A and B, respectively. A 32-min-long binary gradient elution was performed; eluent B was increased from 27% to 97% in 20 min and the mixture was maintained for 5 min; eluent B was then decreased to 27% and the mixture maintained for another 7 min for column re-equilibration. The flowrate was 0.3 mL min<sup>-1</sup> and the column oven temperature was maintained at 45°C. Lipid identification was based on retention time and on mass accuracy peaks from the MS survey scan compared with theoretical masses and fragment ions from MS/MS scan. Relative quantification was achieved with multiquant software (AB Sciex, USA) on the basis of intensity values of extracted masses of the different lipids previously identified. Detector response was normalized by the quantity of FAME previously measured by GC–MS for each sample.

### Phylogenetic analysis and logo sequences

The CvFAP protein sequence was used as bait and blasted against different databases using tBLASTn or BLASTp (including NCBI, Phytozome, Fernbase, and Unigene TARA ocean databases). Sequences from the BLAST were pooled with reference sequences from a previous tree of the GMC oxidoreductase superfamily (Zámocký et al., 2004). Alignment of sequences was done with Muscle algorithm (Edgar, 2004) and viewed with Seaview software (Gouy et al., 2010). Selection of conserved sites was done with Gblock. A set of 226 conserved positions were used for tree construction using maximum likelihood algorithm (PhyML, with LG matrix) with 100 replicates for bootstrap analysis. Annotation of the tree was done using annotation data provided by TARA. FAP Logo sequence was based on 35 sequences including at least one sequence of each taxa from the phylogeny. GMC logo sequence was based on sequences of non-FAP GMC oxidoreductases. Alignment of sequences was done using Muscle algorithm and viewed with Seaview software. Construction of Logo sequences was done using WebLogo (<https://weblogo.berkeley.edu/logo.cgi>).

### Heterologous expression of FAPs in *E. coli*

Coding sequences of putative FAPs from *G. sulphuraria* (Supplemental Figure S13) and *N. gaditana* (Supplemental Figure S14) were directly amplified from cDNAs obtained by reverse transcription of total RNAs. The FAP homologs from *Galdieria*, *Chondrus*, and *Ectocarpus* were codon-optimized and synthesized, then cloned into pLIC07 as described before for CvFAP and CrFAP (Sorigué et al., 2017). Potential transit peptides were removed for better expression in *E. coli*, and N-terminal sequences were as follows: GFDRSREFDYVIVGGG for *Galdieria*; SSEAATTYDYIVGGG

for *Chondrus*; LQSVSMKAPAAVASSTYDYIVGGG for *Nannochloropsis*; SMSVAEEGHKFIIVGGG for *Ectocarpus*. *Escherichia coli* was cultivated in Terrific broth medium at 37°C until OD<sub>600nm</sub> reached 0.8. Expression was then induced with 0.5-mM IPTG and cultures transferred at 22°C under 100 μmol photons m<sup>-2</sup> s<sup>-1</sup>.

### 77K fluorescence emission spectra

Low temperature spectra were measured on whole cells at 77K using a SAFAS Xenius optical fiber fluorescence spectrophotometer (Dang et al., 2014). About 200 μL of light-adapted cell suspension was frozen in a liquid nitrogen bath cryostat. The excitation wavelength used was 440 nm, and detection wavelength ranged from 600 to 800 nm with a 5 nm split. Fluorescence emission spectra were all normalized to the 686 nm signal.

### Accession numbers

FAP is referenced under the enzyme classification number E.C. 4.1.1.106. Accession numbers of FAP genes expressed in this study are XP\_005842992 (*C. variabilis* NC64A), XP\_001703004 (*C. reinhardtii*), CBJ25560 (*E. silicosus*), and XP\_005714951 (*C. crispus*).

### Supplemental data

The following materials are available in the online version of this article.

**Supplemental Figure S1.** Phylogenetic tree of GMC oxidoreductase superfamily.

**Supplemental Figure S2.** Logo sequences for FAPs and other GMC oxidoreductases.

**Supplemental Figure S3.** FAP is predicted to be localized in plastids in many algae

**Supplemental Figure S4.** Ultrastructure of *C. reinhardtii* WT and fap strains.

**Supplemental Figure S5.** FAP gene expression during day–night cycles and hypothetical mechanism that may explain HC loss.

**Supplemental Figure S6.** Growth curves and cell volume of WT and fap strains.

**Supplemental Figure S7.** Growth of WT and fap strains using various concentrations of salt.

**Supplemental Figure S8.** Fatty acid profile in mixotrophic conditions.

**Supplemental Figure S9.** Variation in the proportion of each fatty acid in total fatty acids during cell cycle.

**Supplemental Figure S10.** Profiles of major glycerolipid species in WT, fap, and complementation strains.

**Supplemental Figure S11.** Photosynthetic activity in fap and WT strains.

**Supplemental Figure S12.** Fatty acid acclimation to cold conditions.

**Supplemental Figure S13.** Sequence of *G. sulphuraria* FAP deduced from the cDNA cloned.

**Supplemental Figure S14.** Sequence of *N. gaditana* FAP deduced from the cDNA cloned.

**Supplemental Table S1.** Multiple alignment of FAP sequences used for the phylogenetic analysis (text file).

**Supplemental Table S2.** List of GMC oxidoreductases used for the phylogenetic analysis.

## Acknowledgments

We thank Dr Olivier Vallon for help with analysis of some sequenced algal genomes and helpful discussions. Thanks are due to Dr Quentin Carradec and Dr Patrick Wincker for providing access to TARA sequences and to Dr Florence Corellou, Dr Eric Maréchal, and Prof. Stefano Caffarri for useful discussions. Help of Dr Philippe Ortet and Emmanuelle Billon with analysis of TARA sequences is also acknowledged.

## Funding

This project received funding from CEA (DRF Impulsion Invention E2FAP to F.B.) and from Agence Nationale de la Recherche (PHOTOALKANE, N° ANR-18-CE43-0008-01, to G.P.). This work was also supported by the HelioBiotec platform funded by the EU, the Région Sud, the French Ministry of Research, and the C commissariat à l'énergie atomique et aux énergies alternatives (CEA). S.L.Y.M. has received a PhD scholarship from Ecole Normale Supérieure Paris and the French Ministry of Education and Research.

*Conflict of interest statement.* None declared.

## References

- Beller HR, Goh E-B, Keasling JD (2010) Genes involved in long-chain alkene biosynthesis in *Micrococcus luteus*. *Appl Environ Microbiol* **76**: 1212
- Berla BM, Saha R, Maranas CD, Pakrasi HB (2015) Cyanobacterial alkanes modulate photosynthetic cyclic electron flow to assist growth under cold stress. *Sci Rep* **5**: 14894
- Bernard A, Domergue F, Pascal S, Jetter R, Renne C, Faure J-D, Haslam RP, Napier JA, Lessire R, Joubès J (2012) Reconstitution of plant alkane biosynthesis in yeast demonstrates that *Arabidopsis* ECERIFERUM1 and ECERIFERUM3 are core components of a very-long-chain alkane synthesis complex. *Plant Cell* **24**: 3106–3118
- Björn LO, editor (2015) Photoactive proteins. In *Photobiology, The Science of Light and Life*, Ed 3. Springer, New York, pp 139–150
- Blaby-Haas CE, Merchant SS (2019) Comparative and functional algal genomics. *Annu Rev Plant Biol* **70**: 605–638
- Burlacot A, Sawyer A, Cuiné S, Auroy-Tarrago P, Blangy S, Happe T, Peltier G (2018) Flavodiiron-mediated O<sub>2</sub> photoreduction links H<sub>2</sub> production with CO<sub>2</sub> fixation during the anaerobic induction of photosynthesis. *Plant Physiol* **177**: 1639–1649
- Chua NH, Bennoun P (1975) Thylakoid membrane polypeptides of *Chlamydomonas reinhardtii*: wild-type and mutant strains deficient in photosystem II reaction center. *Proc Natl Acad Sci USA* **72**: 2175–2179
- Coates RC, Podell S, Korobeynikov A, Lapidus A, Pevzner P, Sherman DH, Allen EE, Gerwick L, Gerwick WH (2014) Characterization of cyanobacterial hydrocarbon composition and distribution of biosynthetic pathways. *PLoS One* **9**: e85140
- Dang K-V, Plet J, Tolleter D, Jokel M, Cuiné S, Carrier P, Auroy P, Richard P, Johnson X, Alric J, et al. (2014) Combined increases in mitochondrial cooperation and oxygen photoreduction compensate for deficiency in cyclic electron flow in *Chlamydomonas reinhardtii*. *Plant Cell* **26**: 3036–3050
- de Vargas C, Audic S, Henry N, Decelle J, Mahé F, Logares R, Lara E, Berney C, Le Bescot N, Probert I, et al. (2015) Ocean plankton. Eukaryotic plankton diversity in the sunlit ocean. *Science* **348**: 1261605
- Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinform* **5**: 113
- Gouy M, Guindon S, Gascuel O (2010) SeaView Version 4: a multi-platform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* **27**: 221–224
- Gruber A, Rocap G, Kroth PG, Armbrust EV, Mock T (2015) Plastid proteome prediction for diatoms and other algae with secondary plastids of the red lineage. *Plant J* **81**: 519–528
- Harris EH (1989) The *Chlamydomonas* Sourcebook. A Comprehensive Guide to Biology and Laboratory Use. Academic Press, San Diego, CA
- Herman NA, Zhang W (2016) Enzymes for fatty acid-based hydrocarbon biosynthesis. *Curr Opin Chem Biol* **35**: 22–28
- Heyes DJ, Lakavath B, Hardman SJO, Sakuma M, Hedison TM, Scrutton NS (2020) Photochemical mechanism of light-driven fatty acid photodecarboxylase. *ACS Catal* **10**: 6691–6696
- Jetter R, Kunst L (2008) Plant surface lipid biosynthetic pathways and their utility for metabolic engineering of waxes and hydrocarbon biofuels. *Plant J* **54**: 670–683
- Jiménez-Díaz L, Caballero A, Pérez-Hernández N, Segura A (2017) Microbial alkane production for jet fuel industry: motivation, state of the art and perspectives. *Microb Biotechnol* **10**: 103–124
- Knoot CJ, Pakrasi HB (2019) Diverse hydrocarbon biosynthetic enzymes can substitute for olefin synthase in the cyanobacterium *Synechococcus* sp. PCC 7002. *Sci Rep*
- Lakavath B, Hedison TM, Heyes DJ, Shanmugam M, Sakuma M, Hoeven R, Tilakaratna V, Scrutton NS (2020) Radical-based photoinactivation of fatty acid photodecarboxylases. *Anal Biochem* **600**: 113749.
- Lea-Smith DJ, Biller SJ, Davey MP, Cotton CAR, Perez Sepulveda BM, Turchyn AV, Scanlan DJ, Smith AG, Chisholm SW, Howe CJ (2015) Contribution of cyanobacterial alkane production to the ocean hydrocarbon cycle. *Proc Natl Acad Sci USA* **112**: 13591–13596
- Lea-Smith DJ, Ortiz-Suarez ML, Lenn T, Nürnberg DJ, Baers LL, Davey MP, Parolini L, Huber RG, Cotton CAR, Mastroianni G, et al. (2016) Hydrocarbons are essential for optimal cell size, division, and growth of cyanobacteria[OPEN]. *Plant Physiol* **172**: 1928–1940
- Lee RE, editor (2008) Basic characteristics of the algae. In *Phycology*, Ed 4. Cambridge University Press, New York, p 3
- Lee SB, Suh MC (2013) Recent advances in cuticular wax biosynthesis and its regulation in *Arabidopsis*. *Mol Plant* **6**: 246–249
- Li X, Zhang R, Patena W, Gang SS, Blum SR, Ivanova N, Yue R, Robertson JM, Lefebvre PA, Fitz-Gibbon ST, et al. (2016) An indexed, mapped mutant library enables reverse genetics studies of biological processes in *Chlamydomonas reinhardtii*. *Plant Cell* **28**: 367–387
- Liu K, Li S (2020) Biosynthesis of fatty acid-derived hydrocarbons: perspectives on enzymology and enzyme engineering. *Curr Opin Biotechnol* **62**: 7–14
- Los DA, Mironov KS, Allakhverdiev SI (2013) Regulatory role of membrane fluidity in gene expression and physiological functions. *Photosynth Res* **116**: 489–509
- Mendez-Perez D, Begemann MB, Pflieger BF (2011) Modular synthase-encoding gene involved in  $\alpha$ -olefin biosynthesis in *Synechococcus* sp. strain PCC 7002. *Appl Environ Microbiol* **77**: 4264–4267
- Moulin S, Légeret B, Blangy S, Sorigué D, Burlacot A, Auroy P, Li-Beisson Y, Peltier G, Beisson F (2019) Continuous photoproduction of hydrocarbon drop-in fuel by microbial cell factories. *Sci Rep*
- Moustafa A, Beszteri B, Maier UG, Bowler C, Valentin K, Bhattacharya D (2009) Genomic footprints of a cryptic plastid endosymbiosis in diatoms. *Science* **324**: 1724–1726

- Pagliano C, Barera S, Chimirri F, Saracco G, Barber J (2012) Comparison of the  $\alpha$  and  $\beta$  isomeric forms of the detergent n-dodecyl-D-maltoside for solubilizing photosynthetic complexes from pea thylakoid membranes. *Biochim Biophys Acta* **1817**: 1506–1515
- Qiu Y, Tittiger C, Wicker-Thomas C, Le Goff G, Young S, Wajnberg E, Fricaux T, Taquet N, Blomquist GJ, Feyereisen R (2012) An insect-specific P450 oxidative decarboxylase for cuticular hydrocarbon biosynthesis. *Proc Natl Acad Sci USA* **109**: 14858–14863
- Rude MA, Baron TS, Brubaker S, Alibhai M, Del Cardayre SB, Schirmer A (2011) Terminal olefin (1-alkene) biosynthesis by a novel P450 fatty acid decarboxylase from *Jeotgalicoccus* species. *Appl Environ Microbiol* **77**: 1718–1727
- Rui Z, Li X, Zhu X, Liu J, Domigan B, Barr I, Cate JHD, Zhang W (2014) Microbial biosynthesis of medium-chain 1-alkenes by a non-heme iron oxidase. *Proc Natl Acad Sci USA* **111**: 18237–18242
- Schirmer A, Rude MA, Li X, Popova E, del Cardayre SB (2010) Microbial biosynthesis of alkanes. *Science* **329**: 559–562
- Smith DR, Lee RW (2014) A plastid without a genome: evidence from the nonphotosynthetic green algal genus *Polytomella*[W][OPEN]. *Plant Physiol* **164**: 1812–1819
- Sorigué D, Légeret B, Cuiné S, Blangy S, Moulin S, Billon E, Richaud P, Brugière S, Couté Y, Nurizzo D, et al. (2017) An algal photoenzyme converts fatty acids to hydrocarbons. *Science* **357**: 903–907
- Sorigué D, Légeret B, Cuiné S, Morales P, Mirabella B, Guédénay G, Li-Beisson Y, Jetter R, Peltier G, Beisson F (2016) Microalgae synthesize hydrocarbons from long-chain fatty acids via a light-dependent pathway. *Plant Physiol* **171**: 2393
- Tardif M, Atteia A, Specht M, Cogne G, Rolland N, Brugière S, Hippler M, Ferro M, Bruley C, Peltier G, et al. (2012) PredAlgo: a new subcellular localization prediction tool dedicated to green algae. *Mol Biol Evol* **29**: 3625–3639
- Terashima M, Specht M, Hippler M (2011) The chloroplast proteome: a survey from the *Chlamydomonas reinhardtii* perspective with a focus on distinctive features. *Curr Genet* **57**: 151–168
- Valentine DL, Reddy CM (2015) Latent hydrocarbons from cyanobacteria. *Proc Natl Acad Sci USA* **112**: 13434–13435
- Yamamori T, Kageyama H, Tanaka Y, Takabe T (2018) Requirement of alkanes for salt tolerance of Cyanobacteria: characterization of alkane synthesis genes from salt-sensitive *Synechococcus elongatus* PCC7942 and salt-tolerant *Aphanothece halophytica*. *Lett Appl Microbiol* **67**: 299–305
- Yunus IS, Wichmann J, Wördenweber R, Lauersen KJ, Kruse O, Jones PR (2018) Synthetic metabolic pathways for photobiological conversion of CO<sub>2</sub> into hydrocarbon fuel. *Metab Eng* **49**: 201–211
- Zámocký M, Hallberg M, Ludwig R, Divne C, Haltrich D (2004) Ancestral gene fusion in cellobiose dehydrogenases reflects a specific evolution of GMC oxidoreductases in fungi. *Gene* **338**: 1–14
- Zhou YJ, Kerkhoven EJ, Nielsen J (2018) Barriers and opportunities in bio-based production of hydrocarbons. *Nature Energy* **3**: 925–935
- Zones JM, Blaby IK, Merchant SS, Umen JG (2015) High-resolution profiling of a synchronized diurnal transcriptome from *Chlamydomonas reinhardtii* reveals continuous cell and metabolic differentiation. *Plant Cell* **27**: 2743–2769

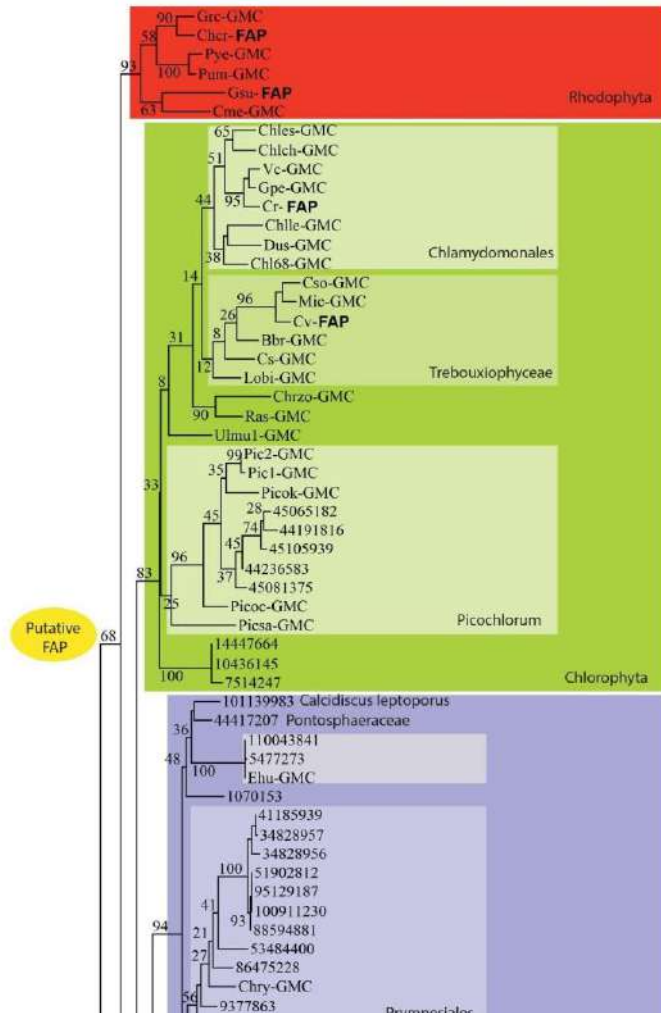


## *Supplementary material*

Fatty acid photodecarboxylase is an ancient photoenzyme that forms hydrocarbons in the thylakoids of algae

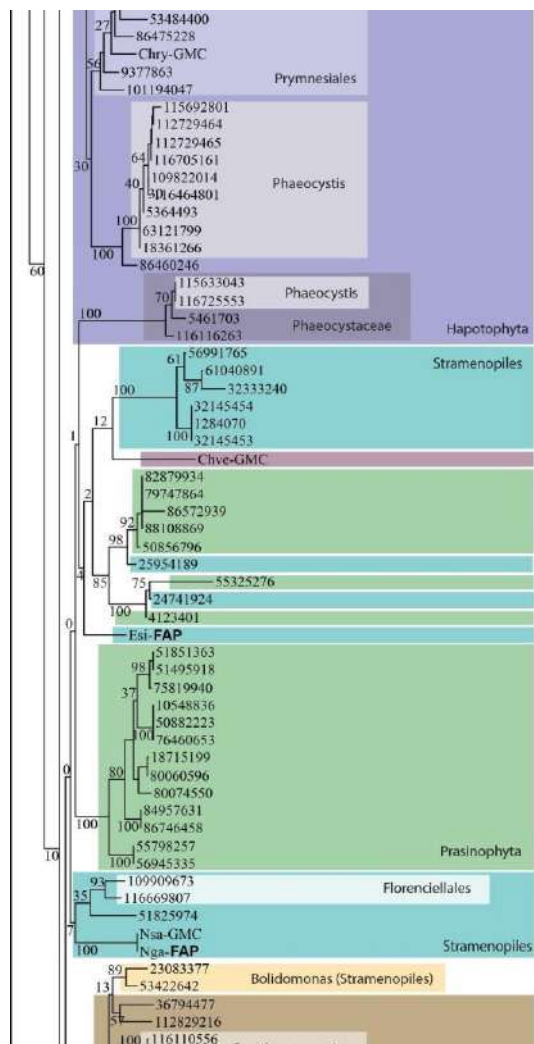
Moulin et al 2021

0.2

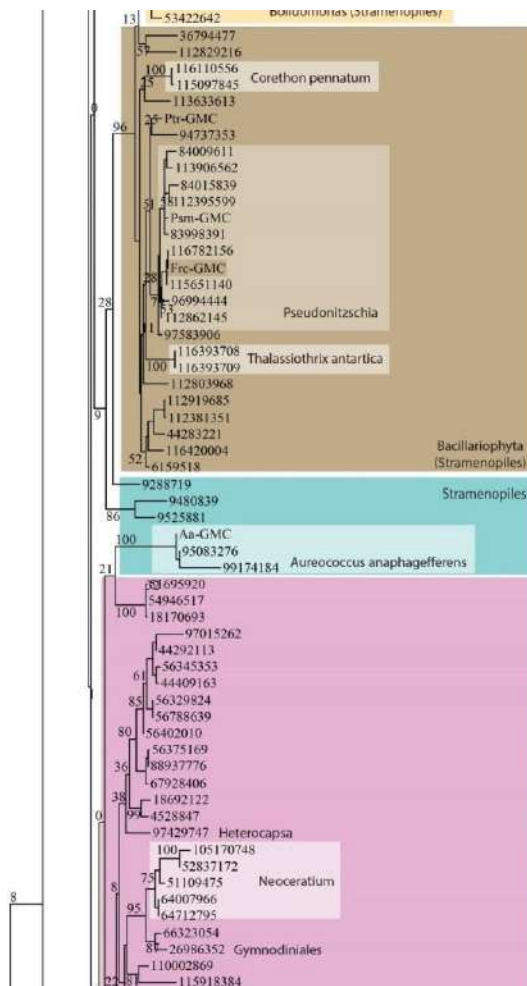


(Fig. S1, continued on next page)

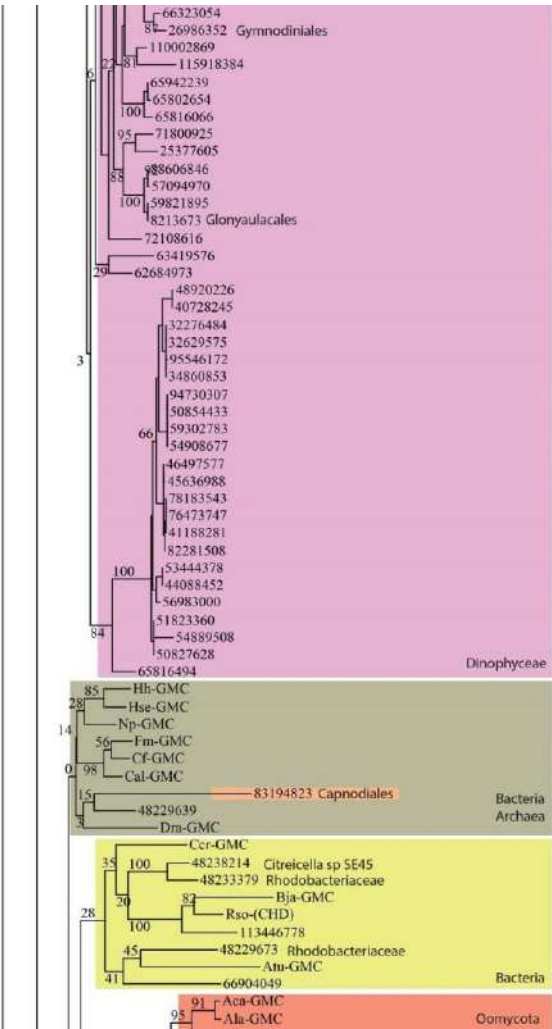
(continued on next page)



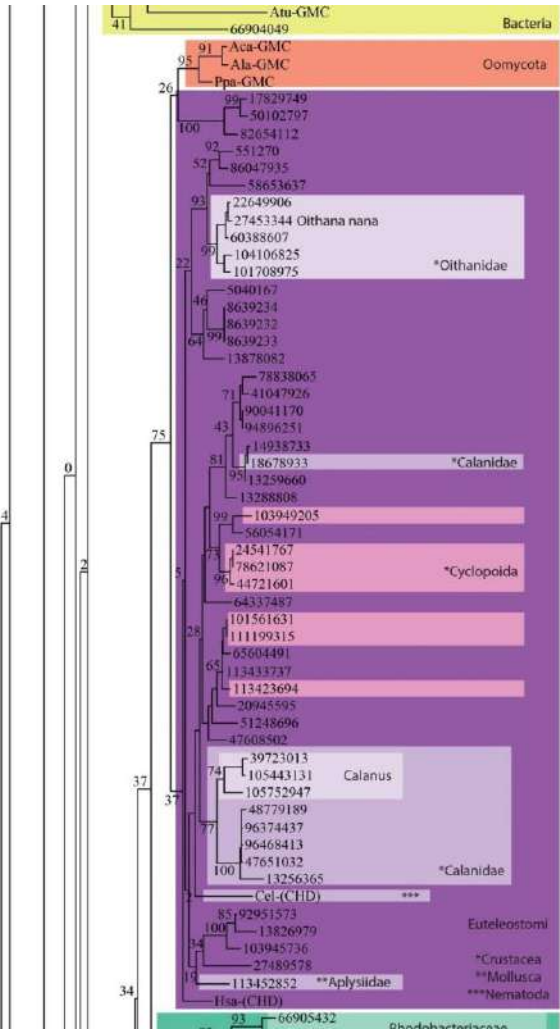
(continued on next page)

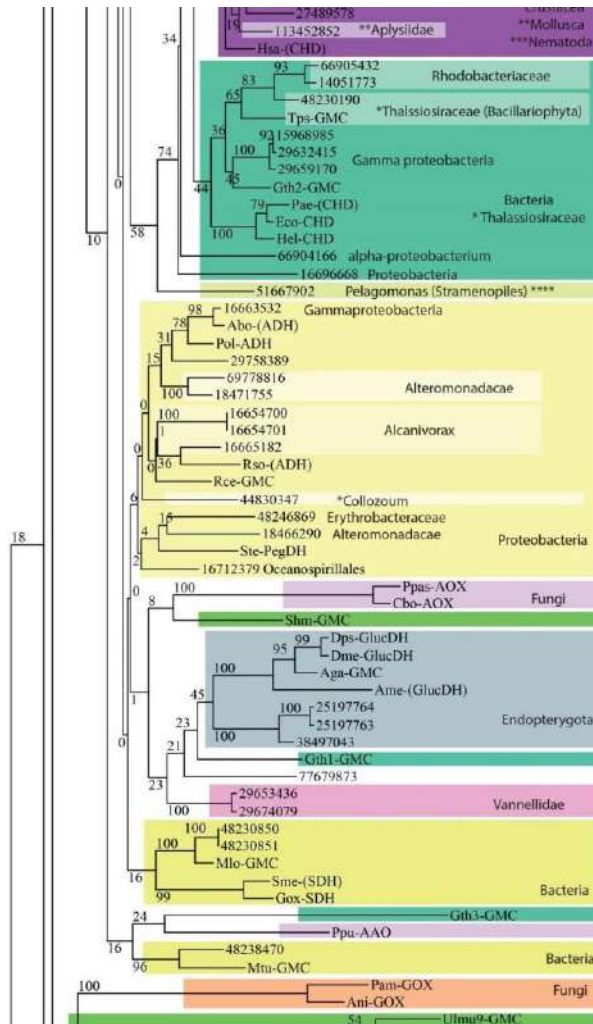


(continued on next page)



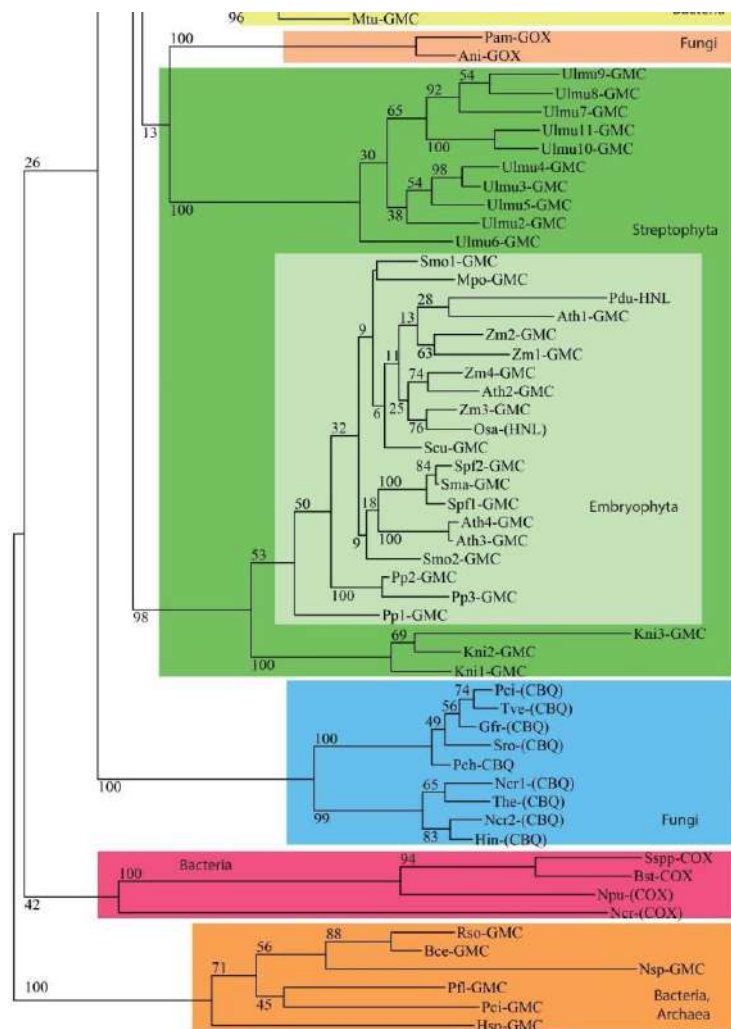
(continued on next page)



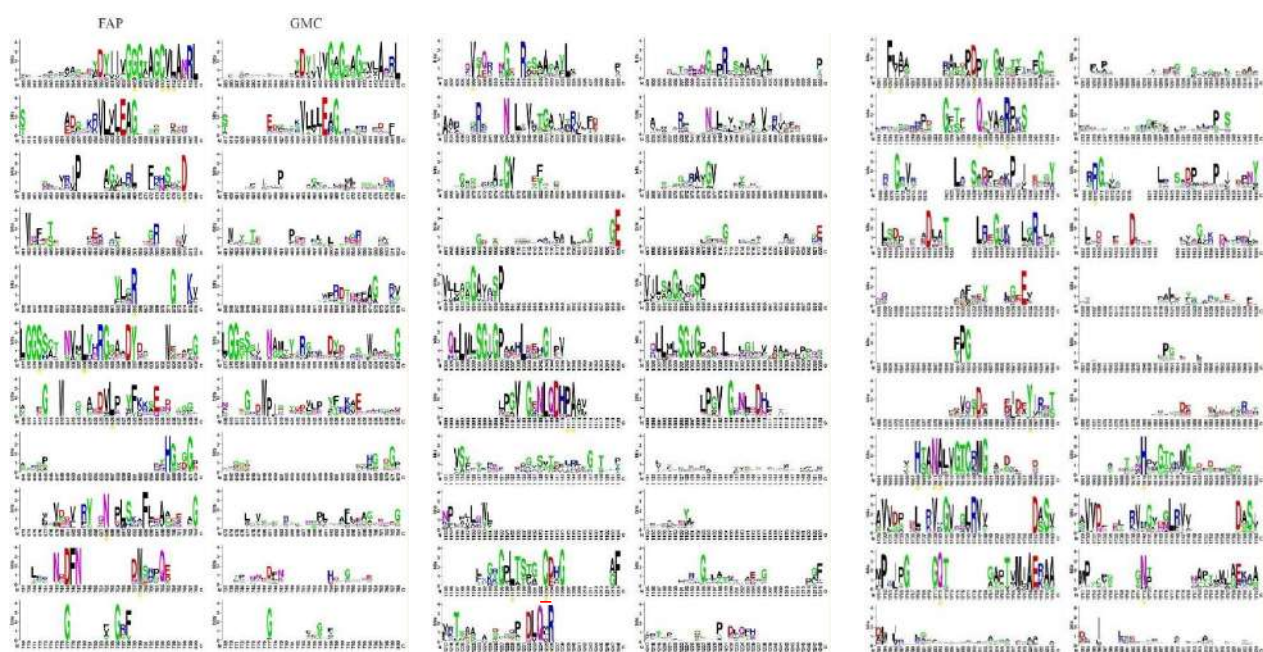


(continued on next page)

**Figure S1. Phylogenetic tree of GMC oxidoreductase superfamily.** The tree of GMC oxidoreductases sequences was built using maximum of likelihood algorithm. Each branch is defined by the lowest taxonomic indication of all the sequences present in the branch. Colors help delineate taxonomic groups. The scale bar shows evolutionary distance in substitutions per site. Numbers on branches are bootstrap values. When biochemical activity is demonstrated capital letters other than GMC indicate it : AOX, alcohol oxidase; FAP, fatty acid photodecarboxylase; CBQ, cellobiose dehydrogenase; CHD, choline dehydrogenase; COX, cholesterol oxidase; GlucDH, glucose dehydrogenase; GOX, glucose oxidase, HNL hydroxymandelonitrile lyase. When indicated in brackets, function has been shown by phylogenetic approach but is not yet supported by an enzyme activity assay. Sequence names consisting only of numbers correspond to TARA sequences. The correspondence between sequence names and species is given in Table S2.



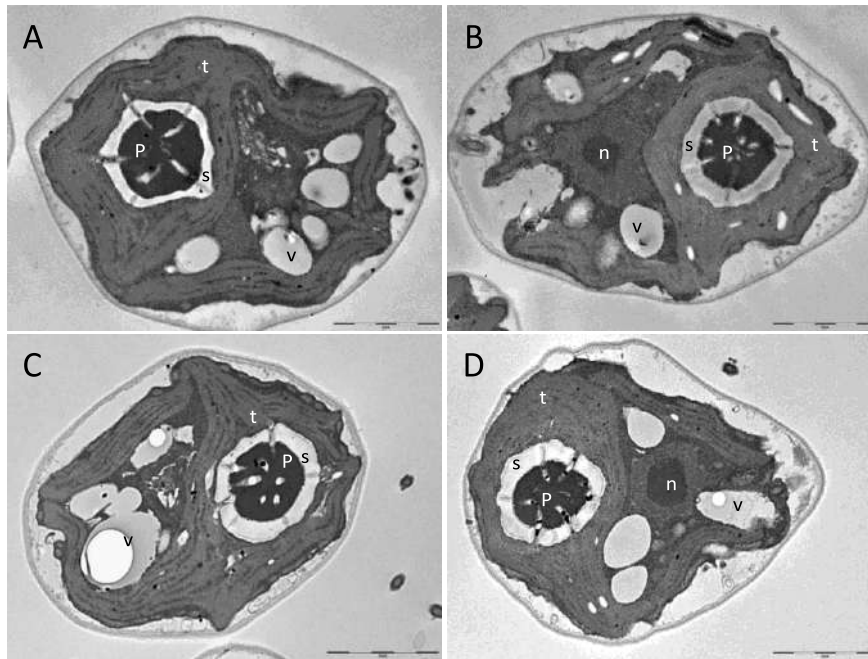




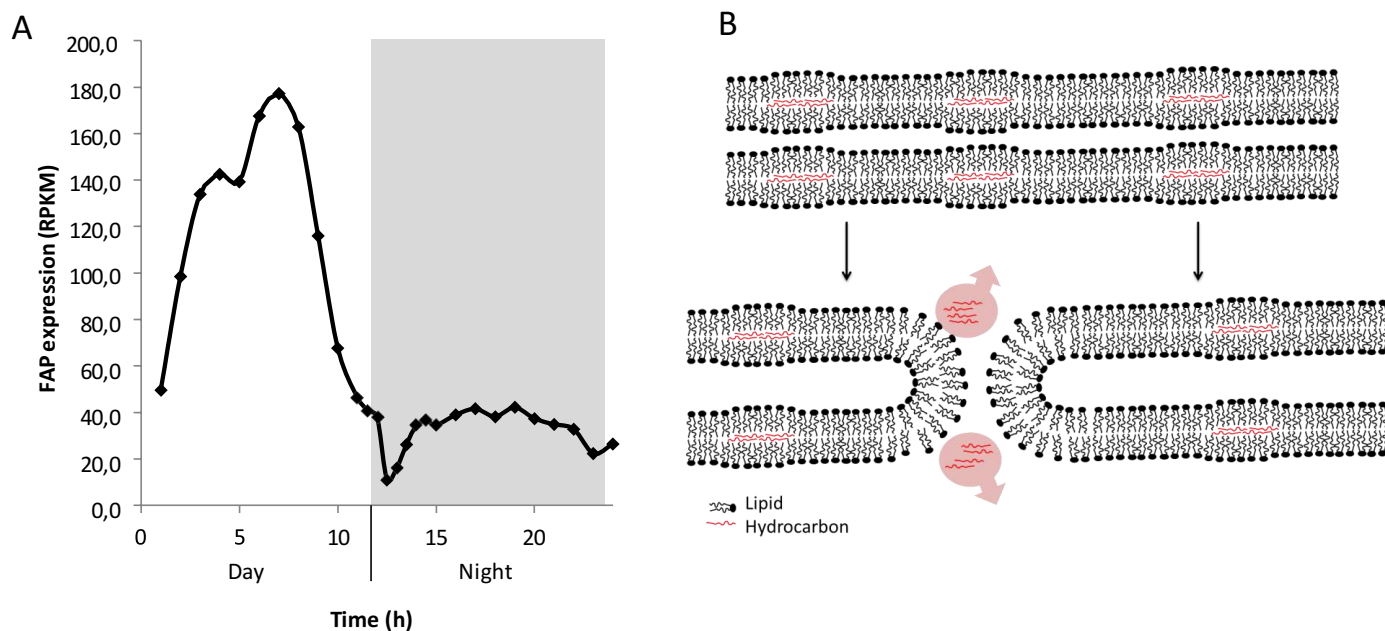
**Figure S2. Logo sequences for FAPs and other GMC oxidoreductases.** Sequences used for alignments are in supplemental table S1. Letters represent amino acids at successive positions in the multiple alignment. The overall height of the stack indicates the sequence conservation at that position, while the height of symbols within the stack indicates the relative frequency of each amino acid at that position. Red, blue, black, purple and green letters: acidic, basic, hydrophobic, amide and other polar residues respectively. Yellow dots indicate conserved residues specific to FAPs. Residues C432 and R451 of CvFAP active site are indicated by a yellow dot underlined in red (bottom of third column).

	PredAlgo					
	Score M	Score C	Score SP	Prediction		
<i>Chlorella variabilis</i>	1,846	3,452	0,004	C		
<i>Chlorella sorokiniana</i>	2,416	2,287	0,016	M		
<i>Chlamydomonas reinhardtii</i>	0,356	3,55	8E-04	C		
<i>Coccomyxa subellipsoidea</i>	0,176	1,898	2E-04	C		
<i>Volvox carteri f. nagariensis</i>	0,007	1,612	6E-04	C		
<i>Dunaliella salina</i>	0,625	0,796	1E-04	C		
<i>Raphidocelis subcapitata</i>	0,119	4,151	0,033	C		
<i>Gonium pectorale</i>	0,246	3,358	0,003	C		
<i>Micractinium conductrix</i>	4E-04	1E-06	5E-05	-		
<i>Chondrus crispus</i>	0,445	1,24	0,048	C		
<i>Gracilariopsis chorda</i>	0,004	0,898	0,002	C		
<i>Galdieria sulphuraria</i>	2E-04	0,092	0,005	-		
<i>Cyanidioschyzon merolae</i>	0,551	0,577	0,005	C	Signal Peptide	ASAFind Plastid (Confidence)
<i>Nannochloropsis gaditana</i>	0,083	0,107	1,657	SP	+	YES (High)
<i>Nannochloropsis salina</i>	0,083	0,107	1,657	SP	+	YES (High)
<i>Pseudo-nitzschia multistriata</i>	0,005	0,241	0,897	SP	+	NO
<i>Chrysochromulina</i>	0,405	1,581	0,667	C	+	NO
<i>Emiliana huxleyi</i>	2,904	0,584	0,377	M	+	NO
<i>Ectocarpus siliculosus</i>	0,512	1,527	0,203	C	+	YES (High)

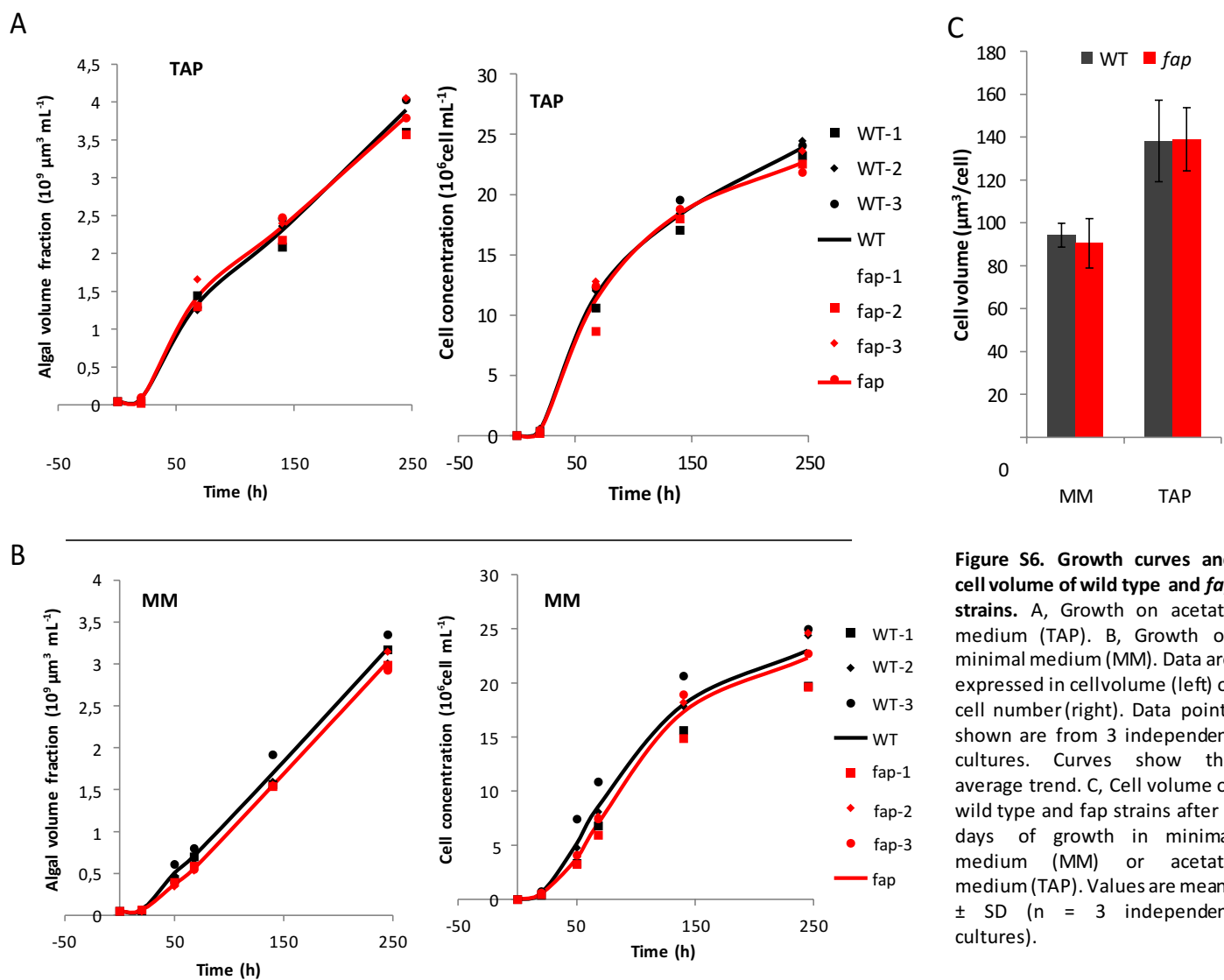
**Figure S3. FAP is predicted to be localized in plastids in many algae.** Sequences from various algae with sequenced genomes were analyzed with two algorithms adapted to algal sequences. Names of green and red algae species are shaded in green and red respectively; algal names that are not shaded correspond to secondary algae. ASAFind has been developed specifically for algae with secondary plastids. M: mitochondria, C: chloroplast, SP: secretory pathway. For secondary endosymbiosis algae, ASAFind results are presented with presence of transit peptide according to SignalP and confidence of plastidial localization according to ASAFind. Hyphens: no localization predicted by PredAlgo. Predicted plastidial localization are shaded in bright green.



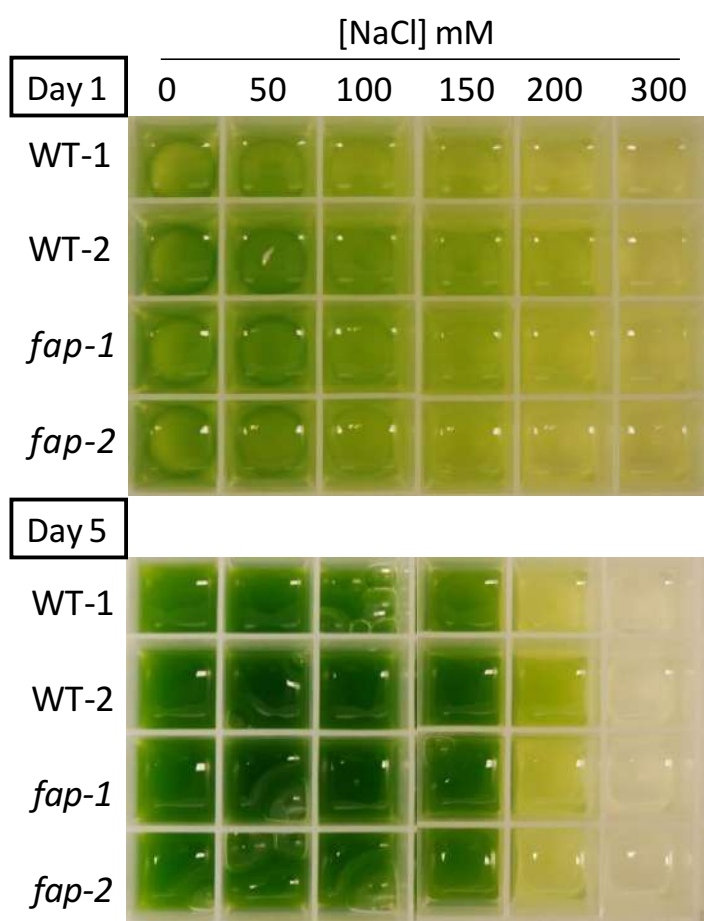
**Figure S4. Ultrastructure of *C. reinhardtii* wild type and *fap* strains.** Transmission electron microscopy of wild type (A,B) and *fap* strain (C,D). Thylakoids (t), nucleus (n), vacuoles (v), pyrenoid (p) and starch (s). Entire scale bar: 2 μm.



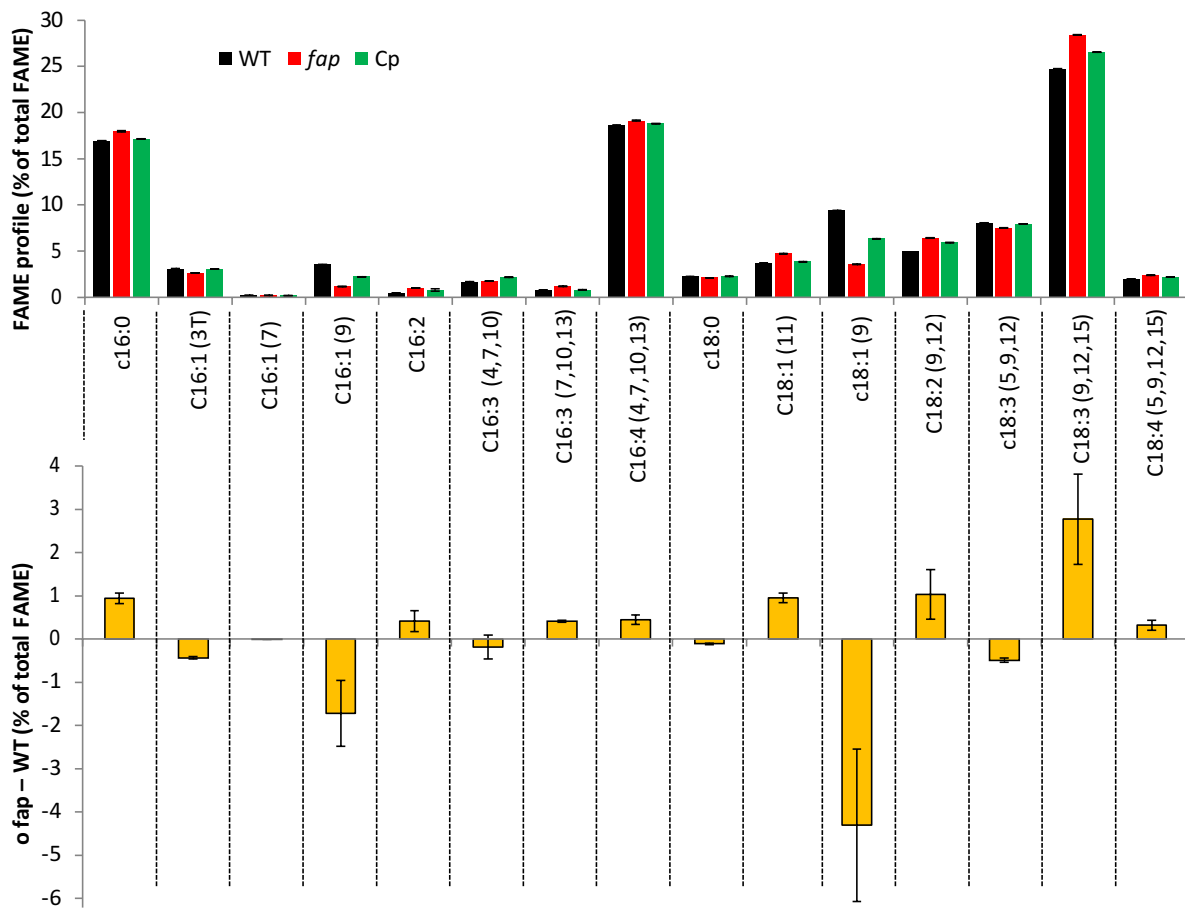
**Figure S5. FAP gene expression during day-night cycle and hypothetical mechanism that may explain HC loss.** **A**, FAP expression from transcriptomic data (from Zones et al., 2015). RPKM: reads per kilobase of transcript per million reads mapped. Diamonds indicate time of actual samples. **B**, Proposed model for HC loss in *Chlamydomonas* chloroplasts during cell division. The HCs in the bubbles are those lost. Alternatively, HCs could be metabolized.



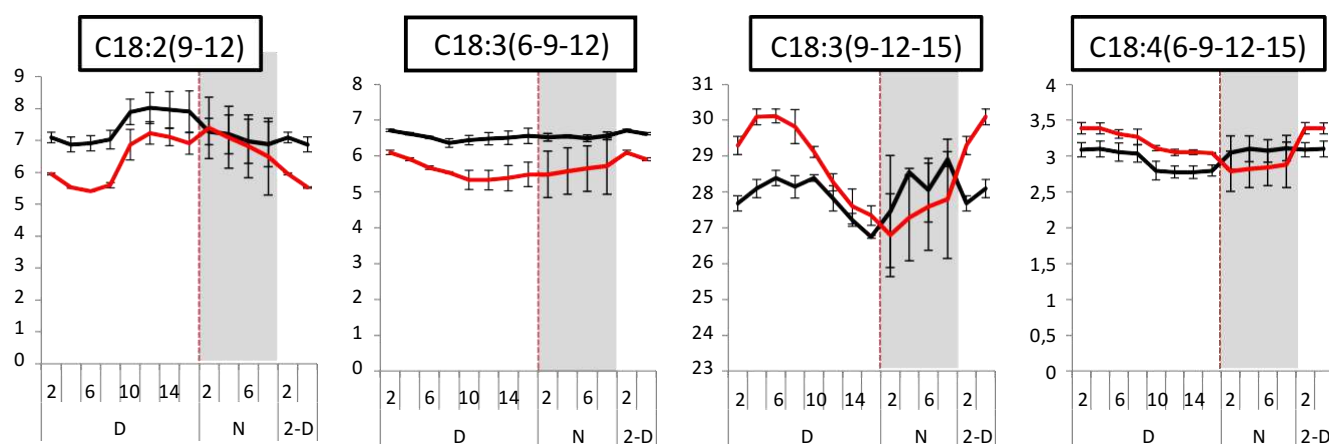
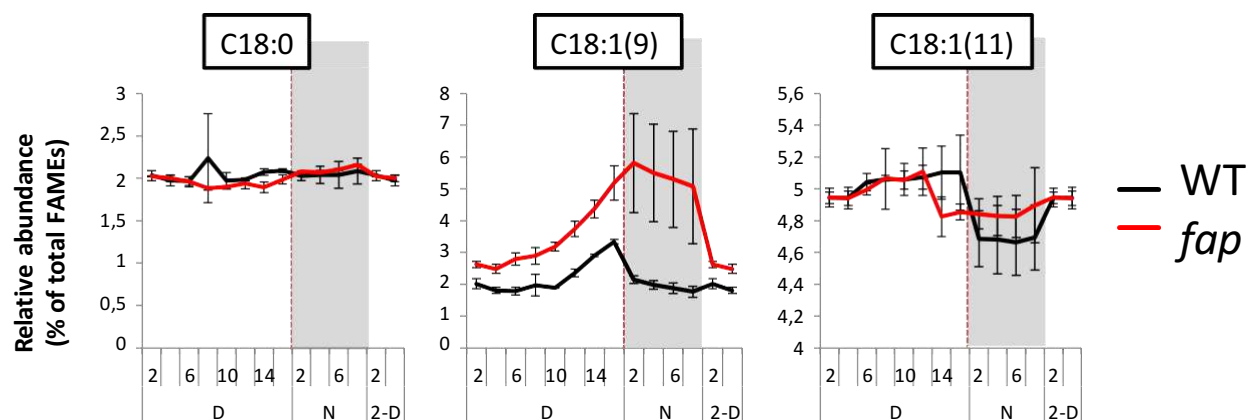
**Figure S6. Growth curves and cell volume of wild type and *fap* strains.** A, Growth on acetate medium (TAP). B, Growth on minimal medium (MM). Data are expressed in cell volume (left) or cell number (right). Data points shown are from 3 independent cultures. Curves show the average trend. C, Cell volume of wild type and *fap* strains after 6 days of growth in minimal medium (MM) or acetate medium (TAP). Values are means  $\pm$  SD ( $n = 3$  independent cultures).



**Figure S7. Growth of wild type and *fap* strain using various concentrations of salt.** Cultures in acetate medium (TAP) were exposed for 1 or 5 days to salt concentrations from 0 to 300 mM NaCl.

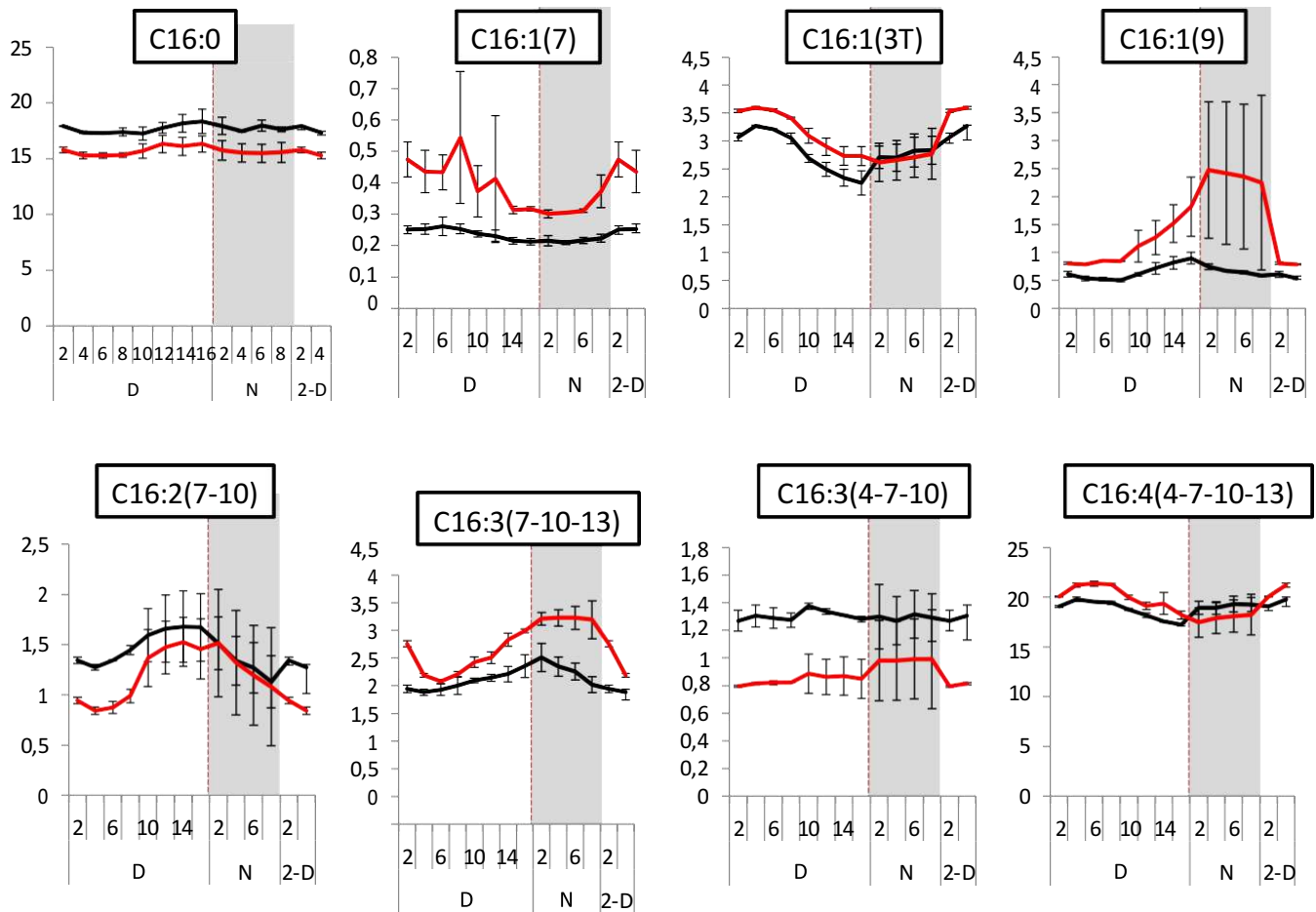


**Figure S8. Fatty acid profile in mixotrophic conditions.** Relative abundance of fatty acids methyl esters measured after transmethylation of whole cells and analysis by GC-MS. FAMES are separated by dashed lines. The bottom panel shows for each fatty acid the difference of relative abundance between *fap* and WT strains. Cells were grown in TAP medium, under 80  $\mu\text{mol photon m}^{-2} \text{s}^{-1}$  in Erlenmeyers (means  $\pm$  SD; n = 3 biological replicates).

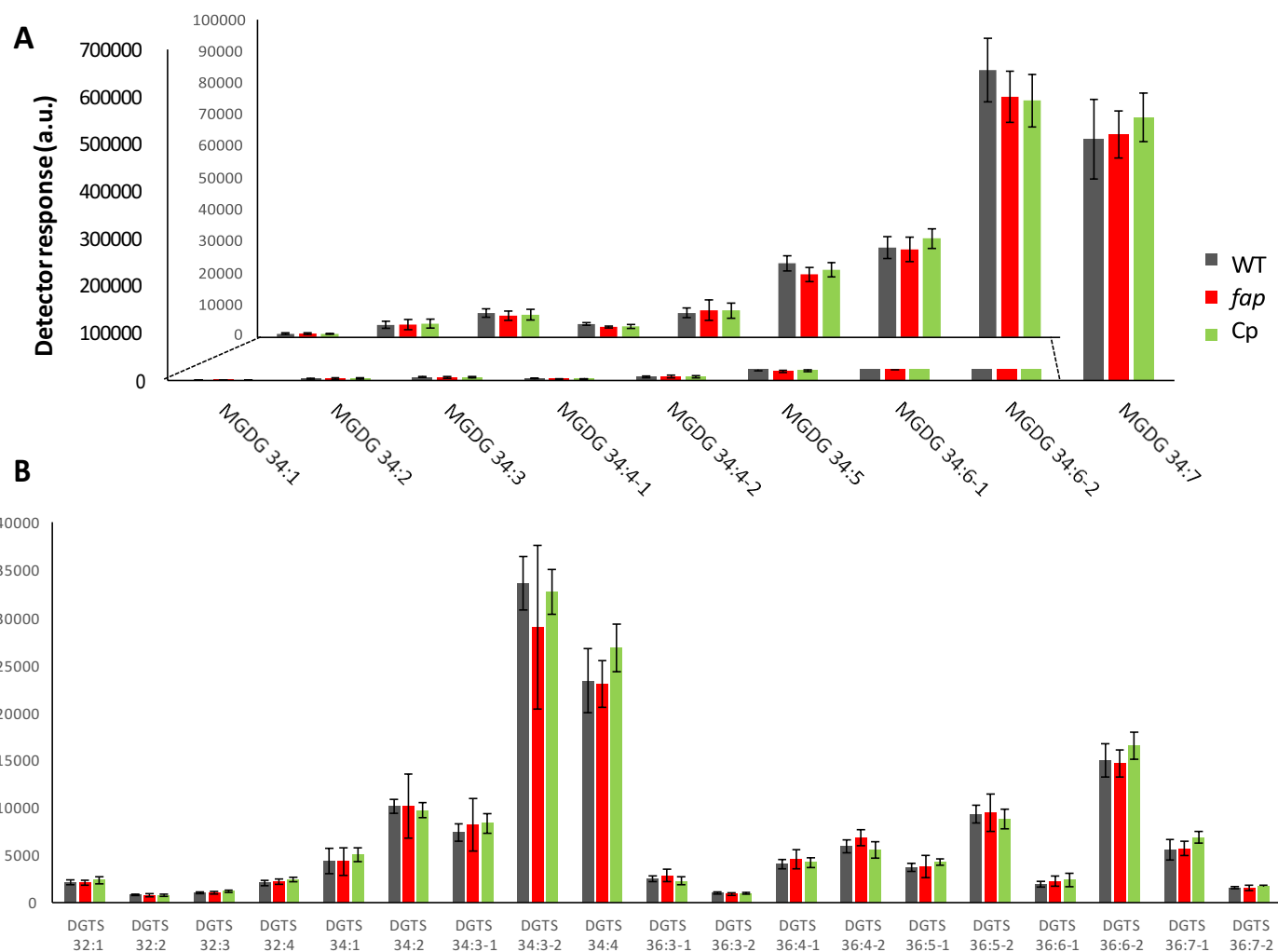


(continued on next page)

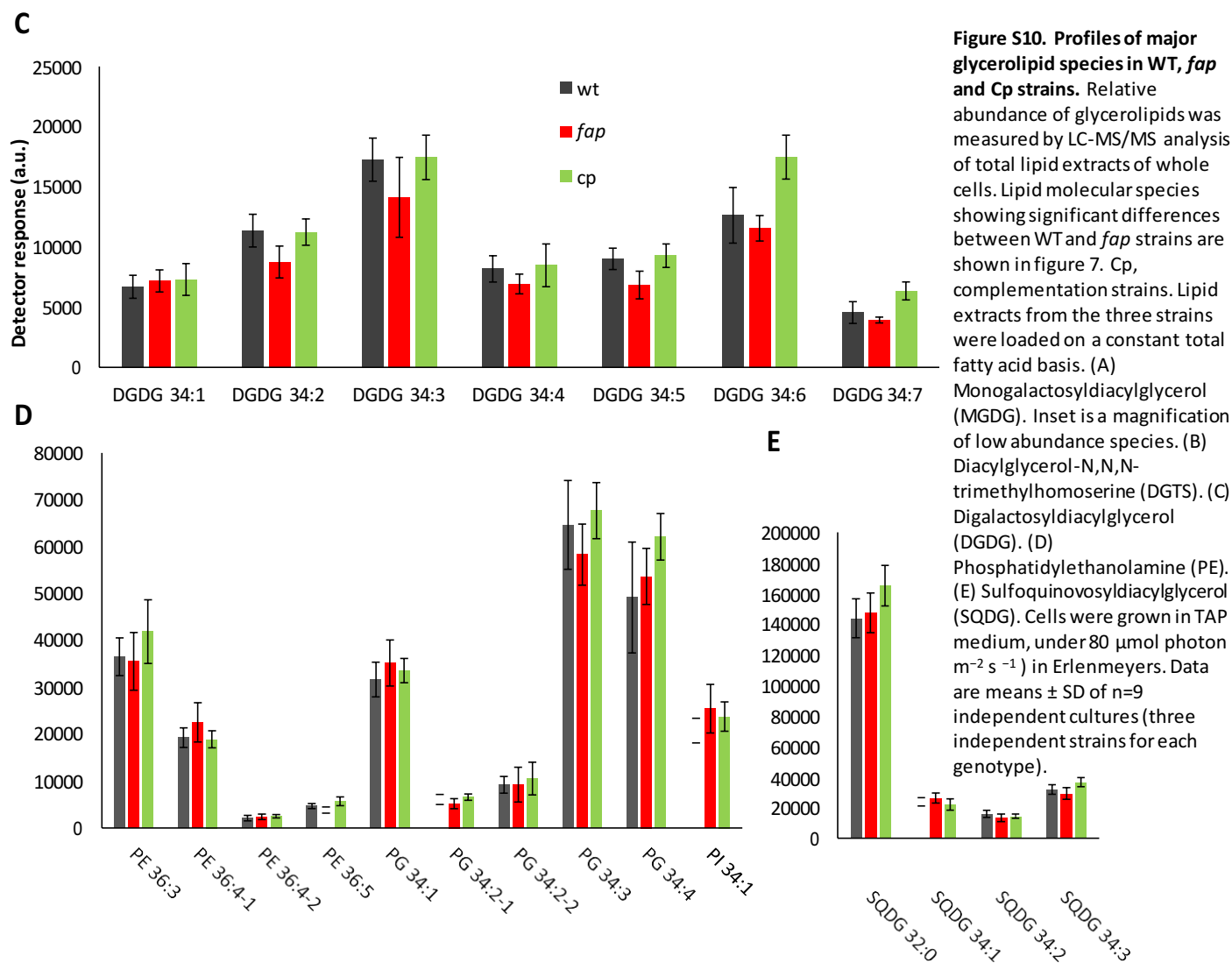


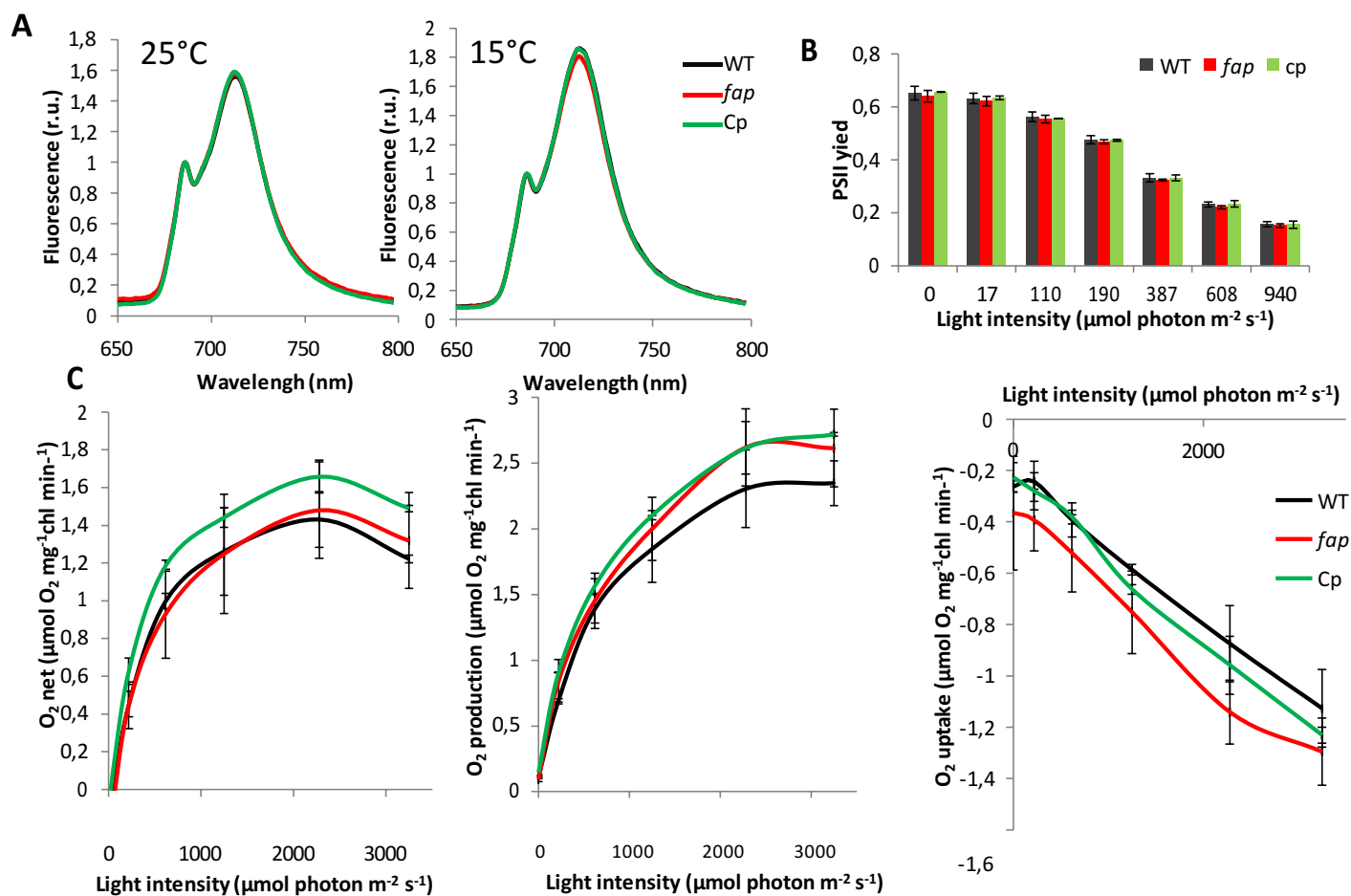


**Figure S9. Variation in the proportion of each fatty acid in total fatty acids during cell cycle.** Relative abundance of fatty acid methyl esters from transmethylation of synchronized cells analyzed by GC-MS and expressed as percentage of total FAMES along a day-night cycle to visualize the cycle (means  $\pm$  SD; n = 3 biological replicates). D: day (16 hours, white), N: night (8 hours, gray), 2-D: first 2 measurements of the following day.

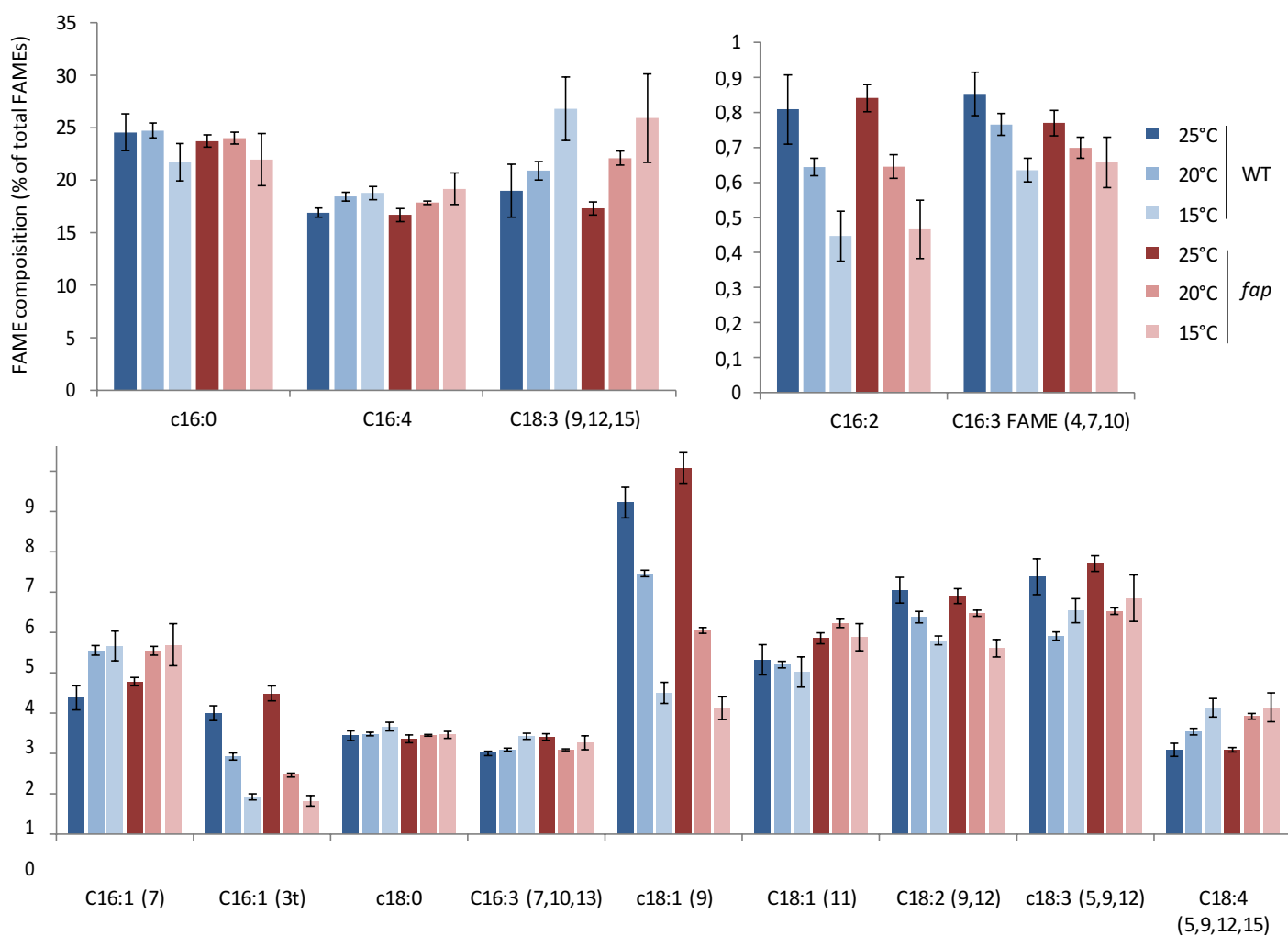


(continued on next page)





**Figure S11. Photosynthetic activity in *fap* and WT strains.** A, 77K fluorescence spectrum for cells grown at 25°C and 15°C. Relative units (r.u.) normalized to the 686 nm signal. B, Photosystem II operating yield under various light intensities. C, Net oxygen, oxygen uptake and oxygen production measured by membrane inlet mass spectrometry after acclimation for 2 minutes at each light intensity. Values are mean  $\pm$  SD ( $n = 3$  independent cultures using three independent strains for each genotype).



```

>CDS-Gsu
ATGAATCTCATGAAGACCTCATTTATTTCTGTGTTGTTTACCAAGTTGTACGTCGTCAGGACTACATCAAAGGTTTGGAAAGAAGCTTCTCTAAAGTGTTCAATTTAGAAAAAGCTCGCGCGCTT ATTTTATAAATGG
AAAAGATTTTCAAGTAGAATAACCGCAACAGCTTCTGTGGAACAATGAGGTCAGTCTCAGTATGTCCTACGAGAGAGTTTGATTATGTAATTGTTGGAGCTGGTGACGACGGGTG TGTTCTTG CTTCCTGTCTGTCCGA
GGATAAAGCTTCCACAGTTTATTATTGGAAAGCTGGCAAGGAAATGAGAACTTTACATCCACGTTCCCATGGGCTTTCCTTATCTCGTTGGAAAGCAGCTAGACTGGAAGTACCAGTCGACCT CTGAAGAGAGATTGCT
TGATAGAAAATAATGTGGCCAAAGAGGAAGGTTTGGGAGGAAGTCATGCGATATC CGTTATGTTATATCTCCGGGGAAGTGCCAA CGATTATAGAGAGTGGGAGAGATGTGGAGCGCTTGCT GGGGTCTGAAGA
CGTTTGCCATATTATTTGAAGGCAGAAAATAAATTGCGCGGTGTTGGTCCCTATCATGGAAGGTTGCTTCTTAACTGTCA GCGATCTCCCCAGTCCGAATGAAATGTCAATGCTTTTGTA TTGCAAGGCAGACTTTG
GATATTCTTACAACAAAGACTTCAACGATTGGTCTCATTCAAGAGGGAATAGGCTT GTTTCAGTGACTCAAGCAAATGGAAACGTGTTTCACCTGCCACCGCATATTTACATCCTGTTGAAAGTAGACGCAATCTTT
TTATTGAAACTCAAAGACAGTGGAGAAAGATCTTGTTTACAAAAGAAATCGGAAGTTGCTCGCGCTATCGGTGTTTCTTATATCAATTCGAATGGTAAGAGAGACGAGCCATGATACGTAAA GAAGTTGTCGTTTCAG
CTGGTG CATATGGCAGTCCACAGCTATTGATGTTATCGGGCATAGGACCATCGAACGTTCTTGATTCTATTGGCATTCCGACTGTTATGCCTCTTGAAAGAGTTGGAAAGAAATTTA CAA GATCAC TTGCAGTTATGTTTCA
TGCAAGTCACCCGACCTTCGAAAGGACCGAAAACGTCGAAATTTGATTATACAGATGATACTGGAAGAACTGGAAGACAATACTGACTTGGTTCTTAAGTGGAAGGTCCTTTAACAGTAC GATGTGTGAAGCTGG
AGCATTTCTAAAAACAAACCTACTTTGGACGATCCCGATTGCGAGCTTCGGTTTATTCCTTTTTTTTCGGAAGCTGATCCGTAATTTTCTTGAGTGATTACTCATCAAAGGGTATGTTTCTTA GAAATCGTTCTTACCGACCC
TCAGGTTTCACCATCCAATCCGTTGCTATACGTCCAAAATCCCGAGGAAAACCTGAGAAATCCAGTCGAAAGATCCAACGTGAGCGCAATTATCGAGTCTGGTTGGTTTT CTTCACAAGAAAGACTTGGAAACGCTTCTTCGAG
GCATTTCAATTTCAAAAAGTTGCTCGGACTTGTCCACTTTCTCCTATTGTTGGTGAACAATGCTTTCCATCACCTTCGCTATCCAAGGAAGAAGATATTATTCGTTATATTTAGGCACCTGT CATACTGCAAAATGCAGTCGT
TGGTACATGTCGATGGGAACGGATAAACAAGCAGTTGTCATCCCAATCTCAAGTCATGGGAGTAGAACGCTTA AGAGTGCTCCTACAATTATGA

>Protein-Gsu
MNLMTKSFISCCCLPSCSSRTTSKVLERSFPKVFQFRKCSRAYFINGKRFSSRITATASVDNEVQSQQYVRTREFDYVIVGAGAAGCVLASRLSEDKRSTVLLLEAGKEDENFYIHVPMGFPLYV GSDLDWKYQSTSEERLLDRKIMWPR
GKVLGGSHAISVMILYLRGSANDYREWERCALGWGPEDVLPYYLKAENNLRGVGPYHGKGGPLTVSDLPSPNEMSHAFVIAQGTLDPYNKDFNDWSHSQEGIGLFQVTQANGKRVSPATAYLHPVRSRRNLFIEQRHVEKILF
TKEIGSCPRAIGVSYINSGKRERAMIRKEVVWSAGAYGSPQLMLLSGIGPSNVLDIGIPTVMPLEGVGNLQDHFVAFMFSKSPDPSKDRKRNNLYYTDGKDWKTLTWFWSGKGPLTSTM CEAGAFLKTKPTLDDPDLQLRFI
PFFSEADPYFSLSDYSSKGMFLRNRYSRPSGFTIQSVAIRPKSRGKLRIQSKDPTVAPIESGWFSSEQEDLETLLRGISFSQKLSRCTLSSYCGEQCFPSLSKEEDIIRIYILGTCHTANAVVGTCTCRMGTQKQAVVNPNLQVMGVERLRV
LLQL

```

**Figure S13. Sequence of *Galdieria sulphuraria* FAP deduced from the cDNA cloned.**

>CDS-Nga  
 ATGTCATCAAAATGGGTACCTTCGGGCCTACCACTCCTGATCGCCCTCCTGATACAGCCAACGCTTTTAAATTACGCTCCACGCCTCTCCAAGACCACTACGGCCTCCAATC  
 CTTCGTGACAGCAAAATATGGAGTGAGGAGCAATCTCTCCGAGGGGGGCTACAATCTGTTTCGATGAAGGCAACGGCTGCCGTGCTTCGTCCACTACGACTACATCA  
 TAGTCGGTGGCGGCATCGGCGTTGTGTTTCGCCAACCGACTGACAGAAAGCGGCGTTTCAAGGTCTTCTCTGGAGGCCGGAATCAGCAGAGCGAAACCCGTACGT  
 GAACATCCCCGCGGCGTGGTCCGCTCTTAAATCGGCCTTGGATTGGCAGTTTGAGAGCGCCCCGAAAGACACCTGGACGGGAAGGAGGTGTAATGTCAGGGGGAA  
 GGCATGGGAGGCTCAGTGGGTGAATGTGATGCTCGTACACCGCGCTCGGCCTCCGACTACGCGAAGTGGGAGGCGGAGGGAGCTCAGGGATGGGGCCGGAGGAA  
 GCCTTACGCTACTTCAAGAAAATGGAAAGACAATTGGTGGGCGGGGAAGGGCGGTGGCACGGCCAGGGGGGCATGTACCCGGTGGACGACGTCAAGTACCAAAACCTTTG  
 TCCAAGCGCTTTTACAGGCGTGCAGGAGTACGGGTGGCGGGCGAAACCCGATTTTAAACGACTGGTCCCATCCCAGGACGGATACGGGAGCTTCAAAGTGGCGCAGAAAG  
 CACGGCAAACGGGTGACGGCGCCTCCGGGTACTTGAACAAGGCCGTTGACGTCGACCAAACTAGACATTCTGTGAGAGGCCCTTGTGACCCGGGTTTTGCTTGAGGGAG  
 AGGGGAGTGTGAAGGCGGTGGGGTGGAGTTACGGGCAAAGATGGGAAGACGACCAAGTTCCGACACGGGGAAGGCGGGAAGTTCCTTGGCGGGCGGGCGG  
 TCAACAGCCCCCACTCCTGATGCTGAGTGGGATCGGACCGGAGGAGACCTCCAGGCAGTAGGCATCGCGACCAAGTAAACCGTCCGGGCGTGGGGGAGAACTTGAAG  
 ACCATCCCGCGTCACTATTGCGCACAAACATCACGCGCCCTATCTCGCTCTGCGACGACCTCTTCTTTTCCACACCTGTCCCTAAGCCCCACAGGTGCTCCGGTGGACCT  
 GACCGGATCGGGCCCGCTGACCAACCCAGGCTGCGACCAACGGCGCGTTCCTGAAAGCCCGAGAAGACCTCCAGGAACCCAACTACAGTTTCGTTTCATAGCCGGGCGGGT  
 TCGGACCTTGACGCGGTGCGCTCTACATCATGGGAGGCTCAGCACGCCCCCTGTGGGATTGACGCTTCAAGTCGTAACATCCGCCC CAAAAGCAAAGGGAAGCTGACCTT  
 GCGAGCAAGGACCCGTTGAAGAACCCCGGATCGAAGTCCGTACTTGTCCGCGCCGAAGACTTGACAGGCCCTTGCACGGGCATGAGGA TTGGC CGGGATCTGATCAA  
 GCAACGGGCTTTTCCGATATCTTGGACGGAAGTGTCCCGGGCCCCCGCTCAGACGGAAGGAACTGGACGCTACATCCGGGATAGCTTGACACGGCCAAACGC  
 CCTGGTAGGCACCTGCAAGATGGGGAGTGTGGAGGATCGGAACGCAAGTGGTGGACCCGAGTCCGGGTATCGGCGTGGGAGGCCTGCGGGTGGTGGATGCAAGCGTG  
 ATGCCCTTATCCCCGAGGGCAAACGGGAAGTGGACGACCATGTGCGCCGAAAAAGCGGCCGACCTGGTCAGGGCCACGCTGGGATTTGGTGGAGATGGGAGTGCA  
 AGATGAGGAGAGGAAGGGAGGATGTTTTAACGGCTGTTAGGGCGCAAAACAGAAAGCGCGCGACGTGA

>Protein-Nga  
 MSSNGYLRAYHLLIALISANAFITPPRLSKTTIGLQSFVTANYGVRRRAISLRGGLQSVSMKAPAAVASSTYDYIIVGGGIGGCVL  
 ANRLTESGRFKVLLLEAGKSAERNPYVNIPAGVVRLFKSALDWQFESAPERHLDGKEVYLVVRGKAMGSSAVNVMLVHRGSAS  
 DYAKWAEGAQGWGPPEALRYFKKMEDNLVGGEGRWHGQGGMYPVDDVKYQNP LSKRFLQACEEYGRANPDFNDWSHP  
 QDGYGSFKVAQKHGKRVTAAASGYLNKAVRRRPNLDILSEALVTRV LLEGEEDVKAVGVEFTGKDGTKHQVVRTTGKAGEVLLAG  
 GAVNSPQLLLMSLIGPEADLQAVGIATKVNRPVGVGENLQDHPAVTIAHNITRPI SLCDLFLFHTPVPKPHQVLRWTLTGSGPLT  
 TPGCDHGAFLKTRDLQEPNVQFRFIAGRGSDPDGVRYSIMGG SARPLSGLTLQVVNIRPKSKGLTLASKDPLKKPRIEVRYLS  
 AAEDLQALRTGMRIGRDLIKQRAFDILDEEVFPGPAAQTDEELD AYIRDSLHTANALVGTCKMGSVEDRNAVVDPECRVIGVG  
 GLRVVDASVMPVIPGGQDTGSGTTMLAEKAADLVRAHADLVLEMGVQDEERKGGWFNGLLGRKQKAAT

**Figure S14. Sequence of *Nannochloropsis gaditana* FAP deduced from the cDNA cloned.**





## **Chapter 3: Study of the mechanism of fatty acid photodecarboxylase**



## Objectives, personal contribution and main results

Following the initial work that led to the first proposal for the photocycle and mechanism of CvFAP, several open questions remained, including: (i) Which structural features of the FAP active site promote substrate stabilization and favor the abstraction of an electron from the fatty acid carboxylate by the blue-light excited FAD cofactor (forward electron transfer)? (ii) Is decarboxylation instantaneous upon this forward electron transfer step or is it slowed by an activation barrier? (iii) Does conversion of alkyl radical to the alkane occur by back electron transfer from  $\text{FAD}^{\bullet-}$  coupled to a proton transfer or by hydrogen atom transfer from a nearby amino acid? (iv) What is the origin of the proton or the hydrogen atom? The objectives of the large international consortium studying FAP mechanism in which my laboratory participated was therefore to continue to elucidate the mechanism of FAP using a wealth of biophysical, biochemical and computational approaches and mutants of the active site.

**Personal contribution:** The study of the detailed mechanism of the enzyme started before the beginning of my thesis but I had the chance to contribute significantly to this work. I participated to the production and purification of the protein on a large scale for the different pluridisciplinary biophysical approaches (mainly for the crystallography experiments using free electron laser x-ray sources (XFEL)). I also participated with Dr. Damien Sorigué and Dr. Pascal Arnoux to the biochemical and static crystallography experiments, especially on the wild type protein but also on the R451K mutant. Finally, I helped to perform some transient absorption spectroscopy experiments during a stay of several weeks in the laboratory of Dr. Pavel Müller and Dr. Klaus Brettel in CEA Saclay. This whole experience was enriching because it allowed me to familiarize myself with the purification and manipulation of CvFAP but also because it introduced me to different biophysical tools necessary to the study of the enzyme and photoenzymes in general.

**Main results:** A 1.8-Å-resolution x-ray crystal structure of CvFAP was obtained and revealed position of water molecules as well as a second fatty acid in the entrance of the active site. Two strictly conserved residues, Arg451 and Cys432 were shown to be essential for proper positioning of the substrate and some water molecules and for electron abstraction from the fatty acid carboxylate. Several putative steps of the mechanism were also demonstrated : decarboxylation is shown to occur quasi-instantaneously upon this forward electron transfer ; back electron transfer from  $\text{FAD}^{\bullet-}$  is demonstrated to be coupled to the transfer of an

exchangeable proton or hydrogen atom ; unexpectedly, most of the CO<sub>2</sub> coproduct formed in the active site of CvFAP is shown to be converted catalytically in situ in 100 ns (transformation most likely into bicarbonate).

The article was accepted and published in Science in 2021. The link of the publication is the following: <https://www.science.org/doi/10.1126/science.abd5687>

## **Article 2:**

### **Mechanism and dynamics of fatty acid photodecarboxylase**



## RESEARCH ARTICLE SUMMARY

## PHOTOENZYMES

## Mechanism and dynamics of fatty acid photodecarboxylase

D. Sorigué, K. Hadjimetriou, S. Blangy, G. Gotthard, A. Bonvalet, N. Coquelle, P. Samire, A. Aleksandrov, L. Antonucci, A. Benachir, S. Boutet, M. Byrdin, M. Cammarata, S. Carbajo, S. Cuiné, R. B. Doak, L. Foucar, A. Gorel, M. Grünbein, E. Hartmann, R. Hienerwadel, M. Hilpert, M. Kloos, T. J. Lane, B. Légeret, P. Legrand, Y. Li-Beisson, S. L. Y. Moulin, D. Nurizzo, G. Peltier, G. Schirò, R. L. Shoeman, M. Sliwa, X. Solinas, B. Zhuang, T. R. M. Barends, J.-P. Colletier, M. Joffe, A. Royant, C. Berthomieu\*, M. Weik\*, T. Domratheva\*, K. Brettel, M. H. Vos\*, I. Schlichting\*, P. Arnoux\*, P. Müller\*, F. Beisson\*

**INTRODUCTION:** Photoenzymes are rare biocatalysts driven by absorption of a photon at each catalytic cycle; they inspire development of artificial photoenzymes with valuable activities. Fatty acid photodecarboxylase (FAP) is a natural photoenzyme that has potential applications in the bio-based production of hydrocarbons, yet its mechanism is far from fully understood.

**RATIONALE:** To elucidate the mechanism of FAP, we studied the wild-type (WT) enzyme from *Chlorella variabilis* (CvFAP) and variants with altered active-site residues using a wealth of techniques, including static and time-resolved crystallography and spectroscopy, as well as biochemical and computational approaches.

**RESULTS:** A 1.8-Å-resolution CvFAP x-ray crystal structure revealed a dense hydrogen-bonding network positioning the fatty acid carboxyl group in the vicinity of the flavin adenine dinucleotide (FAD) cofactor. Structures solved from free electron laser and low-dose synchrotron x-ray crystal data further highlighted an

unusual bent shape of the oxidized flavin chromophore, and showed that the bending angle (14°) did not change upon photon absorption (step 1) or throughout the photocycle. Calculations showed that bending substantially affected the energy levels of the flavin. Structural and spectroscopic analysis of WT and mutant proteins targeting two conserved active-site residues, R451 and C432, demonstrated that both residues were crucial for proper positioning of the substrate and water molecules and for oxidation of the fatty acid carboxylate by  $^1\text{FAD}^*$  (~300 ps in WT FAP) to form  $\text{FAD}^{\bullet-}$  (step 2). Time-resolved infrared spectroscopy demonstrated that decarboxylation occurred quasi-instantaneously upon this forward electron transfer, consistent with barrierless bond cleavage predicted by quantum chemistry calculations and with snapshots obtained by time-resolved crystallography. Transient absorption spectroscopy in  $\text{H}_2\text{O}$  and  $\text{D}_2\text{O}$  buffers indicated that back electron transfer from  $\text{FAD}^{\bullet-}$  was coupled to and limited by transfer of an exchangeable proton or hydrogen atom

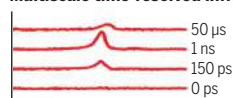
(step 3). Unexpectedly, concomitant with  $\text{FAD}^{\bullet-}$  reoxidation (to a red-shifted form  $\text{FAD}_{\text{RS}}$ ) in 100 ns, most of the  $\text{CO}_2$  product was converted, most likely into bicarbonate (as inferred from FTIR spectra of the cryotrapped  $\text{FAD}_{\text{RS}}$  intermediate). Calculations indicated that this catalytic transformation involved an active-site water molecule. Cryo-Fourier transform infrared spectroscopy studies suggested that bicarbonate formation (step 4) was preceded by deprotonation of an arginine residue (step 3). At room temperature, the remaining  $\text{CO}_2$  left the protein in 1.5  $\mu\text{s}$  (step 4'). The observation of residual electron density close to C432 in electron density maps derived from time-resolved and cryocrystallography data suggests that this residue may play a role in stabilizing  $\text{CO}_2$  and/or bicarbonate. Three routes for alkane formation were identified by quantum chemistry calculations; the one shown in the figure is favored by the ensemble of experimental data.

**CONCLUSION:** We provide a detailed and comprehensive characterization of light-driven hydrocarbon formation by FAP, which uses a remarkably complex mechanism including unique catalytic steps. We anticipate that our results will help to expand the green chemistry toolkit. ■

The list of author affiliations is available in the full article online.  
\*Corresponding author. Email: frederic.beisson@cea.fr (F.B.), pavel.muller@i2bc.paris-saclay.fr (P.M.), pascal.arnoux@cea.fr (P.A.), ilme.schlichting@mpimf-heidelberg.mpg.de (I.S.), marten.vos@polytechnique.edu (M.H.V.), tatjana.domratheva@mpimf-heidelberg.mpg.de (T.D.), martin.weik@ibs.fr (M.W.), catherine.berthomieu@cea.fr (C.B.)  
Cite this article as D. Sorigué et al., *Science* 372, eabd5687 (2021). DOI: 10.1126/science.abd5687

**READ THE FULL ARTICLE AT**  
<https://doi.org/10.1126/science.abd5687>

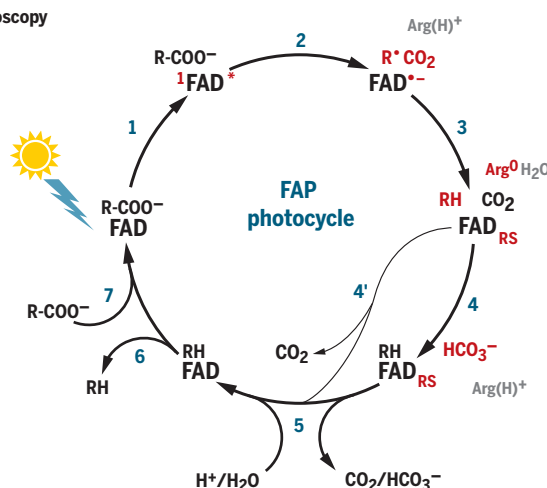
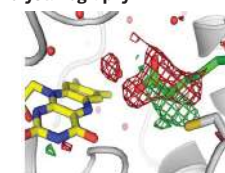
## Multiscale time-resolved infrared spectroscopy



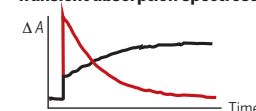
## Fourier transform infrared spectroscopy



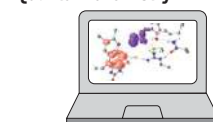
## Time-resolved serial femtosecond crystallography



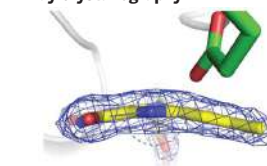
## Transient absorption spectroscopy



## Quantum chemistry



## X-ray crystallography



Elucidation of the FAP photocycle by combining spectroscopic, biochemical, crystallographic, and computational studies.

## RESEARCH ARTICLE

## PHOTOENZYMES

## Mechanism and dynamics of fatty acid photodecarboxylase

D. Sorigué<sup>1</sup>, K. Hadjidemetriou<sup>2</sup>, S. Blangy<sup>1</sup>, G. Gotthard<sup>3</sup>, A. Bonvalet<sup>4</sup>, N. Coquelle<sup>5</sup>, P. Samire<sup>1,6</sup>, A. Aleksandrov<sup>4</sup>, L. Antonucci<sup>4</sup>, A. Benachir<sup>4</sup>, S. Boutet<sup>7</sup>, M. Byrdin<sup>2</sup>, M. Cammarata<sup>8,†</sup>, S. Carbajo<sup>7</sup>, S. Cuiné<sup>1</sup>, R. B. Doak<sup>9</sup>, L. Foucar<sup>9</sup>, A. Gorel<sup>9</sup>, M. Grünbein<sup>9</sup>, E. Hartmann<sup>9</sup>, R. Hienerwadel<sup>1</sup>, M. Hilpert<sup>9</sup>, M. Kloos<sup>9,‡</sup>, T. J. Lane<sup>7</sup>, B. Légeret<sup>1</sup>, P. Legrand<sup>10</sup>, Y. Li-Beisson<sup>1</sup>, S. L. Y. Moulin<sup>1</sup>, D. Nurizzo<sup>3</sup>, G. Peltier<sup>1</sup>, G. Schirò<sup>2</sup>, R. L. Shoeman<sup>9</sup>, M. Sliwa<sup>11</sup>, X. Solinas<sup>4</sup>, B. Zhuang<sup>4,6</sup>, T. R. M. Barends<sup>9</sup>, J.-P. Colletier<sup>2</sup>, M. Joffre<sup>4</sup>, A. Royant<sup>2,3</sup>, C. Berthomieu<sup>1,\*</sup>, M. Weik<sup>2,\*</sup>, T. Domratheva<sup>9,12,\*</sup>, K. Brettel<sup>6</sup>, M. H. Vos<sup>4,\*</sup>, I. Schlichting<sup>9,\*</sup>, P. Arnoux<sup>1,\*</sup>, P. Müller<sup>6,\*</sup>, F. Beisson<sup>1,\*</sup>

Fatty acid photodecarboxylase (FAP) is a photoenzyme with potential green chemistry applications. By combining static, time-resolved, and cryotrapping spectroscopy and crystallography as well as computation, we characterized *Chlorella variabilis* FAP reaction intermediates on time scales from subpicoseconds to milliseconds. High-resolution crystal structures from synchrotron and free electron laser x-ray sources highlighted an unusual bent shape of the oxidized flavin chromophore. We demonstrate that decarboxylation occurs directly upon reduction of the excited flavin by the fatty acid substrate. Along with flavin reoxidation by the alkyl radical intermediate, a major fraction of the cleaved carbon dioxide unexpectedly transformed in 100 nanoseconds, most likely into bicarbonate. This reaction is orders of magnitude faster than in solution. Two strictly conserved residues, R451 and C432, are essential for substrate stabilization and functional charge transfer.

By far most enzymatic reactions in living cells are thermally activated, whereas reactions driven by light are much less common (1). Apart from photosynthetic reaction centers, three natural photoenzymes have been identified to date: DNA photolyases, which are involved in the repair of UV-damaged DNA (2); light-dependent prochlorophyllide oxidoreductase (LPOR), which is required for the maturation of chlorophyll (3); and the recently discovered fatty acid photodecarboxylase (FAP, EC 4.1.1.106), which converts fatty acids to hydrocarbons and CO<sub>2</sub> (4).

Photoenzymes require a photon for each turnover of a substrate molecule and offer the possibility to trigger and monitor catalytic steps and associated structural changes on very short time scales that are generally not accessible for thermally activated enzymes (5). Thus, over the past 20 years, the photochemical mechanism has been studied in detail for DNA photolyases (6–8) and has started to be identified for LPOR (9–13). Because of its readily available substrate, FAP may become a model of choice to understand catalytic steps that occur in enzymology.

Light-driven enzymes are also interesting for practical applications. Increasing our understanding and the repertoire of photoenzymatic mechanisms may help in the design of catalysts performing new reactions (14, 15) or in the development of new light-controlled proteins for optogenetics (16). FAP has potential biotechnological applications in green chemistry because hydrocarbons are important as cosmetics emollients, chemical synthons, solvents, and fuels (17–19). FAP complements routes previously identified for bio-based synthesis of hydrocarbons (20) by providing a one-step light-driven pathway from fatty acids. Rational design approaches have recently allowed improvement of the efficiency of FAP on high-value functionalized carboxylic acids (21) and on short-chain fatty acids to produce liquefied petroleum gas (20). Understanding the reaction mechanism of FAP in detail is thus of utmost importance, both from a fundamental research and an application point of view.

FAP is an algae-specific enzyme from the glucose-methanol-choline (GMC) oxidoreductase family harboring a flavin adenine dinucleotide (FAD) cofactor. It allows the decarboxylation

of C16–C18 free (nonesterified) fatty acids to the corresponding *n*-alka(n)es (4). These hydrocarbon products are mostly located in chloroplast thylakoids, but their exact role is still unknown (22, 23). The initial spectroscopic characterization of FAP, which was based on monitoring the electronic state of the flavin after excitation by a laser flash, led to the first model of the FAP photocycle (4). The cycle starts with the quenching of the singlet excited state by forward electron transfer (fET) from bound fatty acid R-CO<sub>2</sub><sup>−</sup> in ~300 ps (with a quantum yield >80%), forming a flavin anion radical FAD<sup>•−</sup> and a fatty acid radical R-CO<sub>2</sub><sup>•</sup>. The latter decarboxylates, yielding an alkyl radical R<sup>•</sup> and CO<sub>2</sub>. FAD<sup>•−</sup> is reoxidized in ~100 ns by back electron transfer (bET), which ultimately provides the electron for the reduction of R<sup>•</sup> to the alkane RH. FAD<sup>•−</sup> reoxidation results in a transiently red-shifted flavin state FAD<sub>RS</sub> that reverts to the initial state in 4 ms.

Despite this insight, several open questions remained, including: Which structural features of the FAP active site promote substrate stabilization and favor the fET? Is decarboxylation instantaneous upon this fET step or is it slowed by an activation barrier? Does conversion of R<sup>•</sup> to the alkane RH occur by bET from FAD<sup>•−</sup> coupled to a proton transfer (PT/PCET) or by hydrogen atom transfer (HAT) from a nearby amino acid (4, 24)? What is the origin of the proton or the hydrogen atom? Here, we report a high-resolution structure of FAP and characterize key steps along the FAP photocycle using a wealth of static and time-resolved crystallographic and spectroscopic techniques, as well as computational approaches. Our detailed characterization of FAP reveals unforeseen mechanistic complexity.

## High-resolution structure of CvFAP

Crystals of FAP from *C. variabilis* NC64A (CvFAP), diffracting x-rays to high resolution and without the twinning fault reported earlier (4), were obtained upon removal of the N-terminal helix involved in crystal packing (residues 61 to 76) in native CvFAP. The resulting structure, solved at 1.8-Å resolution (table S1), now provides a detailed view of the active-site architecture (Fig. 1, A to D, and fig. S1). Although no substrate was added during protein purification and crystallization, two C18 fatty acids copurified with FAP were clearly identified, one in the active site and the other

<sup>1</sup>Aix-Marseille University, CEA, CNRS, Institute of Biosciences and Biotechnologies, BIAM Cadarache, 13108 Saint-Paul-lez-Durance, France. <sup>2</sup>Université Grenoble Alpes, CEA, CNRS, Institut de Biologie Structurale, 38000 Grenoble, France. <sup>3</sup>European Synchrotron Radiation Facility, 38043 Grenoble, France. <sup>4</sup>LOB, CNRS, INSERM, Ecole Polytechnique, Institut Polytechnique de Paris, 91128 Palaiseau, France. <sup>5</sup>Large-Scale Structures Group, Institut Laue Langevin, 38042 Grenoble Cedex 9, France. <sup>6</sup>Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198 Gif-sur-Yvette, France. <sup>7</sup>Linac Coherent Light Source (LCLS), SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA. <sup>8</sup>Department of Physics, UMR UR1-CNRS 6251, University of Rennes 1, F-Rennes, France. <sup>9</sup>Max-Planck-Institut für medizinische Forschung, Jahnstrasse 29, 69120 Heidelberg, Germany. <sup>10</sup>Synchrotron SOLEIL, L'Orme des Merisiers Saint-Aubin, BP 48, 91192 Gif-sur-Yvette, France. <sup>11</sup>Univ. Lille, CNRS, UMR 8516, LASIRE, Laboratoire de Spectroscopie pour les Interactions, la Réactivité et l'Environnement, 59000 Lille, France. <sup>12</sup>Department of Chemistry, Lomonosov Moscow State University, Moscow 119991, Russia.

<sup>†</sup>Present address: European Synchrotron Radiation Facility (ESRF), F-38043 Grenoble, France. <sup>‡</sup>Present address: Sample Environment and Characterization European XFEL, 22869 Schenefeld, Germany.

\*Corresponding author. Email: frederic.beisson@cea.fr (F.B.); pavel.muller@l2bc.paris-saclay.fr (P.M.); pascal.arnoux@cea.fr (P.A.); ilme.schlichting@mpimf-heidelberg.mpg.de (I.S.); marten.vos@polytechnique.edu (M.H.V.); tatjana.domratheva@mpimf-heidelberg.mpg.de (T.D.); martin.weik@ibs.fr (M.W.); catherine.berthomieu@cea.fr (C.B.)



on the surface of the protein. The latter is stabilized by hydrophobic interactions with nonpolar side chains (L427, Y419, I126, I416, and L420), as well as the aliphatic parts of R132 and R122. In the active site, the carboxyl group of the fatty acid is stabilized by hydrogen bonds with water molecules (Wat1 and Wat2) and the side chains of R451 and N575 (Fig. 1D). The fatty acid substrate is also stabilized on the dimethylbenzene side of FAD, in contrast to other GMC oxidoreductase enzymes, which have their substrate stabilized near the N5 atom of FAD (fig. S2).

#### Conformation of oxidized FAD in CvFAP

In the high-resolution dark-state structure of FAP obtained from synchrotron data “100 K dark,” the isoalloxazine ring of the FAD cofactor was found to be bent, with the dihedral angle C4-N5-N10-C9 (butterfly bending angle) deviating by  $17.4^\circ$  from planarity (Fig. 1E and fig. S2A). Such bending is usually interpreted as being caused by x-ray photoreduction con-

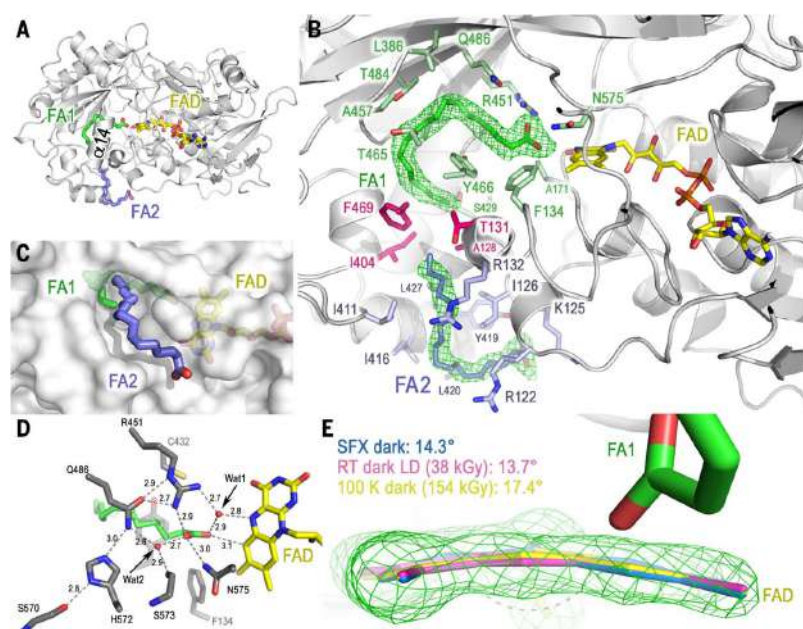
verting supposedly planar oxidized flavin to the bent reduced form (8, 25, 26). For FAP, however, in crystallo ultraviolet-visible (UV-Vis) absorption microspectrophotometry (fig. S3), low-dose crystallography, (fig. S4A and table S1), Raman microspectrophotometry (fig. S4, B and C), and molecular dynamics simulations (fig. S5, A and B) using a recent flavin force field (27) indicated that the FAD cofactor is bent in its oxidized form (supplementary text S1). To obtain a definite answer concerning the conformation of the oxidized FAD in FAP, we performed room temperature (RT) serial femtosecond crystallography (SFX) at an x-ray free-electron laser (XFEL; see below for details), which allows collecting essentially radiation damage-free diffraction data (28, 29). This SFX dark-state structure of FAP (“SFX dark”), solved at 2.0-Å resolution (table S2), features a FAD with a similar bending angle ( $14.3^\circ$ ) as observed in the synchrotron 100 K dark and the “RT dark low-dose” structures (Fig. 1E and figs. S4, S6, and S7), supporting

the notion that in FAP, the FAD cofactor is in a bent conformation in the oxidized state. To our knowledge, a butterfly bent conformation of the oxidized flavin has not been firmly established for any other flavoprotein. In fact, bent conformations were either not discussed in the literature or were attributed to x-ray-induced flavin reduction. Future radiation-damage-free structures of oxidized flavoproteins should reveal whether the bending is a feature specific to FAP.

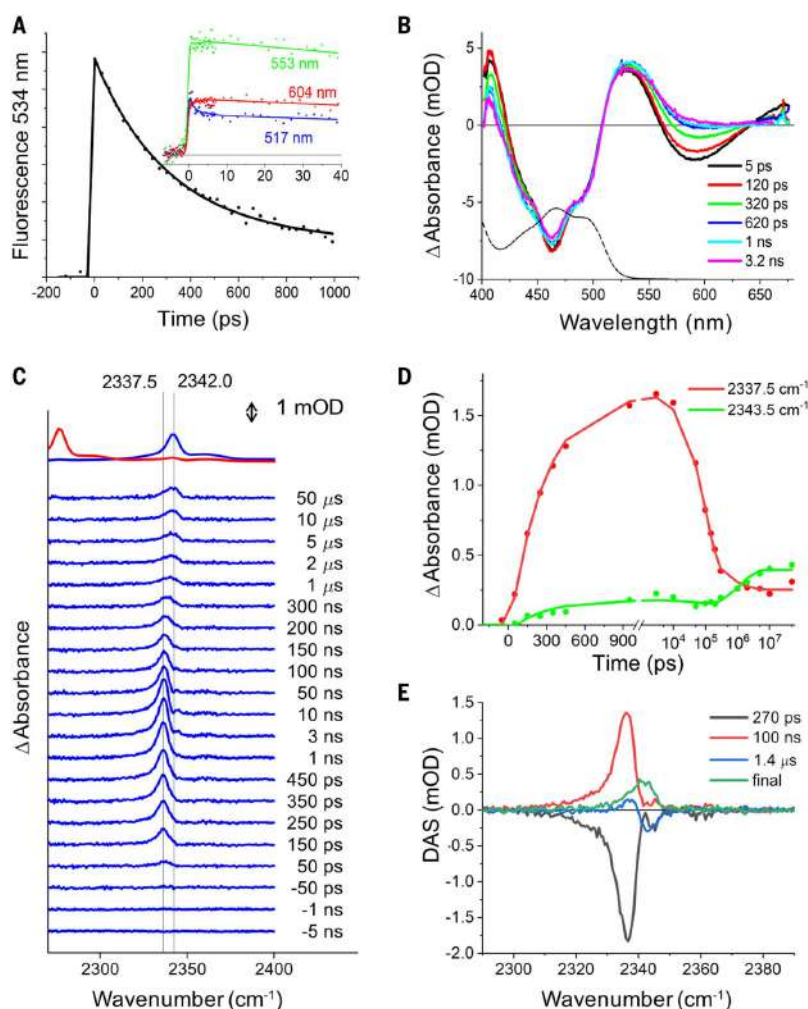
#### New insights into the CvFAP photocycle by time-resolved spectroscopy

Our previous single-shot fluorescence and transient absorption experiments in the presence of substrate, with 100 ps and 10 ns temporal resolution, respectively, showed a decay of the singlet excited flavin ( $^1\text{FAD}^*$ ) in  $\sim 300$  ps and formation of  $\text{FAD}^{\bullet-}$  within 10 ns (4). Here, we extended the fluorescence and visible absorption experiments to the ultrafast time scale with 100-fs-resolution pump-probe spectroscopy (Fig. 2, A and B, and fig. S8A), showing that no faster  $^1\text{FAD}^*$  decay phases occurred [only an  $\sim 2$ -ps thermal relaxation phase was observed in the fluorescence as in other flavoproteins (30)] and that, as predicted (4),  $\text{FAD}^{\bullet-}$  is formed concomitantly with  $^1\text{FAD}^*$  decay (fig. S8B).

We investigated whether the carboxylate of the fatty acid is cleaved off concomitantly with electron transfer from the fatty acid to  $^1\text{FAD}^*$  ( $\sim 300$  ps), after this oxidation, or accompanying the bET from flavin ( $\sim 100$  ns). Light-induced Fourier transform infrared (FTIR) difference spectroscopy performed at 298 K showed a  $\text{CO}_2$  band at  $2342\text{ cm}^{-1}$  (Fig. 2C), corresponding to  $\text{CO}_2$  in solution (37). The kinetics of  $\text{CO}_2$  formation in FAP were monitored by picosecond to microsecond time-resolved infrared spectroscopy (TR-IR) in a frequency range centered around  $2350\text{ cm}^{-1}$  and arbitrary detuning asynchronous optical sampling (ADASOPS) (32, 33). This experiment showed the appearance of a  $\text{CO}_2$  absorption band centered at  $2337.5\text{ cm}^{-1}$  with a time constant of  $\sim 270$  ps (Fig. 2, C to E, and fig. S9). We conclude that decarboxylation is rate-limited by electron transfer from substrate to  $^1\text{FAD}^*$ , occurring in  $\sim 300$  ps. The initial  $\text{CO}_2$  band frequency is  $\sim 5\text{ cm}^{-1}$  down-shifted with respect to that of  $^{12}\text{CO}_2$  in aqueous solution, a finding that we assign to the protein environment. Subsequently, the  $\text{CO}_2$  signal diminishes about fourfold with a time constant of 100 ns without changing much in shape, followed by an upshift toward  $2342\text{ cm}^{-1}$ , a process fitted with a time constant of  $\sim 1.5\text{ }\mu\text{s}$ . The latter process likely reflects migration of  $\text{CO}_2$  toward the solvent. The data indicate that the 100-ns process implies transformation of  $\sim 75\%$  of the initially formed  $\text{CO}_2$  within the protein into another molecule, possibly bicarbonate, concomitant with flavin reoxidation to the red-shifted form  $\text{FAD}_{\text{RS}}$  (see next



**Fig. 1. High-resolution crystal structures of CvFAP.** (A) Structure of CvFAP determined from synchrotron data at 100 K (100 K dark), including the FAD cofactor and two C18 fatty acid substrates (FA1 and FA2). (B) Binding of the two substrate molecules. The “omit” electron density map ( $5.0\text{ }\sigma$  contour level) is shown as a green mesh, and amino acid side chains in a radius of 4 Å around the binding site are shown as sticks (green: active site, blue: secondary binding site, purple: between sites). (C) Position of the peripheral substrate (FA2) partly obstructing the channel leading to the active site tunnel. (D) Close-up view of the catalytic site showing the water molecules Wat1 and Wat2 and the interactions with FAD (yellow sticks) and the substrate in the active site FA1 (green sticks). Distances are indicated in angstroms. The shortest distance between the substrate and the FAD cofactor (carboxylate O1-isoalloxazine C6) is 3.1 Å. The tail of the peripheral substrate points toward the entrance of a tunnel leading to the active site that is lined by A128, T131, I404, and F469. (E) Superposition of the FAD isoalloxazine rings from the SFX dark structure (blue; molecule A) and the synchrotron structures (pink: RT dark low-dose; yellow: 100 K dark). The SFX  $F_{\text{obs}} - F_{\text{calc}}$  omit map at  $3\text{ }\sigma$  (green) is overlaid, and the FAD bending angles are indicated.



**Fig. 2. Time-resolved infrared and UV-Vis spectroscopies of CvFAP.** (A) Fluorescence kinetics reflecting  $^1\text{FAD}^*$  decay. The solid line is a fit with time constants of 300 ps (85%) and 5 ns [the latter time constant was imposed according to (4)]. The inset shows an additional  $\sim 2$  ps wavelength-dependent phase reflecting red shifting of the fluorescence spectrum caused by excited-state relaxation. (B) Transient visible absorption spectra at different delay times in the same time domain. The negative absorption features reflect bleaching of the  $\text{FAD}_{\text{ox}}$  resting state ( $<525$  nm, the black dashed line is the ground state absorption spectrum) and stimulated emission (550 to 650 nm). Full transient fluorescence spectra and global analysis of the transient absorption spectra are shown in fig. S8. (C) Transient infrared spectra in the  $\text{CO}_2$  spectral region on the picosecond-microsecond time scale. The vertical lines are guides for the eye and correspond to the maximum of released  $\text{CO}_2$  in the protein and to the known maximum for  $\text{CO}_2$  in aqueous solution,  $2342\text{ cm}^{-1}$  (31). The upper traces correspond to independent steady-state 298 K light-induced FTIR difference spectra with  $^{12}\text{C}$ -palmitate and  $^{13}\text{C}$ -palmitate substrates (blue trace and red trace, respectively). (D) Kinetics at frequencies close to the initial and final maxima of released  $\text{CO}_2$ . The lines are the result of a global fit. The time scale is linear up to 1 ns and logarithmic thereafter. (E) Decay-associated spectra (DAS) corresponding to a global fit of the data with three exponential phases (time constants of the fit are indicated) and a constant phase.

section on the red-shifted intermediate). This process was not foreseen in our previously proposed reaction scheme (4). Whereas our present data covering six orders of magnitude in time can be reasonably well described with

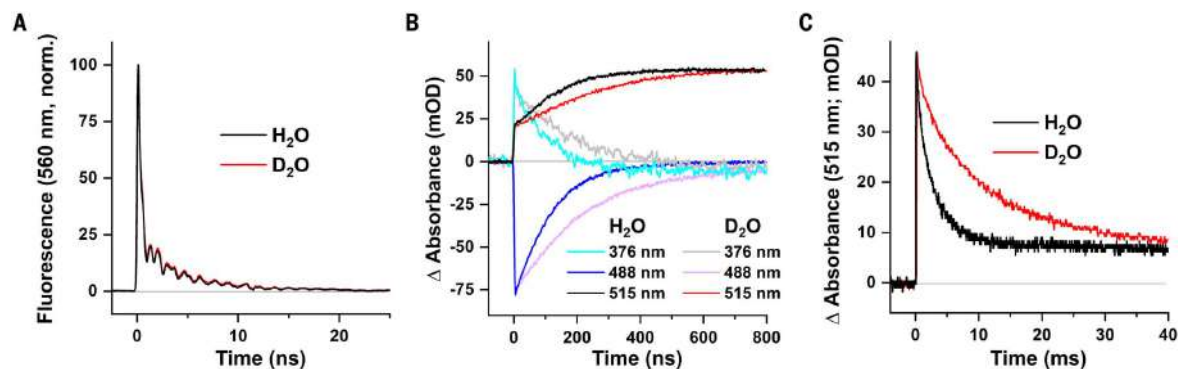
these three exponential processes and a constant phase (Fig. 2E), it is possible that an additional phase of  $\text{CO}_2$  release into the solvent occurs on a time scale exceeding 50  $\mu\text{s}$ , the temporal window of our present experiments.

To determine whether any of the reaction steps are coupled to or directly reflect PT or HAT, we compared the kinetics of all steps observable by continuous-probe time-resolved fluorescence or transient absorption spectroscopy in  $\text{H}_2\text{O}$  and  $\text{D}_2\text{O}$  buffers. The biphasic fluorescence signal showing the decay of  $^1\text{FAD}^*$  [caused by quenching by ET from the substrate (300 ps phase;  $>80\%$ ) and by the intersystem-crossing to a nonreactive triplet state ( $\sim 6.5$  ns phase;  $<20\%$ )] was not visibly affected by the isotope exchange (Fig. 3A), as expected for reactions that are not substantially coupled to movements of exchangeable hydrogen species [e.g., back-ET from  $\text{Q}_\text{A}^-$  to  $\text{P680}^+$  in photosystem II (34)]. By contrast, the kinetics of  $\text{FAD}^*$  reoxidation to  $\text{FAD}_{\text{RS}}$  observed by transient absorption spectroscopy at three characteristic wavelengths slowed down by a factor of  $\sim 2$ , from  $\sim 100$  ns in  $\text{H}_2\text{O}$  to  $\sim 200$  ns in  $\text{D}_2\text{O}$  (Fig. 3B), suggesting that this bET step is coupled to and limited by transfer of an exchangeable proton or hydrogen atom. Change of buffer pH in the interval between 7.5 and 9.1 had no effect on the kinetics of this step (fig. S10A), indicating that the proton or hydrogen atom donor involved in the reaction has a  $\text{pK}_\text{a}$  value  $>9.1$  (where  $K_\text{a}$  is the acid dissociation constant). The significantly lower kinetic isotope effect (KIE) of 1.2 reported by Heyes *et al.* (24) might be due to insufficient time-resolution of their experiment (see supplementary text S4).

The last step observable by transient absorption spectroscopy was the disappearance of the transient red shift of reoxidized FAD occurring in a few milliseconds. This process was previously (4) assigned to reprotonation of  $\text{X}^-$ , the conjugate base of  $\text{XH}$ , an unidentified proton donor to the alkyl radical. Kinetics of this process also slowed down from  $\sim 3$  to  $\sim 10$  ms when  $\text{H}_2\text{O}$  in the buffer was replaced by  $\text{D}_2\text{O}$  (Fig. 3C). Consistent with this observation and the recent report of a pH increase (in an unbuffered solution) associated with this step (24), this process apparently reflects a proton transfer from bulk solvent. Again, change of buffer pH in the interval between 7.5 and 9.1 had no effect on the kinetics of this step (fig. S10B), nor did the consumption of the native substrate (fig. S10C). Time-resolved spectroscopic findings are summarized in fig. S11.

#### Photodecarboxylation and characterization of the red-shifted intermediate at cryogenic temperatures

Our time-resolved spectroscopic results predicted that a fraction of the  $\text{CO}_2$  product is present in the active site in the red-shifted photoproduct intermediate. For a detailed characterization by static methods, we tried to stabilize this intermediate using cryotrapping. UV-Vis absorption spectra of CvFAP crystals



**Fig. 3. Effects of H<sub>2</sub>O-D<sub>2</sub>O exchange on time-resolved fluorescence and UV-Vis spectroscopy of CvFAP.** (A) Normalized signals of time-resolved fluorescence of <sup>1</sup>FAD\* at 560 nm in H<sub>2</sub>O and D<sub>2</sub>O buffers. (B) Effect of H<sub>2</sub>O-to-D<sub>2</sub>O exchange on the kinetics of FAD<sup>•-</sup> reoxidation to FAD<sub>RS</sub>, followed by transient absorption spectroscopy at selected wavelengths on a submicrosecond time scale. (C) Effect of H<sub>2</sub>O to D<sub>2</sub>O buffer exchange on the decay of FAD<sub>RS</sub> measured at 515 nm on the millisecond time scale.

and solutions exposed to increasing amounts of blue light at 100 K indeed revealed the gradual conversion into a red-shifted form similar to that previously observed at RT [(4) and fig. S12]. These conditions were used to further characterize the red-shifted intermediate using light-induced FTIR difference spectroscopy. To identify IR modes of the substrate and products of the reaction, we replaced the native substrate with 1-<sup>12</sup>C or 1-<sup>13</sup>C palmitate. The identification of characteristic <sup>13</sup>C-sensitive IR bands of carboxylate in the FTIR difference spectra at 1541 and 1391 cm<sup>-1</sup> shows that the substrate initially was in the deprotonated form (Fig. 4A, i and ii; details supporting all IR bands assignments mentioned below are given in supplementary text S2). A peak at 2340 cm<sup>-1</sup> (Fig. 4A, v) was assigned to formation of CO<sub>2</sub> from 1-<sup>12</sup>C-palmitate and at 2274 cm<sup>-1</sup> from 1-<sup>13</sup>C-palmitate. When comparing FTIR spectra recorded with FAP samples containing 1-<sup>12</sup>C- and 1-<sup>13</sup>C-palmitate, small bands were also observed at 1356 to 1335 cm<sup>-1</sup> and 1312 cm<sup>-1</sup> (Fig. 4A, iii), which is indicative of the formation of trace amounts of <sup>12</sup>C and <sup>13</sup>C bicarbonate, respectively. We repeated the FTIR experiments at 150 K. At this temperature, the CO<sub>2</sub> band was small (Fig. 4A, v), whereas large positive bands at 1646 (1614) and 1352 (1318) cm<sup>-1</sup> were detected that could be unambiguously assigned to IR modes of <sup>12</sup>C- (<sup>13</sup>C-) bicarbonate (Fig. 4A, iv). The FTIR data thus demonstrate the buildup of bicarbonate at 150 K and indicate that its formation at 100 K is limited by an energy barrier.

The structure derived from a CvFAP crystal exposed to 470 nm light at 100 K ("100 K light") featured FAP trapped in a FAD<sub>RS</sub> state, as shown by in crystallo microspectrophotometry (fig. S12). The difference electron density map between 100 K light and 100 K dark displays strong and significant peaks only

around the substrate and FAD cofactor. These peaks can be fitted with a CO<sub>2</sub> molecule in addition to the alkyl chain stabilized in the active site (Fig. 4, B and C, and fig. S13, A to C). The positive difference electron density along the aliphatic tail of the substrate (from C13 to C15) and the 8° rotation of the side chain of Y466 indicates a correlated motion consistent with the strong electronic coupling observed by quantum chemistry between the substrate and Y466. To reflect the optimum pH of FAP in solution [pH 8.5 (4)] and to overcome thermal activation barriers, a FAP crystal was soaked at pH 8.5 and illuminated at 150 K ("150 K light"). After modeling alkane and a CO<sub>2</sub> molecule, the residual  $F_o - F_c$  electron density map showed a positive density with a triangular shape close to C432, which we tentatively attributed to bicarbonate with 30% occupancy (Fig. 4D and fig. S14).

#### Time-resolved serial femtosecond crystallography of CvFAP

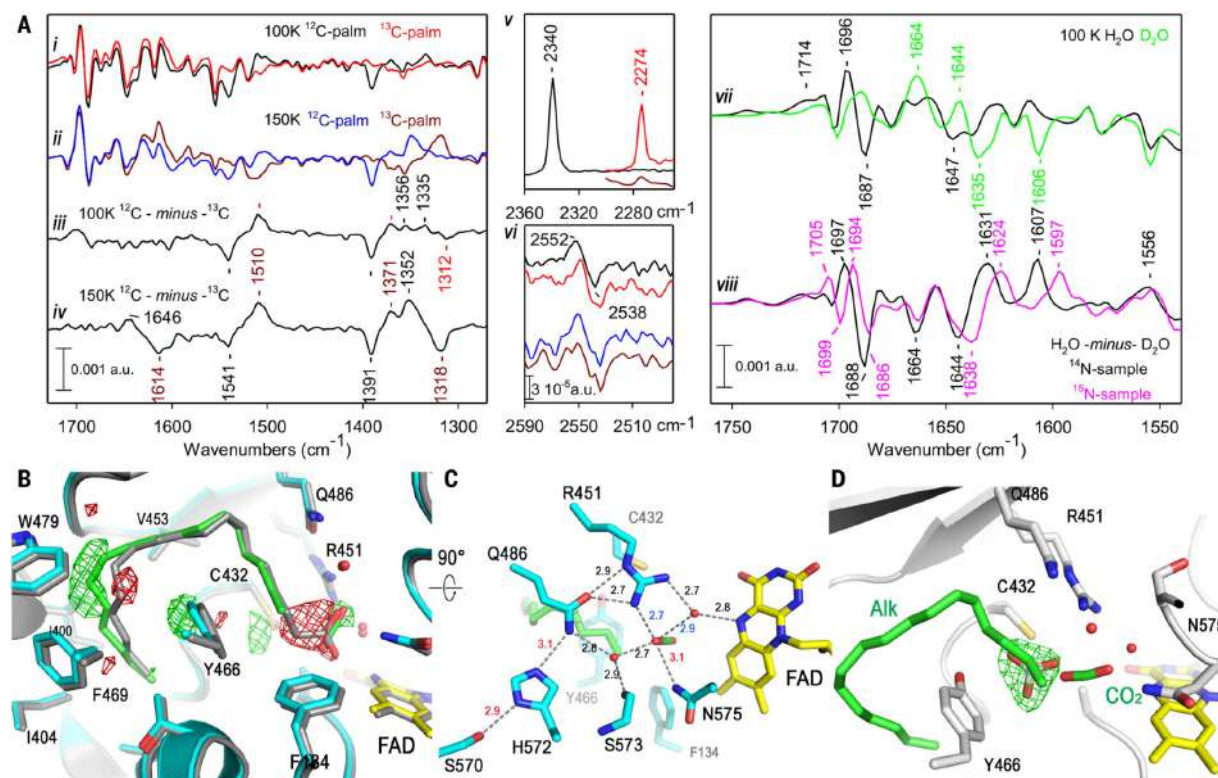
We investigated the structural changes occurring in FAP after photoexcitation at RT by a time-resolved SFX (TR-SFX) experiment (35) using a pump-probe scheme. Time-resolved fluorescence spectroscopy on FAP microcrystals established that the kinetics of fET are the same as in solution (supplementary text S5 and fig. S15). For TR-SFX at the Linac Coherent Light Source (LCLS), FAP microcrystals were photoexcited by picosecond 400-nm pump pulses and probed by femtosecond XFEL pulses after pump-probe delays of 20 ps, 900 ps, 300 ns, and 2 μs to cover time scales on which FAD reduction and FAD<sup>•-</sup> reoxidation occur. The SFX dark state structure mentioned above was determined from the data in the absence of a pump laser pulse. Structural changes after photoexcitation were visualized as positive and negative peaks in difference Fourier electron density maps calculated between the

light and dark datasets ( $F_{\text{obs}}^{\text{light}, \Delta t} - F_{\text{obs}}^{\text{dark}}$ , Fig. 5A) at 2.2-Å resolution. The most prominent difference electron density peaks at all four time points were at the active site, with the highest negative peak (-5.5 to -10.1 σ, depending on the time point) at the position of the carboxyl group of the substrate, showing that light-induced decarboxylation occurred (Fig. 5A). At 900 ps, decarboxylation had occurred to a considerable extent, consistent with the 270-ps time constant determined by multiscale time-resolved IR spectroscopy (Fig. 2, C to E). At 300 ns and 2 μs, a strong negative peak was observed at Wat1 (-5.8, and -6.2 σ, respectively; Fig. 5A), but not on Wat2. We note the absence of positive difference electron density peaks associated with the photo-dissociated CO<sub>2</sub> in the vicinity of the substrate carboxyl group. A possible reason could be the small initial displacement of cleaved CO<sub>2</sub> relative to its position in the fatty acid, consistent with the structure determined based on an illuminated cryocooled crystal 100 K light (see fig. S13C for comparison). It is conceivable that the positive difference densities close to C432 in the 300-ns and 2-μs datasets (Fig. 5A) correspond to the feature(s) observed in the data obtained from cryocooled crystals illuminated at 150 K (Figs. 5B and 4D), tentatively assigned to a bicarbonate (fig. S14). In the SFX data, attempts to fit unambiguously this positive difference density with a bicarbonate, a CO<sub>2</sub> molecule, or a mixture thereof remained unsatisfactory. The absence of significant difference electron density peaks at the FAD at all four time points suggests that the isoalloxazine ring does not undergo significant light-induced conformational changes.

#### Quantum chemistry study of CvFAP

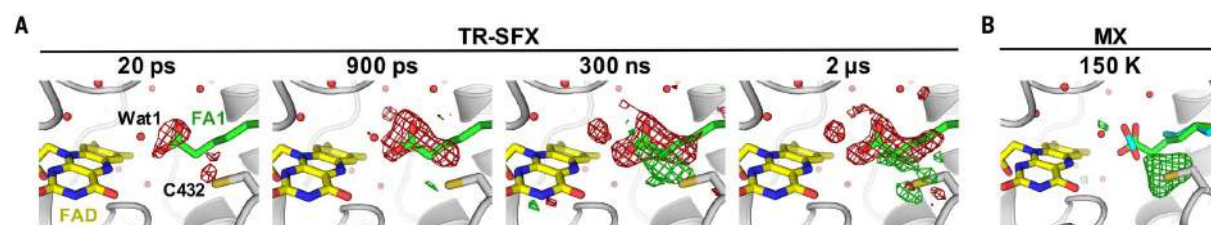
We performed quantum chemistry calculations for the decarboxylation reaction in the active site of FAP, considering electronic





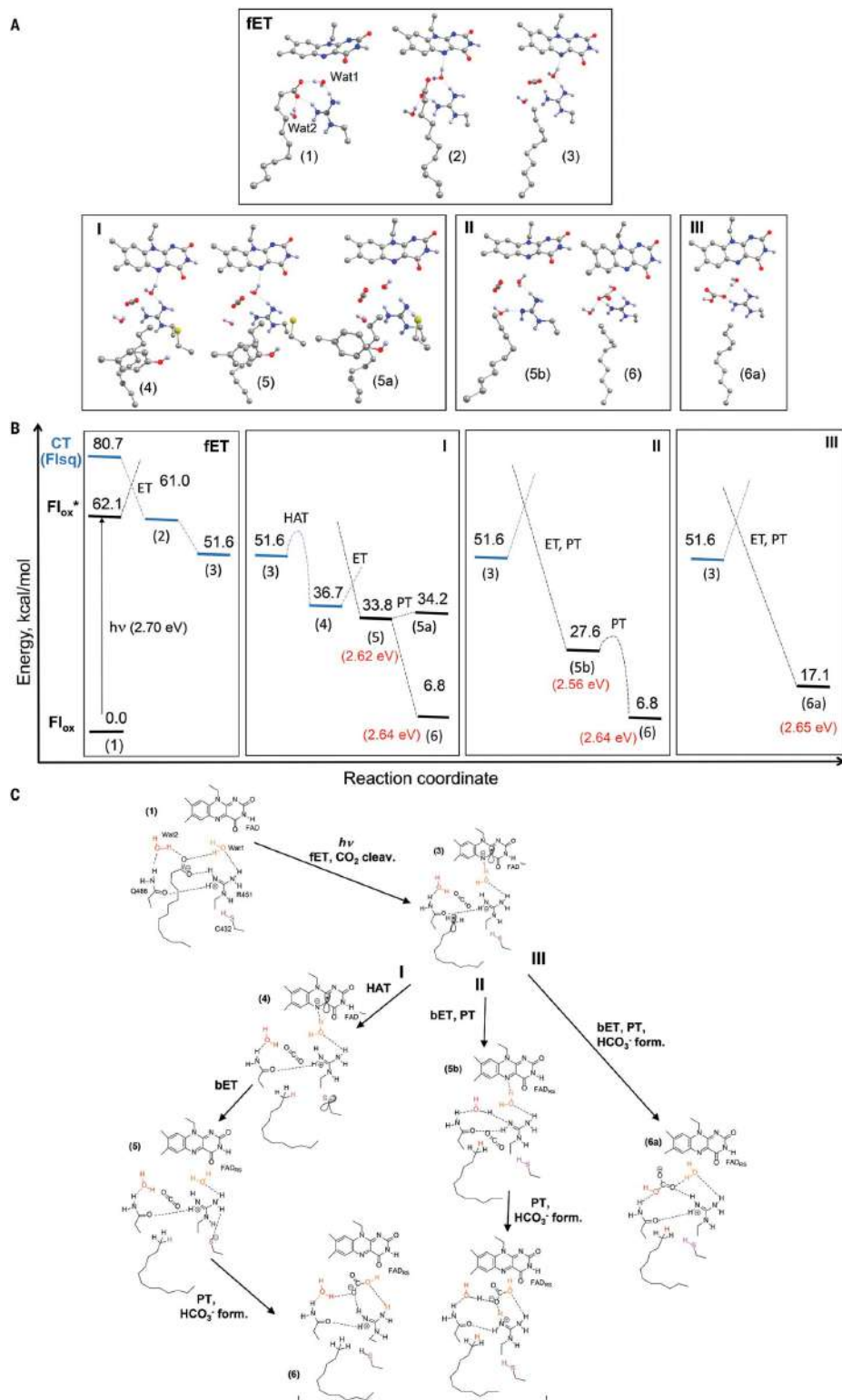
**Fig. 4. Characterization of the CvFAP red-shifted intermediate at cryogenic temperatures.** (A) Light-induced FTIR difference spectra recorded from FAP solutions (i) and (ii) and corresponding  $^{12}\text{C}$ -minus- $^{13}\text{C}$  difference spectra (iii) and (iv). The 2360 to 2260  $\text{cm}^{-1}$  and 2590 to 2490  $\text{cm}^{-1}$  regions of the spectra in (i) and (ii) are shown in (v) and (vi). For the spectrum in black, the FTIR spectrometer was continuously purged with  $\text{N}_2$  to avoid contamination by gaseous  $\text{CO}_2$  that might originate from a dry air purge. This was not the case for spectra recorded from FAP samples with  $1\text{-}^{13}\text{C}$ -palmitate; these were cut off above 2300  $\text{cm}^{-1}$  to avoid bands from gaseous  $\text{CO}_2$ . (vii) Overlay of the spectra recorded at 100 K from FAP with  $^{12}\text{C}$ -palmitate in  $\text{H}_2\text{O}$  (black) and in  $\text{D}_2\text{O}$  (green). (viii)  $\text{H}_2\text{O} - \text{D}_2\text{O}$  difference spectra calculated for  $^{14}\text{N}$ - and  $^{15}\text{N}$ -labeled FAP samples. (B) Experimental difference

density map ( $F_{\text{light}} - F_{\text{dark}}$ , 100 K data) contoured at  $\pm 4\sigma$  around the active-site substrate superimposed on the refined structures of the dark state (gray) and the red-shifted form (cyan, with FAD in yellow and alkane and  $\text{CO}_2$  in green). The cleavage of the C1-C2 bond is clearly visible. (C) Details of the active site of FAP<sub>RS</sub> formed upon illumination (UV-Vis spectra in fig. S12). All distances shorter than 3.2 Å are labeled, except the fatty acid O1-FAD C6 distance, which increases from 3.1 Å in the dark to 3.6 Å upon illumination (in  $\text{CO}_2$ ). (D)  $F_0 - F_c$  electron density omit map ( $3.5\sigma$ ) derived from crystals kept at 150 K and pH 8.5, showing positive difference electron density next to C432, consistent with a bicarbonate. Its potential interactions with the environment are shown in fig. S14. Uncleaved fatty acid (~30% occupancy) is omitted for clarity.



**Fig. 5. Time-dependent changes in the CvFAP active site followed by TR-SFX.** (A) Locally averaged  $q$ -weighted difference Fourier maps calculated between the SFX light and dark datasets ( $F_{\text{obs}}^{\text{light}, \Delta t} - F_{\text{obs}}^{\text{dark}}$ , with  $\Delta t = 20$  ps, 900 ps, 300 ns, 2  $\mu\text{s}$ ). The 2.2-Å resolution maps are shown at  $+4\sigma$  (green) and  $-4\sigma$  (red). The SFX dark-state model of molecule B is overlaid, with FAD in yellow, fatty acid in green, and the protein in gray. The  $q$ -weighted difference

maps for molecules A and B, i.e., without local averaging, are shown in fig. S31. (B)  $F_{\text{obs}} - F_{\text{calc}}$  electron density (contoured at  $3.5\sigma$ ) of Fig. 3E, shown in a different orientation, which features unmodeled positive electron density next to C432 that is reminiscent of a bicarbonate (Fig. 4D and fig. S14, see text), in a location similar to where positive difference density is present in the time-resolved maps at 300 ns and 2  $\mu\text{s}$ .



**Fig. 6. Plausible CvFAP photoreaction pathways according to quantum chemistry calculations.** (A) Structural changes accompanying decarboxylation and alkane formation in the FAP active site. Flavin butterfly-bending angle is 15° and 19° in the oxidized and semiquinone states, respectively. For clarity, only a small part of the active-site model is shown. The complete model is presented in fig. S16. (B) Energies of the reactants, intermediates, and products indicated in (A). Common to all pathways is fET mediated by Wat1 rearrangement (1) → (2) followed by CO<sub>2</sub> cleavage (2) → (3). Alkane formation occurs either

through HAT (pathway I) or PCET (pathways II and III). Pathway I consist of alkane formation through HAT from C432 (3) → (4), flavin reoxidation by the C432 radical (4) → (5), and C432 reprotonation and bicarbonate formation (5) → (6). Pathway II involves alkane formation through bET coupled to PT from Wat2 and subsequent R451 deprotonation (3) → (5b). Bicarbonate formation from CO<sub>2</sub> and Wat1 recovers protonated R451 (5b) → (6). In pathway III, bET is coupled to Wat2 deprotonation and formation of bicarbonate (3) → (6a). (C) Chemical scheme detailing intermediate states of fET and pathways I to III.

states relevant for flavin photoexcitation, electron transfer, decarboxylation, and alkane formation, as well as the effects of flavin bending (see supplementary text S6). To perform the reaction pathway calculations, a large active-site model (consisting of 272 atoms; fig. S16) was prepared using the coordinates of the high-resolution crystal structure of the dark state (Fig. 1D). Unexpectedly, computations indicated multiple routes to form an alkane in the FAP active site, as summarized in Fig. 6.

Common to all routes, photoexcitation triggered charge transfer (CT) from the substrate to the flavin. Interactions with two water molecules stabilized the anionic carboxylate substrate (1) (Fig. 6A). Removing these water molecules from the active site or rearranging Wat1 to form a hydrogen bond with flavin substantially decreased the energy of the CT state, i.e., of the flavin radical anion and the fatty acid radical (table S6). After Wat1 rearrangement, the CT energy was 15 kcal/mol lower than the excited-flavin energy (table S6). Concomitantly, electronic coupling between the flavin and substrate increased almost threefold, reaching 47 meV (table S7), thus favoring the transfer of an electron from the carboxylate. The formed carboxylic radical underwent a barrierless decarboxylation (fig. S17), affording the alkyl radical (3). The energy of the alkyl radical intermediate was 52 kcal/mol above the dark state (1), which corresponds to >80% of the photon energy (Fig. 6B). CO<sub>2</sub> dissociation increased the distance separation in the radical pair, and, accordingly, the electronic coupling between the alkyl and flavin radicals was only 7 meV (table S7).

From the alkyl radical (3), the alkane product (6/6a) can be formed through several reaction pathways (Fig. 6, pathways I to III). The chemical changes characterizing pathways I to III are summarized in Fig. 6C. HAT from a nearby residue has been discussed previously (4, 24). Our results indicate strong interactions with the Y466 phenolic side chain (electronic coupling 142 meV; table S7) that may facilitate migration of the alkyl radical toward Y466 and C432. The HAT reaction from C432 (Fig. 6, pathway I) proceeded through an energy barrier (9 kcal/mol; fig. S18) and led to a Cys-radical state (4) with a 15 kcal/mol lower energy than that of the alkyl radical intermediate (3). Re-oxidation of the flavin by C432 (5) further reduced the energy by 3 kcal/mol (Fig. 6B). The

resulting thiolate anion was stabilized by a hydrogen bond with positively charged R451; however, proton transfer from R451 to C432 in (5) along the hydrogen bond yielding (5a) did not further reduce the energy (table S5). Therefore, it is likely that C432 reprotonation proceeds by a different mechanism (see below).

Alternatively to the HAT reaction involving C432, a proton can be transferred from Wat2, which stabilizes the alkyl radical (3). The shift of the negative charge from the flavin to the alkane by bET coupled to proton transfer from Wat2 led to a transient formation of a hydroxyl anion which either interacted with CO<sub>2</sub>, forming bicarbonate (6a) directly (Fig. 6, pathway III), or deprotonated R451 (5b) (Fig. 6, pathway II). The presence of CO<sub>2</sub> and water molecules in the active site, in particular Wat1, allows R451 reprotonation concomitant with bicarbonate formation (6), even after the alkane was formed, with a small activation energy (4 kcal/mol; fig. S19). The bicarbonate product derived from Wat1 (6) was 10 kcal/mol lower than the bicarbonate originating from Wat2 (6a), rendering pathway II more energetically favorable than pathway III (Fig. 6B). Energy lowering by bicarbonate formation may also further stabilize product (5) of the HAT reaction (pathway I); in the first step, the thiolate C432 obtains a proton from R451 and/or Wat1 and in the second step, bicarbonate is formed from CO<sub>2</sub> and OH<sup>−</sup> derived from Wat1 (6).

In view of the experimentally observed red-shifted reoxidized flavin intermediate, we compared the flavin excitation energy of the species resulting from the bET reaction and that of the initial dark state (Fig. 6B and table S8). Our computations suggest that the red shift can be explained by formation of various species in which the initial negative charge of the deprotonated carboxylate (1) is either neutralized by formation of CO<sub>2</sub> and deprotonated R451 (5) or shifted away from the flavin by formation of the anionic bicarbonate (6) and (6a) (fig. S20). Additionally, a red-shifted spectrum is caused by hydrogen-bonding interactions of Wat1 with flavin in (5a).

Consistent with the x-ray structures, all active-site models contained the butterfly-bending conformation of the flavin isoalloxazine ring. Flavin bending persisted during geometry optimization of the FAP active-site models (Fig. 6A), in contrast to the essentially

planar optimized geometry of the oxidized and semireduced forms obtained in computations of flavins (36). The significant bending angle in the optimized FAP active-site structures is consistent with the notion that interactions with the protein modulate flavin bending (37). As previously discussed, this bending biases the energy levels of the flavin, favoring flavin reduction (36) and decreasing the excitation energy (table S9), with possible functional implications. In particular, bending diminishes the vertical electron affinity even more than the excited-flavin energy (fig. S21) and thus preferentially reduces the CT-state energy, which should facilitate fET. Thus, flavin bending explains the strongly red-shifted (20-nm) ground-state FAD absorption spectrum of FAP in the dark state compared with free FAD or most flavoproteins, allowing photoexcitation and facile fET up to as far as 530 nm. This is well into the so-called “green gap” (between 500 and 600 nm) in the absorption spectra of chlorophylls that dominate the absorption of algae, thus enhancing the net light-harvesting capacities of FAP.

#### Role of conserved amino acids in the FAP active site

The active site of CvFAP contains three residues (Y466, C432, and R451) that are strictly conserved and specific to FAPs compared with other GMC oxidoreductases (fig. S23). Y466 and C432 were previously considered as hydrogen atom donor to the alkyl radical R<sup>•</sup> (followed by bET from FAD<sup>•−</sup> to the tyrosyl or cysteinyl radical) or, alternatively, proton donor in a proton-coupled back-ET from FAD<sup>•−</sup> to R<sup>•</sup> (4, 24). Our quantum chemical calculations additionally suggest R451 as a potential proton donor. To identify the proton donor experimentally, we performed FTIR; in particular, we analyzed possible contributions from deprotonated forms of these three residues in the cryotrapped red-shifted intermediate. Light-induced FTIR difference spectra of CvFAP recorded at 100 K showed only a shift of a  $\nu(\text{S-H})$  IR mode of cysteine, and no changes in D<sub>2</sub>O compared with H<sub>2</sub>O that could be assigned to tyrosine vibrations, which does not support deprotonation of either of the two residues under conditions in which the flavin red shift was observed (Fig. 4A, vi, and supplementary text S2). However, a strong negative band was observed at 1606 cm<sup>−1</sup> in



FTIR difference spectra recorded in D<sub>2</sub>O (Fig. 4A, vii). In <sup>15</sup>N-labeled FAP samples, this band downshifted to 1597 cm<sup>-1</sup> (Fig. 4A, viii). Both observations support assignment of the band to the guanidinium IR mode of an arginine side chain. Absence of a positive counterpart of this band after illumination is indicative of arginine deprotonation in the FAD<sub>RS</sub> state at 100 K (see supplementary text S2).

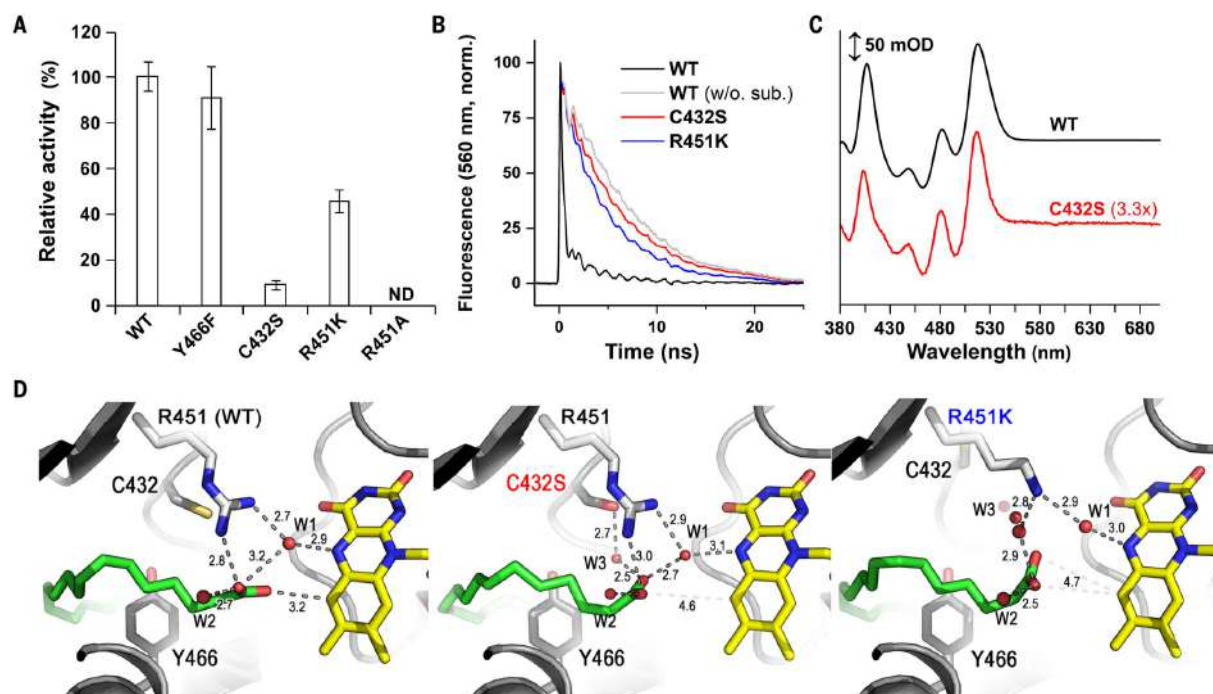
To gain further insight into the role of Y466, C432, and R451, we prepared Y466F, C432S, R451A, and R451K mutants and performed activity measurements and extensive structural and spectroscopic characterization (UV-Vis absorption spectra are shown in fig. S24). Mutation of Y466 to phenylalanine affected the catalytic activity only slightly (Fig. 7A), consistent with a previous report (24). The kinetics of both fET to FAD and bET from FAD<sup>•-</sup> did not differ much from wild-type (WT) CvFAP (fig. S25C). Furthermore, almost identical FTIR difference spectra were obtained with WT CvFAP and the Y466F mutant (fig.

S26A). Altogether, these results suggest that Y466 does not directly participate in any PT/HAT or ET step in the FAP photocycle, and its role is only briefly discussed in the supplementary text S4.

Mutation of C432 to serine (a much poorer proton and hydrogen atom donor) strongly affected the catalytic activity (Fig. 7A), as reported earlier by Heyes *et al.*, who considered it indicative of HAT from C432 to the alkyl radical R<sup>•</sup> in the WT protein (24). In contrast to these authors, we detected a low but significant catalytic activity (~10% of WT) for the C432S mutant. Because impairment of the catalytic activity by a point mutation may result from structural changes rather than suppression of a direct function of the replaced residue, we examined the C432S mutant protein in more detail. The dark-state crystal structure of C432S was highly similar to WT except for a new water molecule, Wat3, interacting with S432 (2.7 Å) and the O1 oxygen atom of the fatty acid carboxylate (2.5 Å),

which was rotated by ~42° (Fig. 7D). The distance of the carboxylate O2 to the flavin N5 increased from 4.0 Å in WT to 5.0 Å.

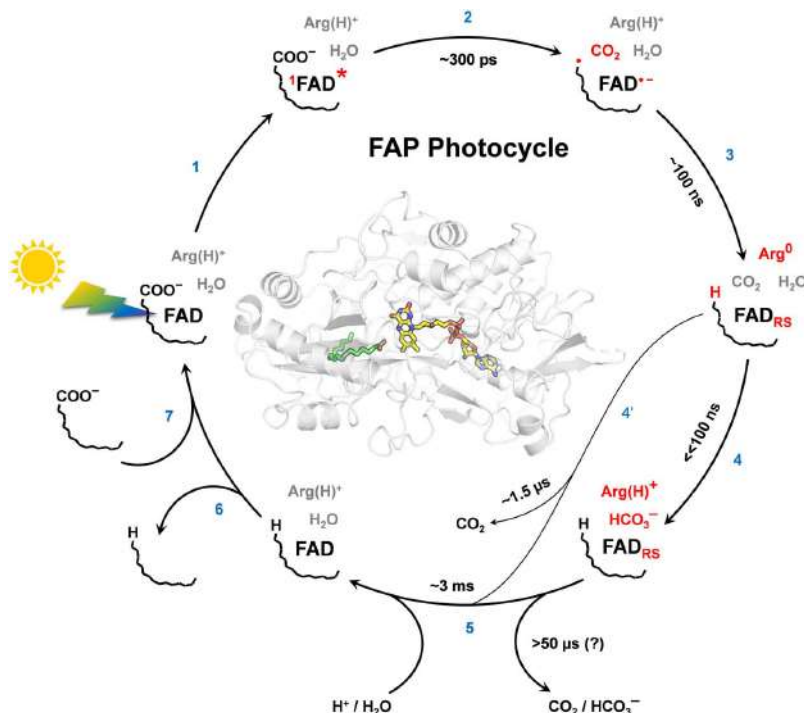
The time-resolved fluorescence signal of C432S resembled that of WT after the consumption of native substrate(s); however, the fluorescence decay was slightly (~10%) accelerated (Fig. 7B), indicating that the signal reflects mostly intrinsic <sup>1</sup>FAD<sup>•-</sup> decay [mostly because of intersystem crossing (ISC) (4)], with only a small contribution of competing fET to <sup>1</sup>FAD<sup>•-</sup>. The intrinsic fET rate was ~10× slower than <sup>1</sup>FAD<sup>•-</sup> decay in this mutant (see supplementary text S3). Transient absorption signals on the submillisecond time scale were thus dominated by the triplet (fig. S25F), and the formation of FAD<sub>RS</sub> could not be resolved. There was, however, a small (~15% of WT) long-lived absorption change at 515 nm (where the transient flavin red shift is most prominent in the WT). About two-thirds of it decayed in a few milliseconds, i.e., with similar kinetics as FAD<sub>RS</sub> in WT [see upper inset of fig.



**Fig. 7. Impact of mutations of conserved residues on CvFAP activity, <sup>1</sup>FAD<sup>•-</sup> fluorescence decay, spectrum of cryotrapped FAD<sub>RS</sub>, and active-site structure.** (A) Activities of purified recombinant mutant CvFAPs relative to WT (measured by gas chromatography coupled with mass spectrometry for the Y466F, R451K, and R451A mutants) in the presence of *cis*-vaccenic acid as substrate. Relative activity of the C432S mutant was measured by membrane inlet mass spectrometry to avoid activity underestimation due to low photostability of this particular mutant under continuous illumination conditions. All activities were normalized to FAD content. Mean ± SD is shown (*n* = 5 repeats). ND, not detected. (B) Normalized time-resolved

fluorescence at 560 nm of WT, C432S, and R451K CvFAP in the presence of native substrates (and after their consumption in WT); for Y466F and R451A mutants, see fig. S25B. (C) Light minus dark spectrum of WT and mutant C432S at 200K obtained by cryo-UV-Vis spectroscopy and normalized on FAD content. (D) Structure of the active site of the WT (left), C432S (middle), and R451K (right) mutants. Distances (in angstroms) between substrate (green), FAD (yellow), water molecules (red), and amino acid residues (white or gray) are shown. Compared with WT, in the C432S and R451K mutants, the FA carboxylate is rotated ~50° around an axis defined by the fatty acid atoms O2 and C2.

**Fig. 8. Suggested CvFAP photocycle.** Upon light excitation (1), fET in ~300 ps from the fatty acid anion to  $^1\text{FAD}^*$  (observed by ultrafast fluorescence and transient absorption spectroscopies) leads to its quasi-instantaneous decarboxylation (2), as observed by TR-IR and TR-SFX and supported by the computed absence of an energy barrier. bET in ~100 ns from  $\text{FAD}^{*-}$  (presumably to the alkyl radical) results in formation of red-shifted (re-) oxidized flavin  $\text{FAD}_{\text{RS}}$ ; the H/D KIE suggests that bET is coupled to and/or limited by PT. Cryotrapping FTIR experiments suggest arginine as the final proton donor to the alkyl (3). Concomitantly, most  $\text{CO}_2$  (~75%) is transformed (4) to bicarbonate as indicated by TR-IR and cryotrapping FTIR.  $\text{FAD}_{\text{RS}}$  disappears in ~3 ms (5) with a H/D KIE > 3, indicating coupling to PT. Upon alkane release (6), new substrate binds (7). Changes after individual steps are marked in red; time constants are for RT.



S25F or figure 3D in (24)]. With these indirect indications of  $\text{FAD}_{\text{RS}}$  formation in C432S, we attempted to accumulate this species at a cryogenic temperature. Illumination of the sample at 200 K yielded a spectrum with similar shape (but lower amplitude) as the  $\text{FAD}_{\text{RS}}$  spectrum obtained with WT FAP (Fig. 7C).

When R451 was mutated to alanine, the enzymatic activity was completely abolished (Fig. 7A) and no fET was observed (fig. S25B); in fact, the structure showed that the fatty acid was oriented very differently from that in the WT (see fig. S26D and supplementary text S3 and S4 for details).

When R451 was replaced by lysine (which is also positively charged), the catalytic activity amounted to ~45% of WT (Fig. 7A). The fluorescence decay (Fig. 7B) was distinctly faster (~4.5 ns) than in WT without substrate (~6.5 ns), consistent with fET occurring for ~30% of the excited flavins. Reoxidation of  $\text{FAD}^{*-}$  and formation of  $\text{FAD}_{\text{RS}}$  was clearly resolved and found to be markedly faster (~30 ns; fig. S25D) than in WT (~100 ns; Fig. 3B, fig. S10A, and supplementary text S3 and S4). The acceleration of the (proton-coupled) bET would be consistent with lysine being a better proton donor than arginine (solution  $\text{pK}_a$  of Lys is 10.7 versus 12.1 for Arg).

To better understand how R451 affects the active-site architecture, we determined the crystal structure of the R451K mutant in its

dark state at 100 K (Fig. 7D and fig. S26C). The R451K mutant structure differed significantly from WT: Whereas a new water molecule mimicked the  $\text{NH}_2$  group of R451 and thereby retained the interaction with the O2 oxygen atom of the fatty acid carboxylate (2.5 Å), the interaction between K451 and the fatty acid O1 oxygen atom induced an ~54° rotation of the carboxylate and an increase by 0.6 Å of the distance of the fatty acid O2 to the flavin N5 compared with WT. This created space for a new water molecule, Wat3, akin to the situation in C432S, located between the fatty acid O1 atom (2.4 Å) and K451 (2.8 Å). The close distance of Wat3 and C432 induced a flip of the amino acid stretch T430 to G435, pushing C432 out of the active site and bringing in an additional water molecule. These large structural changes may explain the very different FTIR spectra obtained with the R451K mutant compared with WT (fig. S26B).

#### Reaction cycle of FAP

Consistent with the strict conservation of C432 and R451 in FAP sequences, even conservative substitutions of these residues resulted in drastic reductions in catalytic activity, in both cases by strongly reducing fET. Unexpectedly, the R451K and C432S mutants shared significant structural modifications with respect to the WT: the presence of a new water molecule, Wat3, close to the fatty acid carboxylate, a rotation of

the carboxylate by ~50°, a significant elongation of the distance of the fatty acid carboxylate to the flavin N5, as well as small changes in Wat1 location. Quantum chemistry calculations showed that in the C432S mutant, the CT energy increased by 0.2 eV and the electronic coupling reduced fivefold, which is consistent with a much slower observed fET (see supplementary text 6.7 and fig. S22).

These mutants provide important insights into the molecular constraints affording FAP activity: The active site of the WT enzyme is arranged such as to optimize the configuration of the fatty acid carboxylate for fET. Each carboxylate oxygen atom interacts with catalytically important groups (O1: Wat1, O2: R451 and Wat2) while avoiding an inactivating bidentate interaction with R451.

Three possible routes toward alkane formation in FAP were suggested by quantum chemistry (Fig. 6C). Our experimental findings allow assessing the suggested pathways. Pathway I, involving a HAT mechanism to reduce the alkyl radical, as also suggested previously (4, 24), is chemically plausible. However, a number of experimental findings argue against it. First, the C432S mutant retains significant enzymatic activity, suggesting that C432 is not essential for catalysis. Second, in the  $\text{FAD}_{\text{RS}}$  state, only a shift of a thiol S-H vibration is observed by FTIR but no cysteine deprotonation. Third, despite the fact that C432 is



rotated out of the active site in the R451K mutant, this variant is quite active. Although it is possible that C432 rotates back into the active site upon changes in the water structure after CO<sub>2</sub> cleavage, this would not explain the threefold faster bET in R451K. Reaction pathway I (Fig. 6C) seems unlikely, although we cannot rule out that the C432S mutant uses a different mechanism than the WT enzyme. The observation of residual electron density consistent with bicarbonate close to C432 suggests that C432 may stabilize reaction products such as CO<sub>2</sub> and/or bicarbonate away from the original position of the substrate carboxylate.

The other two pathways proceed through PCET (Fig. 6C). Pathway III implies a catalytic role of Wat2 in bicarbonate formation. This mechanism is unlikely, as no changes were observed for Wat2 in the electron density maps derived by TR-SFX. By contrast, 300 ns after photoexcitation, a significant loss of electron density of Wat1 was observed, supporting pathway II. With the cleaved CO<sub>2</sub> present in the active site, as evidenced by cryocrystallography and IR spectroscopy, Wat2 can serve as a proton donor. The transiently deprotonated R451 activates Wat1, resulting in bicarbonate formation in <100 ns. This transformation of CO<sub>2</sub> to bicarbonate is orders of magnitude faster than in solution [tens of seconds (38)], indicating a strong catalytic effect.

R451's role as transient proton donor was a priori unexpected because proton transfers from and to arginine residues are rare (39, 40) due to their relatively high pK<sub>a</sub> values. In FAP, the strong basicity of OH<sup>−</sup> formed from Wat2 by proton-coupled ET to the alkyl radical may allow proton transfer from R451. In addition to this catalytic function, R451 is crucial for the FAP active-site architecture by precisely positioning and orienting the fatty acid head group with respect to FAD and stabilizing the carboxylate in the catalytically active deprotonated form.

Fig. 8 and its legend summarize our comprehensive understanding of the very complex cycle. By combining results obtained by a multitude of experimental techniques and computations, we provide a detailed mechanistic description of the evolution of the reactant (fatty acid) to the products [alka(e)ne and CO<sub>2</sub>] and the role of the protein moiety involving a proton-coupled electron-transfer mechanism. We demonstrate partly unexpected structural and dynamic properties of FAP, including features that have not been observed in other flavoproteins and other enzymatic reactions in general. Understanding these catalytic features is an important step in incorporating FAP into the green chemistry toolkit.

## Materials and methods summary

The FAP used in all experiments corresponds to residues 76 to 654 of the full-length C<sub>7</sub>FAP

(or single mutants thereof obtained by site-directed mutagenesis). WT C<sub>7</sub>FAP and C<sub>7</sub>FAP mutants expressed in *Escherichia coli* and purified were studied by x-ray crystallography (static and TR-SFX), spectroscopy in solution (FTIR, TR-IR on picosecond to microsecond time scales, time-resolved fluorescence spectroscopy on picosecond to nanosecond time scales, transient absorption spectroscopy on nanosecond to millisecond time scales, and ultrafast visible absorption and fluorescence spectroscopy) and spectroscopy on single crystals (UV-Vis and Raman). Activity assays were based on quantification of hydrocarbons formed (by gas chromatography coupled to mass spectrometry) or CO<sub>2</sub> released (by membrane inlet mass spectrometry). Computational studies involved molecular dynamics simulations and quantum chemistry calculations on C<sub>7</sub>FAP and multiple alignments of GMC oxidoreductase protein sequences. Detailed materials and methods are available in the supplementary materials.

## REFERENCES AND NOTES

1. L. O. Björn, "Photoactive proteins," in *Photobiology: The Science of Light and Life* (Springer, ed. 3, 2015), pp. 139–150.
2. A. Sancar, Mechanisms of DNA repair by photolyase and excision nuclease (Nobel Lecture). *Angew. Chem. Int. Ed.* **55**, 8502–8527 (2016). doi: [10.1002/anie.201601524](https://doi.org/10.1002/anie.201601524); PMID: 27337655
3. M. Gabruk, B. Mysliwa-Kurdziel, Light-dependent protochlorophyllide oxidoreductase: phylogeny, regulation, and catalytic properties. *Biochemistry* **54**, 5255–5262 (2015). doi: [10.1021/acs.biochem.5b00704](https://doi.org/10.1021/acs.biochem.5b00704); PMID: 26230427
4. D. Sorigué et al., An algal photoenzyme converts fatty acids to hydrocarbons. *Science* **357**, 903–907 (2017). doi: [10.1126/science.1246349](https://doi.org/10.1126/science.1246349); PMID: 28860382
5. D. Zhong, Ultrafast catalytic processes in enzymes. *Curr. Opin. Chem. Biol.* **11**, 174–181 (2007). doi: [10.1016/j.cbpa.2007.02.034](https://doi.org/10.1016/j.cbpa.2007.02.034); PMID: 17353141
6. M. J. Maul et al., Crystal structure and mechanism of a DNA (6-4) photolyase. *Angew. Chem. Int. Ed.* **47**, 10076–10080 (2008). doi: [10.1002/anie.200804268](https://doi.org/10.1002/anie.200804268); PMID: 18956392
7. P. Müller, J. Yamamoto, R. Martin, S. Iwai, K. Brettel, Discovery and functional analysis of a 4th electron-transferring tryptophan conserved exclusively in animal cryptochromes and (6-4) photolyases. *Chem. Commun. (Camb.)* **51**, 15502–15505 (2015). doi: [10.1039/C5CC06276D](https://doi.org/10.1039/C5CC06276D); PMID: 26355419
8. A. Mees et al., Crystal structure of a photolyase bound to a CPD-like DNA lesion after in situ repair. *Science* **306**, 1789–1793 (2004). doi: [10.1126/science.1101598](https://doi.org/10.1126/science.1101598); PMID: 15576622
9. D. J. Heyes, C. N. Hunter, I. H. M. van Stokkum, R. van Grondelle, M. L. Groot, Ultrafast enzymatic reaction dynamics in protochlorophyllide oxidoreductase. *Nat. Struct. Mol. Biol.* **10**, 491–492 (2003). doi: [10.1038/nsb929](https://doi.org/10.1038/nsb929); PMID: 12730687
10. N. S. Scrutton, M. L. Groot, D. J. Heyes, Excited state dynamics and catalytic mechanism of the light-driven enzyme protochlorophyllide oxidoreductase. *Phys. Chem. Chem. Phys.* **14**, 8818–8824 (2012). doi: [10.1039/c2cp23789j](https://doi.org/10.1039/c2cp23789j); PMID: 22419074
11. D. J. Heyes et al., Excited-state charge separation in the photochemical mechanism of the light-driven enzyme protochlorophyllide oxidoreductase. *Angew. Chem. Int. Ed.* **54**, 1512–1515 (2015). doi: [10.1002/anie.201409881](https://doi.org/10.1002/anie.201409881); PMID: 25488797
12. S. Zhang et al., Structural basis for enzymatic photocatalysis in chlorophyll biosynthesis. *Nature* **574**, 722–725 (2019). doi: [10.1038/s41586-019-1685-2](https://doi.org/10.1038/s41586-019-1685-2); PMID: 31645759
13. C.-S. Dong et al., Crystal structures of cyanobacterial light-dependent protochlorophyllide oxidoreductase. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 8455–8461 (2020). doi: [10.1073/pnas.1920244117](https://doi.org/10.1073/pnas.1920244117); PMID: 32234783
14. M. A. Emmanuel, N. R. Greenberg, D. G. Oblinsky, T. K. Hyster, Accessing non-natural reactivity by irradiating nicotinamide-dependent enzymes with light. *Nature* **540**, 414–417 (2016). doi: [10.1038/nature20569](https://doi.org/10.1038/nature20569); PMID: 27974767
15. N. S. Scrutton, Enzymes make light work of hydrocarbon production. *Science* **357**, 872–873 (2017). doi: [10.1126/science.aao4399](https://doi.org/10.1126/science.aao4399); PMID: 28860372
16. T. Courtney, A. Deiters, Recent advances in the optical control of protein function through genetic code expansion. *Curr. Opin. Chem. Biol.* **46**, 99–107 (2018). doi: [10.1016/j.cbpa.2018.07.011](https://doi.org/10.1016/j.cbpa.2018.07.011); PMID: 30056281
17. S. Moulin et al., Continuous photoproduction of hydrocarbon drop-in fuel by microbial cell factories. *Sci. Rep.* **9**, 13713 (2019). doi: [10.1038/s41598-019-50261-6](https://doi.org/10.1038/s41598-019-50261-6); PMID: 31548626
18. W. Zhang et al., Hydrocarbon synthesis via photoenzymatic decarboxylation of carboxylic acids. *J. Am. Chem. Soc.* **141**, 3116–3120 (2019). doi: [10.1021/jacs.8b12282](https://doi.org/10.1021/jacs.8b12282); PMID: 30673222
19. M. Am et al., Low carbon strategies for sustainable bio-alkane gas production and renewable energy. *Energy Environ. Sci.* (2020).
20. N. A. Herman, W. Zhang, Enzymes for fatty acid-based hydrocarbon biosynthesis. *Curr. Opin. Chem. Biol.* **35**, 22–28 (2016). doi: [10.1016/j.cbpa.2016.08.009](https://doi.org/10.1016/j.cbpa.2016.08.009); PMID: 27573483
21. J. Xu et al., Light-driven kinetic resolution of  $\alpha$ -functionalized carboxylic acids enabled by an engineered fatty acid photodecarboxylase. *Angew. Chem. Int. Ed.* **58**, 8474–8478 (2019). doi: [10.1002/anie.201903165](https://doi.org/10.1002/anie.201903165); PMID: 31033108
22. D. Sorigué et al., Microalgae synthesize hydrocarbons from long-chain fatty acids via a light-dependent pathway. *Plant Physiol.* **171**, 2393–2405 (2016). doi: [10.1104/pp.16.00462](https://doi.org/10.1104/pp.16.00462); PMID: 27288359
23. S. Moulin, A. Beyly, S. Blangy, B. Légeret, M. Floriani, A. Burlacot, D. Sorigué, Y. Li-Beisson, G. Peltier, F. Beisson, Fatty acid photodecarboxylase is an ancient photoenzyme responsible for hydrocarbon formation in the thylakoid membranes of algae. *bioRxiv* 166330 [Preprint]. 23 June 2020. <https://doi.org/10.1101/2020.06.23.166330>
24. D. J. Heyes et al., Photochemical mechanism of light-driven fatty acid photodecarboxylase. *ACS Catal.* **10**, 6691–6696 (2020). doi: [10.1021/acscatal.0c01684](https://doi.org/10.1021/acscatal.0c01684); PMID: 32905273
25. T. Senda, M. Senda, S. Kimura, T. Ishida, Redox control of protein conformation in flavoproteins. *Antioxid. Redox Signal.* **11**, 1741–1766 (2009). doi: [10.1089/ars.2008.2348](https://doi.org/10.1089/ars.2008.2348); PMID: 19243237
26. A. K. Röhr, H.-P. Hersleth, K. K. Andersson, Tracking flavin conformations in protein crystal structures with Raman spectroscopy and QM/MM calculations. *Angew. Chem. Int. Ed.* **49**, 2324–2327 (2010). doi: [10.1002/ange.200907143](https://doi.org/10.1002/ange.200907143); PMID: 20187055
27. A. Aleksandrov, A molecular mechanics model for flavins. *J. Comput. Chem.* **40**, 2834–2842 (2019). doi: [10.1002/jcc.26061](https://doi.org/10.1002/jcc.26061); PMID: 31471978
28. H. N. Chapman et al., Femtosecond X-ray protein nanocrystallography. *Nature* **470**, 73–77 (2011). doi: [10.1038/nature09750](https://doi.org/10.1038/nature09750); PMID: 21293373
29. K. Hirata et al., Determination of damage-free crystal structure of an X-ray-sensitive protein using an XFEL. *Nat. Methods* **11**, 734–736 (2014). doi: [10.1038/nmeth.2962](https://doi.org/10.1038/nmeth.2962); PMID: 24813624
30. S. P. Liptonok et al., Ultrafast real-time visualization of active site flexibility of flavoenzyme thymidylate synthase ThxY. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 8924–8929 (2013). doi: [10.1073/pnas.1218729110](https://doi.org/10.1073/pnas.1218729110); PMID: 23671075
31. L. H. Jones, E. McLaren, Infrared Absorption Spectra of SO<sub>2</sub> and CO<sub>2</sub> in Aqueous Solution. *J. Chem. Phys.* **28**, 995 (1958). doi: [10.1063/1.1744329](https://doi.org/10.1063/1.1744329)
32. L. Antonucci, A. Bonvalet, X. Solinas, L. Daniault, M. Joffre, Arbitrary-detuning asynchronous optical sampling with amplified laser systems. *Opt. Express* **23**, 27931–27940 (2015). doi: [10.1364/OE.23.027931](https://doi.org/10.1364/OE.23.027931); PMID: 26480451
33. X. Solinas, L. Antonucci, A. Bonvalet, M. Joffre, Multiscale control and rapid scanning of time delays ranging from picosecond to millisecond. *Opt. Express* **25**, 17811–17819 (2017). doi: [10.1364/OE.25.017811](https://doi.org/10.1364/OE.25.017811); PMID: 28789272
34. B. A. Diner, D. A. Force, D. W. Randall, R. D. Britt, Hydrogen bonding, solvent exchange, and coupled proton and electron transfer in the oxidation and reduction of redox-active tyrosine Y(Z) in Mn-depleted core complexes of photosystem II. *Biochemistry* **37**, 17931–17943 (1998). doi: [10.1021/bi981894r](https://doi.org/10.1021/bi981894r); PMID: 9922161
35. J.-P. Colletier, G. Schirò, M. Weik, "Time-resolved serial femtosecond crystallography, towards molecular movies of

- biomolecules in action," in *X-Ray Free Electron Lasers: A Revolution in Structural Biology*, S. Boutet, P. Fromme, M. S. Hunter, Eds. (Springer, 2018), pp. 331–356.
36. J. D. Walsh, A.-F. Miller, Flavin reduction potential tuning by substitution and bending, *J. Mol. Struct. THEOCHEM* **623**, 185–195 (2003). doi: [10.1016/S0166-1280\(02\)00719-4](https://doi.org/10.1016/S0166-1280(02)00719-4)
  37. B. W. Lennon, C. H. Williams Jr., M. L. Ludwig, Crystal structure of reduced thioredoxin reductase from *Escherichia coli*: Structural flexibility in the isoalloxazine ring of the flavin adenine dinucleotide cofactor. *Protein Sci.* **8**, 2366–2379 (1999). doi: [10.1110/ps.8.11.2366](https://doi.org/10.1110/ps.8.11.2366); pmid: [10595539](https://pubmed.ncbi.nlm.nih.gov/10595539/)
  38. N. McCann *et al.*, Kinetics and mechanism of carbamate formation from CO<sub>2</sub>(aq), carbonate species, and monoethanolamine in aqueous solution. *J. Phys. Chem. A* **113**, 5022–5029 (2009). doi: [10.1021/jp810564z](https://doi.org/10.1021/jp810564z); pmid: [19338322](https://pubmed.ncbi.nlm.nih.gov/19338322/)
  39. Y. Xiao, M. S. Hutson, M. Belenky, J. Herzfeld, M. S. Braiman, Role of arginine-82 in fast proton release during the bacteriorhodopsin photocycle: A time-resolved FT-IR study of purple membranes containing 15N-labeled arginine. *Biochemistry* **43**, 12809–12818 (2004). doi: [10.1021/bi049238g](https://doi.org/10.1021/bi049238g); pmid: [15461453](https://pubmed.ncbi.nlm.nih.gov/15461453/)
  40. P. J. Silva, C. Schulz, D. Jahn, M. Jahn, M. J. Ramos, A tale of two acids: When arginine is a more appropriate acid than H<sub>3</sub>O<sup>+</sup>. *J. Phys. Chem. B* **114**, 8994–9001 (2010). doi: [10.1021/jp100961s](https://doi.org/10.1021/jp100961s); pmid: [20553007](https://pubmed.ncbi.nlm.nih.gov/20553007/)

# ACKNOWLEDGMENTS

We thank J. Woodhouse for help with microcrystallization, L. Uriarte for help with spectroscopy experiments on crystals, G. Nass Kovacs for preparation of injection conditions, M. Tarnawski for FAP stability measurements, F. Maia for helpful discussions and uploading the SFX data to CXIDB.org, M. Philibert and the GRAP platform for custom manufacturing of scientific equipment. The ESRF is acknowledged for access to beamlines via its in-house research program. This work used the iCOS Laboratory, which is a platform of the Grenoble Instruct-ERIC Centre (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) within the Grenoble Partnership for Structural Biology (PSB). The IBS acknowledges integration into the Interdisciplinary Research Institute of Grenoble (IRIG, CEA) and financial support by the CEA, the CNRS, and the UGA. The present work has benefited from the platform of Biophysics of I2BC supported by French

Infrastructure for Integrated Structural Biology (FRISBI) ANR-10-INBS-05 and from the platform HelioBiotech (BIAM). The experiments were performed at the Linac Coherent Light Source (LCLS), SLAC National Accelerator Laboratory. Use of the LCLS is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract no. DE-AC02-76SF00515. Part of the sample injector used at LCLS for this research was funded by the National Institutes of Health, P41GM103393, formerly P41RR001209. **Funding:** This work was supported by ANR SNAPS<sub>HOTs</sub> to F.B., M.W., P.M., K.B., M.H.V., P.A., R.H., and C.B.; ANR Photoalkane to G.P.; ANR SignalBioRNJ to Y.L.-B.; ANR BioXFEL to M.W.; ERC Consolidator Grant STepLADDER (724362) to T.R.M.B.; a Chinese Scholarship Council fellowship to B.Z.; and a MENESR-Univ. Grenoble Alpes fellowship to K.H.; a grant from the Chevreul Institute and the Ministère de l'Enseignement Supérieur et de la Recherche and the Région Nord-Pas de Calais (FEDER) to M.S. **Author contributions:** F.B., M.W., P.A., P.M., I.S., K.B., M.H.V., C.B., and T.D. designed and organized the project. J.-P.C., M.J., T.R.M.B., and A.R. supervised parts of the project. D.S., S. Blangy, P.S., S. Cuiné, B.L., S.M., Y.L.-B., G.P., E.H., and F.B. analyzed FAP sequences, generated mutants, expressed and purified the FAPs, and performed activity assays. D.S., S. Blangy, E.H., and P.A. produced crystals for synchrotron experiments. P.A., G.G., A.R., D.N., P.L., I.S., and D.S. acquired synchrotron diffraction data. G.G., P.A., and A.R. performed *in crystallo* optical spectroscopies, and P.A., I.S., G.G., P.L., and A.R. interpreted these data. D.S., R.H., and C.B. built the setup for cryogenic light-induced FTIR difference spectroscopy and acquired and analyzed the data. A. Bonvalet, A. Benachir, and D.S. optimized and performed multiscale TR-IR experiments. L.A., X.S., A. Bonvalet, and M.J. conceived and developed multiscale TR-IR spectroscopy and adapted for FAP experiments. A. Bonvalet., L.A., M.J., and M.H.V. analyzed TR-IR data. M.H.V. performed and analyzed femtosecond pump-probe absorption and fluorescence spectroscopy. B.Z. and M.H.V. performed and analyzed picosecond-pump power dependence spectroscopy. A.A. performed molecular dynamics simulations. D.S., S. Blangy, and K.H. produced microcrystals for SFX experiments, and K.H., J.-P.C., and M.W. tested their diffraction quality at the ESRF. M.S., K.H., and D.S. performed time-resolved emission and absorption spectroscopy on microcrystals. T.R.M.B., M.C., S. Boutet, J.-P.C., N.C., R.B.D.,

L.F., A.G., M.L.G., K.H., M.H., M.K., T.J.L., G.S., I.S., R.L.S., and M.W. performed SFX experiments. R.L.S., R.B.D., M.L.G., and M.K. performed sample injection. S. Carbajo, G.S. and M.C. carried out laser work at the LCLS. S. Boutet and T.J.L. prepared and performed SFX data collection. L.F., M.H., N.C., J.-P.C., and T.R.M.B. performed online SFX data analysis. K.H., N.C., J.-P.C., A.G., and T.R.M.B. performed offline SFX data processing. K.H., J.-P.C., N.C., M.W., T.R.M.B., and I.S. analyzed SFX data. M.B., K.B., and P.M. built the setups for picosecond-nanosecond time-resolved fluorescence and nanosecond-millisecond transient absorption spectroscopy. D.S., P.S., B.Z., P.M., and K.B. performed the measurements. P.M. and K.B. analyzed the data. T.D. performed quantum chemistry calculations and provided mechanistic insight. F.B., P.M., K.B., M.H.V., I.S., P.A., T.D., M.W., C.B., and D.S. wrote the paper. All authors discussed the results. **Competing interests:** L.A., X.S., A. Bonvalet, and M.J. are authors of a patent describing the general principle of ADASOPS (FR1155799, US10190972B2). The authors declare no other competing interests. **Data and materials availability:** All data are available in the manuscript, the supplementary materials, or at publicly accessible repositories. Atomic coordinates of protein structures have been deposited in the Protein Data Bank under accession nos. 6YRU, 6YRV, 6YRX, 6YRZ, 6YSL, 6YS2, 6ZH7, and 7AV4. SFX diffraction images of FAP microcrystals have been deposited in the Coherent X-ray Imaging Data Bank website (CXIDB) with accession ID 177 (<https://cxidb.org/id-177.html>).

# SUPPLEMENTARY MATERIALS

[science.sciencemag.org/content/372/6538/eabd5687/suppl/DC1](https://science.sciencemag.org/content/372/6538/eabd5687/suppl/DC1)  
Materials and Methods  
Supplementary Text S1 to S6  
Figs. S1 to S31  
Tables S1 to S11  
References (41–121)  
MDAR Reproducibility Checklist

[View/request a protocol for this paper from Bio-protocol.](#)

29 June 2020; accepted 17 February 2021  
10.1126/science.abd5687

## Mechanism and dynamics of fatty acid photodecarboxylase

D. Sorigué, K. Hadjidemetriou, S. Blangy, G. Gotthard, A. Bonvalet, N. Coquelle, P. Samire, A. Aleksandrov, L. Antonucci, A. Benachir, S. Boutet, M. Byrdin, M. Cammarata, S. Carbajo, S. Cuiné, R. B. Doak, L. Foucar, A. Gorel, M. Grünbein, E. Hartmann, R. Hienerwadel, M. Hilpert, M. Kloos, T. J. Lane, B. Légeret, P. Legrand, Y. Li-Beisson, S. L. Y. Moulin, D. Nurizzo, G. Peltier, G. Schirò, R. L. Shoeman, M. Sliwa, X. Solinas, B. Zhuang, T. R. M. Barends, J.-P. Colletier, M. Joffe, A. Royant, C. Berthomieu, M. Weik, T. Domratheva, K. Brettel, M. H. Vos, I. Schlichting, P. Arnoux, P. Müller and F. Beisson

*Science* **372** (6538), eabd5687.  
DOI: 10.1126/science.abd5687

### Light makes light work of fatty acids

Photosynthetic organisms are notable for their ability to capture light energy and use it to power biosynthesis. Some algae have gone a step beyond photosynthesis and can use light to initiate enzymatic photodecarboxylation of fatty acids, producing long-chain hydrocarbons. To understand this transformation, Sorigué *et al.* brought to bear an array of structural, computational, and spectroscopic techniques and fully characterized the catalytic cycle of the enzyme. These experiments are consistent with a mechanism starting with electron transfer from the fatty acid to a photoexcited oxidized flavin cofactor. Decarboxylation yields an alkyl radical, which is then reduced by back electron transfer and protonation rather than hydrogen atom transfer. The wealth of experimental data explains how algae harness light energy to produce alka(e)nes and provides an appealing model system for understanding enzyme-catalyzed photochemistry more generally.

*Science*, this issue p. eabd5687

#### ARTICLE TOOLS

<http://science.sciencemag.org/content/372/6538/eabd5687>

#### SUPPLEMENTARY MATERIALS

<http://science.sciencemag.org/content/suppl/2021/04/07/372.6538.eabd5687.DC1>

#### REFERENCES

This article cites 113 articles, 7 of which you can access for free  
<http://science.sciencemag.org/content/372/6538/eabd5687#BIBL>

#### PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)

*Science* (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2021 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works



## **Chapter 4: Substrate specificity of fatty acid photodecarboxylase and autocatalytic effect**



## Objectives, personal contribution and main results

As described in the literature review (FAP/substrate specificity), first characterizations of FAP show that the enzyme prefers LC fatty acids, in line with the features of the hydrophobic pocket accommodating the substrate. There was no detailed substrate specificity study on the purified protein using SC and MC fatty acids. Only low activities were reported on SC fatty acids using cell homogenates or bioconversion experiments with CvFAP. The enzyme was therefore supposed to have very low activities on SC and MC fatty acids. The aim of our study was to understand *in vitro* if the activity is low on SC fatty acids and what can explain this substrate preference by highlighting the possible limitations. FAP being a photoenzyme, we can logically imagine two principal limitations: intrinsic limitation due to the capacity of the enzyme to perform the photocycle and extrinsic limitation due to the way the substrate is presented at the lipid-water interface. We propose to explore these two aspects.

**Personal contribution:** I performed all the *in vitro* activity experiments as well as the bioconversions. I also help to perform all the transient spectroscopy experiments in Saclay in the laboratory of Dr. Pavel Müller. I analyzed data, participated in all scientific discussions, and wrote a draft of the article.

**Main results:** By using time resolved spectroscopy, we screened C6-C18 fatty acids to determine the photochemical yield proper to each substrate. The analysis shows that LC fatty acids (C14-C18) are efficiently converted as expected. Interestingly, we observed an autocatalytic phenomenon on C8-C10 fatty acids which allow these substrates to be converted in maximum quasi-similar yields as LC fatty acids. Molecular Dynamic helps to support the view that alkane produced by the decarboxylation of C8-C10 fatty acids helps the positioning of the initial substrate, thereby improving the photochemical yield. The autocatalytic effect prompted us to better characterize under continuous illumination the conversion yield of MC fatty acid (8:0) compared to a LC fatty acid (16:0). *In vitro*, the optimization of pH and substrate concentration shows that C8 fatty acid can be better converted (4-fold) compared to 16:0 described so far as one of the bests substrates. Since SC fatty acids are very interesting specially to perform continuous photoproduction system in which there is no need to harvest the cell since their corresponding products are volatiles and can diffuse through the cells, we wanted to test the efficiency on C8 fatty acid *in vivo*. This test also showed that C8 fatty acid

can be converted more efficiently (10-fold) compared to C16 fatty acid. Our finding is the first reported efficient activity on MC fatty acids using a non-mutated FAP.

A paper has been published in the journal Science Advances. ( Publication link:  
<https://www.science.org/doi/10.1126/sciadv.adg3881>)



### **Article 3:**

## **Autocatalytic Effect Boosts the Production of Medium-Chain Hydrocarbons by Fatty Acid Photodecarboxylase**



# Autocatalytic Effect Boosts the Production of Medium-Chain Hydrocarbons by Fatty Acid Photodecarboxylase

## Short Title

Photoenzymatic Production of Fuel-like Alkanes

## Authors

Poutoum P. Samire<sup>1,2</sup>, Bo Zhuang<sup>2,3</sup>, Bertrand Légeret<sup>1</sup>, Ángel Baca-Porcel<sup>1</sup>, Gilles Peltier<sup>1</sup>, Damien Sorigué<sup>1</sup>, Alexey Aleksandrov<sup>3\*</sup>, Frédéric Beisson<sup>1\*</sup>, Pavel Müller<sup>2\*</sup>

## Affiliations

<sup>1</sup>Aix-Marseille University, CEA, CNRS, Institute of Biosciences and Biotechnologies, BIAM Cadarache, 13108 Saint-Paul-lez-Durance, France.

<sup>2</sup>Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198 Gif sur Yvette, France.

<sup>3</sup>LOB, CNRS, INSERM, Ecole Polytechnique, Institut Polytechnique de Paris, 91128 Palaiseau Cedex, France.

\*Corresponding authors. E-mails: [pavel.muller@i2bc.paris-saclay.fr](mailto:pavel.muller@i2bc.paris-saclay.fr), [frederic.beisson@cea.fr](mailto:frederic.beisson@cea.fr), [alexey.aleksandrov@polytechnique.edu](mailto:alexey.aleksandrov@polytechnique.edu).

## Abstract

Ongoing climate change and geopolitical tensions are driving the search for renewable, carbon-neutral and local alternatives to fossil fuels. Photocatalytic conversion of fatty acids to hydrocarbons by Fatty Acid Photodecarboxylase (FAP) represents a promising route to green fuels. However, the alleged low activity of FAP on C2-C12 fatty acids seemed to preclude the use for synthesis of gasoline-range hydrocarbons. Here we reveal that *Chlorella variabilis* FAP *in vitro* can convert *n*-octanoic acid four times faster than *n*-hexadecanoic acid, its best substrate reported to date. We also show that *in vivo* this translates into a CvFAP-based production rate over ten-fold higher for *n*-heptane than for *n*-pentadecane. Time-resolved spectroscopy and molecular modeling provide evidence that the high catalytic activity of FAP on *n*-octanoic acid is in part due to an autocatalytic effect of its *n*-heptane product, which fills the rest of the binding

pocket. These results should guide future FAP improvement strategies and represent an important step towards a bio-based and light-driven production of gasoline-like hydrocarbons.

## Teaser

Gasoline-like hydrocarbons can be efficiently produced by a natural algal photoenzyme using nothing but fatty acids and sunlight.

## MAIN TEXT

### Introduction

Fatty Acid Photodecarboxylase (FAP, EC 4.1.1.106), a member of the glucose-methanol-choline (GMC) oxidoreductase family, is an algae-specific enzyme harboring a flavin adenine dinucleotide (FAD) cofactor(1). FAP is one of the few known natural photoenzymes besides photosynthetic reaction centers (which are protein complexes), DNA-repairing enzymes photolyases(2) and light-dependent protochlorophyllide oxidoreductases(3). Despite the short time since the discovery of FAP in the green microalga *Chlorella variabilis* by Sorigué *et al.* in 2017(1), numerous groups have already explored potential applications of this enzyme in biocatalytic processes. Indeed, *Chlorella variabilis* FAP (CvFAP) represents a new, attractive, light-driven and redox-neutral means for the production of *n*-alkanes and *n*-alkenes as a basis for fuels, chemistry and cosmetics(4-7). CvFAP also appears to be a very promising tool for the photocatalytic synthesis of specialty chemicals such as deuterated hydrocarbons(8), enantiomerically pure alpha-amino acids, alpha-hydroxy acids(9), secondary fatty alcohols(10) and aliphatic amines and esters(11).

The natural substrates for the photoproduction of hydrocarbons in the green algae *Chlorella variabilis* and *Chlamydomonas reinhardtii* were previously shown to be C16-C18 linear fatty acids (FAs)(12). Following the discovery of FAP, initial *in vitro* characterization of purified CvFAP and its *Chlamydomonas* homolog indicated that these enzymes indeed exhibited higher affinity for longer (C16-C18) FAs compared to C12-C14 FAs(1). It was also found that the crystal structure of CvFAP expressed in the heterologous host *E. coli*, which contains C10-C18 FAs, had two ‘native’ (likely unsaturated) C18 FA substrates – one at the active site and another one stabilized at the surface of the protein close to the entrance to the tunnel leading to the active site(13). In addition, expression in *E. coli* of four other FAPs chosen in different algal groups among almost two hundred putative FAPs identified in algae genomes or metagenomic

data showed that they were preferentially performing photodecarboxylation of the endogenous C16-C18 FAs rather than C10-C14 FAs(14). Use of *E. coli* cell-free extracts expressing the enzyme confirmed the preference of CvFAP for C16-C18 FAs over C12-C14 FAs(4) and showed that it had low activities on C2-C6 FAs(15). Taken together, all these data lead to the view that CvFAP was adapted to act on C16-C18 substrates and was not likely to be an efficient biological catalyst to produce gasoline-range (C5-C11) hydrocarbons.

Here we reveal that, under the right conditions, CvFAP can in fact be highly active on C8-C10 medium-chain FAs *in vitro*. We also show that the high activity of CvFAP on C8-C10 FAs can be attributed in part to an unexpected autocatalytic effect for which we provide spectroscopic evidence and molecular modeling support. Finally, a bioconversion experiment using bacterial cultures expressing CvFAP provides evidence that this biocatalyst can be used to produce medium-chain hydrocarbons in a much more efficient way than long-chain hydrocarbons.

## Results

### Spectroscopic estimation of C7-C18 FA decarboxylation quantum yield and the observation of autocatalysis

In order to understand the previously reported preference of CvFAP for C16-C18 FAs over shorter ones, we decided to screen the yield of the initial photochemical steps associated with decarboxylation of saturated C7-C18 FA substrates using time-resolved fluorescence (TRF) and transient absorption spectroscopy (TAS). As shown previously(1, 13), the photoexcited singlet flavin adenine dinucleotide cofactor ( $^1\text{FAD}^*$ ) of CvFAP is fluorescent and, in the absence of substrate, its fluorescence decays with a time constant of  $\sim 5\text{-}6$  nanoseconds, yielding an FAD triplet ( $^3\text{FAD}^*$ ), which reverts to the ground state FAD with a time constant of  $\sim 80$   $\mu\text{s}$ (1). However, in the presence of a deprotonated FA substrate ( $\text{R-COO}^-$ ) in the active site, the  $^1\text{FAD}^*$  fluorescence is quenched by forward electron transfer from the substrate to the  $^1\text{FAD}^*$ , yielding an  $\text{FAD}^{\bullet-}$  anion radical, an alkyl radical  $\text{R}^\bullet$  and  $\text{CO}_2$ (13). This is reflected by the appearance of a fast ( $\sim 300$  picoseconds) phase in the TRF signal. The ratio of the amplitudes of the two phases then directly reflects the ratio of the two processes – the unproductive intersystem crossing of  $^1\text{FAD}^*$  to  $^3\text{FAD}^*$  in  $\sim 5\text{-}6$  ns and the productive ET from the FA substrate to  $^1\text{FAD}^*$  in  $\sim 300$  ps, which is followed by a quasi-instantaneous substrate decarboxylation(13). Given that the decarboxylation appears to be essentially irreversible(13),

we suggest that the share of the ~300 ps phase in the TRF signal roughly corresponds to the photodecarboxylation quantum yield.

The subsequent photoreaction step (back ET from  $\text{FAD}^{\bullet-}$  to the alkyl radical  $\text{R}^{\bullet}$  in ~100 ns) leads to the reoxidized flavin with a transiently red-shifted absorption spectrum –  $\text{FAD}_{\text{RS}}$  (see fig. S6C in ref. (1) for  $\text{FAD}_{\text{RS}} - \text{FAD}$  difference spectrum in solution, or fig. S12A,B in ref. (13) for cryo-trapped  $\text{FAD}_{\text{RS}}$ ). The formation (and decay) of  $\text{FAD}_{\text{RS}}$  is best followed by monitoring transient absorption changes around 520 nm (maximum in the  $\text{FAD}_{\text{RS}} - \text{FAD}$  difference spectrum)(1, 13). The red shift disappears within ~3 ms, likely upon binding of a new substrate and/or restoration of the initial charge distribution and hydrogen-bonding network around FAD(1, 13).

To record TRF and TAS signals corresponding to the complex of CvFAP with the given added substrate, the strongly bound native C18 substrates(1, 13) first had to be consumed (see fig. S1A,B) by several strong (~10 mJ per cm<sup>2</sup>) laser flashes at 470 nm (near the absorption maximum of FAD in FAP). In line with the previous studies(1, 15), the activity of CvFAP on all studied medium-chain FA substrates (C7-C12 FAs) as judged by the amplitude ratios in the TRF signals (recorded at 560 nm; see Materials and Methods for reasons why this particular wavelength was chosen) after five (or in some cases less; see Fig. 1 legend for details) strong laser flashes (Fig. 1A,B) was substantially lower than for longer substrates such as C14, C16 or C18 FAs (Fig. 1A,B). However, in the case of *n*-octanoic acid (C8 FA) and, to a lesser extent for *n*-heptanoic, *n*-nonanoic and *n*-decanoic acids (C7, C9 and C10 FAs, respectively), further flashes led to a decrease in the slow fluorescence phase (attributed to non-productive intersystem crossing of the excited flavin) and to an increase in the fast phase reflecting productive forward ET from the substrate to  $^1\text{FAD}^*$  leading to decarboxylation (Fig. 2A,B). After several further laser flashes, the C8 FA TRF signals began to resemble those recorded prior to consumption of the native (likely unsaturated) C18 FA substrate, or those recorded *e.g.*, for saturated C16 FA in one of the control experiments (see fig. S2A,B). This “yo-yo” effect (*i.e.*, an initial decrease in the amplitude of the fast fluorescence phase followed by its recovery upon subsequent flashes) was observed also in the transient absorption signals measured at 515 nm (Fig. 2C), where the initial two flashes led to ever diminishing amounts of  $\text{FAD}_{\text{RS}}$ , but this trend was reversed after the third flash and additional flashes led to an ever increasing amplitude of the ~100 ns kinetic phase corresponding to formation of  $\text{FAD}_{\text{RS}}$  (note that  $\text{FAD}_{\text{RS}}$  is one of the products of a successful decarboxylation). Suspecting an autocatalytic effect of the product (*n*-heptane in the case of C8 FA), we added *n*-heptane (C7 alkane) to a fresh sample of CvFAP

containing native substrate(s) and C8 FA. Upon addition of C7 alkane, the “yo-yo” effect observed in the previous experiments disappeared, and immediately after the consumption of the native C18 FA substrate, the signals stabilized at a similar or even slightly higher (due to the C7 alkane excess) level as after ten flashes in the experiment without added C7 alkane (Fig. 2D vs. 2C), suggesting that our assumption was correct and that the recovered yield of C8 FA photodecarboxylation after multiple flashes (Fig. 2A-C) is indeed due to the presence of C7 alkane generated by decarboxylation of C8 FA upon the preceding flashes (note that the observed quantum yield recovery also shows that the protein was not significantly damaged in the course of the experiments).

These observations inspired us to perform another set of screening experiments, in which we added C5-C12 *n*-alkanes to the CvFAP samples containing C7-C12 FAs. Our results summarized in Fig. 3A-F show that added *n*-alkanes can indeed serve as co-catalysts and enhance the quantum yield of medium-chain FA decarboxylation to ~70%, *i.e.*, to levels as high as those observed for long-chain substrates (Fig. 1).

## MD Simulations

We performed MD simulations with classical force fields to investigate the effects of a co-catalyst in the protein active site (Fig. 4 and figs. S3 and S4). MD simulations lasting 600 ns were performed for FAP with the following molecules in the active site: 1) C8 FA alone, 2) C8 FA with C10 alkane as a co-catalyst, and 3) C18 FA (Fig. 4 and fig. S3). Here, the combination of C8 FA and C10 alkane was preferentially chosen as the native substrate present in the high-resolution crystal structure (13) contains 18 carbon atoms. To investigate the effect of a different chain length, MD simulations lasting 400 ns were also performed for C8 FA with C7 alkane (combination of FA and its decarboxylation product exhibiting the most pronounced autocatalytic effect in the experiment; see Fig. 3) and the corresponding C15 FA (fig. S4). Two additional 500 ns simulations were performed for C8 FA alone starting with different conditions for verification (fig. S3).

As shown in Fig. 4A and fig. S3C, with C10 alkane filling the rest of the substrate-binding tunnel closer to the protein surface, C8 FA was able to stay close to the FAD cofactor throughout the MD simulations. Its position was similar to that of C18 FA in the active site observed in the experimental structure (fig. S3C), with an average distance of 4.1 Å between the terminal carbons of C8 FA and C10 alkane. Similar results were obtained in simulations with C8 FA and C7 alkane (fig. S4). However, due to the longer average distance (4.9 Å)

between the carbons of C8 FA tail and C7 alkane head, C8 FA could transiently move away from FAD and closer to C7 alkane, as shown in fig. S4A-C. Meanwhile, the position of C7 alkane remains unchanged (fig. S4C), preventing C8 FA from further dissociation and eventually returning it to the original position in the active site after ~50 ns (fig. S4A,C). In the absence of alkane, however, substantial structural fluctuations of C8 FA were observed (Fig. 4A,B), regardless of the initial structures, as demonstrated by independent MD simulations (fig. S3A); the C8 FA substrate changed conformations, occupied new positions in the substrate tunnel, and could move as far as ~9 Å away from FAD (Fig. 4A,B and fig. S3). During these rearrangements, carboxyl O atoms of the C8 FA still maintained close interaction with N<sup>η</sup> of the arginine R451 (for most of the simulation time; see fig. S3B), which is a crucial residue for the positioning of the substrate(13), and the space near the FAD previously occupied by the carboxyl group of C8 FA was gradually filled by water molecules.

### **FA decarboxylation efficiency under continuous light**

The selective autocatalytic effect observed for C7-C10 FAs in the spectroscopic experiments described above raised questions about whether this phenomenon also leads to FAP activity enhancement under continuous illumination. To address this question, we compared the *in vitro* activity (chemical yield of decarboxylation) of purified CvFAP on C16 FA (one of the best FAP substrates described so far)(1, 4) and C8 FA (the substrate showing the most pronounced autocatalytic effect in this study). Given that FAP is an interfacial enzyme(16), the saturation of the enzymatic activity can highly depend on the pH and the substrate concentration because these parameters strongly influence the structuring of the FAs and thereby the activity of the enzyme.(17) So far, all FAP activity comparisons on different substrates were done essentially at one pH value and with the same substrate concentrations. This approach may introduce a bias by favoring the photodecarboxylation of certain substrates. To eliminate such bias, we decided to determine the pH optimum for photoconversion of the two compared substrates (C8 and C16 FAs). We screened pH values from 5 to 10 (in a universal Teorell-Stenhagen buffer) and found maximum activity at pH 6.0 for C8 and at pH 8.5 for C16 FAs, see inset of Fig. 5A. Using the optimum pH values, we compared the activity of CvFAP at the activity saturating concentration of the individual substrates (25 mM for C8 FA and 0.5 mM for C16 FA; see fig. S5) by quantifying the corresponding products of decarboxylation (C7 and C15 alkanes, respectively). Our results show that CvFAP produced almost 4-times more C7 alkane from C8 FA than C15 alkane from C16 FA at their respective pH and substrate concentration optima (Fig. 5A).



This result prompted us to evaluate hydrocarbon production in a biotechnological context, where CvFAP would catalyze FA conversion to alkanes inside living cells. *E. coli* bacteria genetically transformed to express CvFAP were therefore grown and C8 or C16 FAs were added (at the same concentrations, 2 mM) to the culture media before illuminating the samples and quantifying the hydrocarbon production. In order to avoid technical biases related to the volatility of the different *n*-alkane products (C7 vs. C15), we used FAs that were <sup>13</sup>C-labeled on the carboxylic group and measured the side product of the reaction - <sup>13</sup>CO<sub>2</sub>. The <sup>13</sup>CO<sub>2</sub> formed by cell cultures contained in sealed flasks upon 1-hour illumination was measured by static headspace (GC/MS analysis of volatile or semi-volatile components in the gas phase above the liquid phase after equilibration in a closed vial; see section ‘Analysis and quantification of <sup>13</sup>CO<sub>2</sub>’ in Materials and Methods for more detail). The results show that CvFAP expressing cell cultures produced almost 13-times more <sup>13</sup>CO<sub>2</sub> from C8 FA than from C16 FA under otherwise identical experimental conditions (Fig. 5B). The large difference in the efficiency of C8 and C16 FAs decarboxylation is not likely to be due to a problem of penetration of C16 FA into the cells because before the start of illumination there was only 10% C16 FAs remaining in the culture medium (see inset of Fig. 5B).

## Discussion

Screening the quantum yield of medium-chain (C7-C12) FA photodecarboxylation by CvFAP using time-resolved spectroscopy, we observed an autocatalytic effect, whereby the initially formed *n*-alkane products enhanced the decarboxylation of further FA substrates by mimicking the missing part of the long-chain, for which the FA binding site of FAP is adapted. This autocatalytic effect was observed for C7-C10 FAs, but it was most pronounced for the C8 FA substrate. These observations inspired us to test the effects of C5-C12 *n*-alkanes added as co-catalysts to C7-C12 FA substrates. Our results clearly show that *n*-alkanes can indeed serve as co-catalysts and decarboxylation quantum yields previously achievable only for long-chain FAs can now be obtained also for medium-chain (C7-C12) FAs, provided the right combination of FA and *n*-alkane is chosen. The highest quantum yields (60-70%) were obtained when the total number of carbon atoms of the FA substrate and the *n*-alkane co-catalyst was  $16 \pm 1$ , in line with the most pronounced autocatalytic effects observed for C8 FA and with C16 FA being one of the best reported substrates for CvFAP(1) (note that the ideal length of ~16 total carbon atoms was suggested also by Zhang *et al.*(15), who attempted to enhance CvFAP enzymatic activity under continuous light on short (C1-C6) substrates by adding C7-C17 *n*-alkanes to crude cell extracts from *E. coli* expressing CvFAP).

Outcomes of our MD simulations suggest that the auto- and/or co-catalytic effect of the *n*-alkanes can be attributed to a steric hindrance that favors the positioning of the medium-chain FAs close enough to the FAD cofactor so that the initial forward electron transfer from the substrate to FAD can occur upon photoexcitation of the latter. The simulations also establish that the conformation of C8 FA and C10 or C7 alkanes closely resembles that observed in the MD simulations with the corresponding long-chain FAs (C18 or C15, respectively).

Our *in vitro* assay demonstrates that the autocatalytic effect (observed in the spectroscopic experiments and rationalized by MD Simulations) together with the optimum pH and substrate concentration can considerably enhance photoconversion of a medium-chain FA (C8 FA) by CvFAP under continuous illumination. The chemical yield of C8 FA decarboxylation exceeding four times that of C16 FA is the best *in vitro* performance of wild-type CvFAP on a medium-chain substrate reported so far. The difference in the optimum pH obtained for C8 and C16 FAs (6.0 and 8.5, respectively) was consistent with experimental data on the  $pK_a$  of C8 to C16 saturated FA salts organized in films (18), as well as quantum chemistry calculations, which show that the lengthening of a fatty acid carbon chain by one methylene group leads to an increase of 0.43 units in the surface  $pK_a$  at the air-water interface (19).

Finally, our *in vivo* assay (photodecarboxylation of 1-<sup>13</sup>C-labeled C8 and C16 FAs inside the living *E. coli* bacteria) shows that the chemical yield of medium-chain *n*-alkanes by FAP under continuous irradiation can greatly exceed that of long-chain ones - by a factor of ~13 for the case of C7 and C15 *n*-alkanes. The autocatalytic effect certainly contributes to the efficient decarboxylation of medium-chain FAs, but the autocatalysis alone is not sufficient to explain the magnitude of this difference. Given that the permeability of the cell membranes seems to be higher for long substrates (such as C16 FA) and that the kinetics of the initial photochemical steps are also similar (for both C8 and C16 FAs), we suggest that the high chemical yield of medium-chain (C8) FA decarboxylation observed both *in vitro* and *in vivo* is most likely due to an acceleration of the 'dark' steps of the photocycle, namely the exchange of the product for a new substrate. Faster replacement of the product with a new substrate (in the case of medium-chain FAs) may also have a positive impact on the (photo)stability of the enzyme(20).

In conclusion, the present work unveils an autocatalytic effect of C7-C10 FA decarboxylation products and documents the co-catalytic effects of C5-C12 *n*-alkanes on the decarboxylation of medium-chain (C7-C12) FAs by CvFAP. It further reveals an unexpectedly high *in vivo* activity of CvFAP on C8 FA acid, which is over ten-fold higher than that on C16 FA, the best FAP

substrate identified prior to our current study. These results should stimulate applied research on the use of CvFAP to convert medium-chain substrates and guide future studies aiming at improving CvFAP for high-yield production of gasoline-like hydrocarbons.

## **Materials and Methods**

### **FAP expression**

The coding sequence corresponding to *Chlorella variabilis* NC64A FAP was cloned into a pLIC07 plasmid (pLIC07 is a pET-28 based expression vector containing downstream FAP gene, a histidine-tagged thioredoxin). The version of CvFAP used here is FAPv2, which corresponds to the full-length protein truncated by the first 75 amino acids at the N-terminus(13, 16). The plasmid pLIC07FAP was then transformed into a BL21 *E. coli* expression strain harboring the pRIL plasmid that allows the synthesis of rare tRNA that are not naturally produced in *E. coli* cells. For protein production, after overnight pre-culture in Luria broth media at 37°C and agitation 180 rpm, the strain was grown in Terrific Broth (TB) medium supplemented with 0.5% (v/v) glycerol at 37°C, 180 rpm, to an absorbance of 1 before induction with 500  $\mu$ M IPTG. The temperature was then lowered to 18°C and the cultures were incubated for 24h. The cells were harvested by centrifugation at 8°C, 5000 rpm for 30 min and the pellets stored in a freezer at -80°C.

### **FAP purification and quantification**

The first purification step consisted in thawing the frozen pellets. The cells were resuspended in a buffer containing 300 mM NaCl, 50 mM Tris-HCl pH 8.0, 10 mM imidazole, 5% (w/v) glycerol. To allow cell lysis, lysozyme was added at a final concentration of 0.25 mg mL<sup>-1</sup>. DNA digestion was done by adding DNase and MgSO<sub>4</sub> (final concentrations 10  $\mu$ g mL<sup>-1</sup> and 20 mM, respectively). Anti-protease tablets (Sigma S8820-20 TAB) were also dissolved in the lysis buffer at the rate of 1 tablet per 100 mL of buffer (as a rule of thumb, for initially 8 liters of cultures, 500 ml of lysis buffer was used). After incubation for 1 hour at room temperature, the cells were treated by sonication to improve bacterial lysis. Sonication parameters were: 20 kHz, 4 cycles of 45 s with 15 s pause and agitation to avoid overheating. Soluble proteins were recovered after centrifugation at 11 000 g for 30 min at 4°C. FAP was purified by affinity chromatography using a nickel column (Ni-NTA) (elution buffer: 5% (w/v) glycerol, 300 mM NaCl, 50 mM Tris-HCl pH 8.0, and 250 mM imidazole. To get rid of the histidine tag in fusion

with thioredoxin, the eluate was then digested for at least 1 h at room temperature using Tobacco Etch Virus protease (TEV) (1 mg of TEV for 20 mg of protein to be digested). The digestate was dialyzed (membrane reference: Spectra por/Standart Rc tubing MWCO: 12-14 kD) overnight against 300 mM NaCl, 50 mM Tris pH 8.0, 10 mM imidazole, 5% (w/v) glycerol. The FAP protein was then separated from the previously cut His tag by a second affinity chromatography using a Nickel column. Final purification step by gel filtration using (Superdex200 26/600 mm GE Healthcare) was necessary to separate any aggregates from the soluble protein. Buffer used for this step contained 150 mM NaCl, 10 mM Tris pH 8.0, 5% (w/v) glycerol. Generally, after all the purification steps, we obtained a protein of >95% purity. The purified protein was quantified based on the absorbance at 280 nm.(13) To determine the amount of active protein (*i.e.*, protein containing FAD), the FAD concentration was estimated by dividing the absorbance at the maximum of the oxidized FAD band in the blue (467 nm)(13), by the typical molar extinction coefficient of FAD in flavoproteins:  $11\,300\text{ M}^{-1}\text{cm}^{-1}$ (21). Typically, 70 to 85% FAP proteins contained bound FAD. After purification, the protein was concentrated using ultracentrifuge filters 50 kDa Amicon® up to  $20\text{ mg.mL}^{-1}$  flash frozen in liquid nitrogen and stored at  $-80\text{ }^{\circ}\text{C}$ .

### **Time-resolved fluorescence (TRF)**

Fluorescence kinetics were monitored on a set-up described previously(1, 13), using ‘weak’ flashes from a Nd:YAG laser as the excitation light source (Continuum Leopard SS-10, pulse duration of 100 ps, 355 nm, energy attenuated to  $\sim 20\text{ }\mu\text{J per cm}^2$ , repetition rate 2 Hz). ‘Strong’ ( $\sim 10\text{ mJ}$ ) flashes at 470 nm of  $\sim 5\text{ ns}$  pulse duration used to consume the native and/or the added substrates were delivered by a Nd:YAG-pumped optical parametric oscillator (Brillant B/Rainbow, Quantel, France). The detection system consisted of a Hamamatsu microchannel plate photomultiplier tube R2566U-11P (under 3.6 kV) connected to a digital oscilloscope (Tektronix MSO64, bandwidth limit set to 6 GHz). The samples were contained in a  $2\times 2\times 10\text{ mm}$  cell with self-masking solid black walls and four clear windows (Starna). Excitation pulses entered through the  $2\times 10\text{ mm}$  window and fluorescence was detected in front of the  $2\times 2\text{ mm}$  window). An 1OG-530 Schott orange glass optical filter and a 560 nm interference filter (with a 10 nm bandwidth) were placed between the sample and the detector. The fluorescence emission spectrum of  $^1\text{FAD}^*$  spreads between 480 and 620 nm, with maxima typically around 530 nm. 560 nm are still well within the  $^1\text{FAD}^*$  emission band (with  $\sim 80\%$  of the maximum intensity) but far enough from 532 nm – the second harmonic of the used Nd:Yag laser, traces of which could have been present also in our 355 nm pulses. The wavelength of 560 nm was

hence chosen to avoid possible contaminations of the fluorescence signals by photons from the excitation source.

Fluorescence traces are averages of 64 signals. ~30  $\mu$ M active (FAD-containing) CvFAP protein was used in the presence (or absence) of ~300  $\mu$ M substrates, ~3 mM *n*-alkane co-catalysts and 5% ethanol (substrates and alkanes were dissolved in ethanol and added to the aqueous solution of CvFAP prior to each experiment; for consistency, ethanol was added (to the final concentration of 5%) also to the samples without added substrates and alkanes).

Normalized TRF signals were fitted using the Levenberg-Marquardt least-square optimization algorithm in Origin 2020 (by OriginLab). The signals with two pronounced phases (~300 ps and ~5ns) could be reasonably fitted by a biexponential function  $y(t) = A_1 \times e^{-t/\tau_1} + A_2 \times e^{-t/\tau_2}$ , where  $A_1$  and  $A_2$  are the amplitudes of the fast and of the slow phase, respectively (their sum should be equal to 1 since the signals are normalized), and  $\tau_1$  and  $\tau_2$  are their respective time constants;  $t$  stands for time (in seconds). However, when the fast phase was small relative to the slow one, the inherent oscillatory artifact at the beginning of our signals (natural response of the photomultiplier to a rapid (~100 ps) change in light intensity) made a correct fit of the fast phase impossible (see Fig. 5A in ref. (1) for the instrument response signal and SI text *ibid.* for more details). Therefore, we fitted only the slow phase, starting the fit at  $t = 1$  ns, *i.e.*, after the end of the ~300 ps process (more than 3-times its time constant), using a simple (monoexponential) function  $y(t) = A_2 \times e^{-t/\tau_2}$ .  $A_1$  was subsequently calculated using  $A_1 = 1 - A_2$ . We verified on several signals with pronounced fast phases that this method yielded the same results for  $A_1$  and  $A_2$  as the biexponential fit of the complete curves fitted from  $t = 0$  ns (with an error of max. 2%). For simplicity, the amplitudes are recalculated to % in the figures ( $A_1 + A_2 = 1 = 100\%$ ).

### **Transient absorption spectroscopy (TAS)**

Transient absorption kinetics were also recorded on a set-up described previously(1, 13). The FAD cofactor in FAP was excited at 470 nm by laser flashes identical to the ‘strong’ flashes in the TRF experiment (*i.e.*, ~5 ns pulse duration and an energy in the order of 10 mJ per cm<sup>2</sup> delivered by a Nd:YAG-pumped optical parametric oscillator Brilliant B/Rainbow, Quantel, France). The monitoring light at 515 nm was provided by the DPSS laser Cobolt Fandango<sup>TM</sup> (150 mW). The sample cell was identical to that used in the TRF experiments (2×2×10 mm). The monitoring light was attenuated by neutral density filters and mechanically chopped to produce a rectangular pulse of ~140  $\mu$ s duration and an energy in the order of 1  $\mu$ J at the

entrance of the cell (2×2 mm window), thus avoiding significant actinic effects. This pulse was synchronized with the excitation laser flash entering the sample through the 2×10 mm window. The signals were recorded using an Alphalas UPD-500-UP photodiode (rise time < 500 ps; sensitive area 0.5 mm<sup>2</sup>) connected via a Femto HCA electronic signal amplifier (DC-325 MHz, 28 dB) to the Tektronix MSO64 digital oscilloscope with bandwidth limit set to 200 MHz. The samples used in TAS experiments were prepared in the same way as those used in the TRF experiments (see above). For consistency with the TRF experiments (to allow the substrates and co-catalysts the same time to reorganize after the strong flashes), the excitation flashes were separated by delays of ~2 minutes (about the time needed for one fluorescence measurement).

### **MD simulations**

The structure of CvFAP was taken from the Protein Data Bank (PDB entry: 6ZH7; resolution: 2.0 Å). In this structure, two molecules of C18 FA (presumably unsaturated) are present; the one located at the surface of the enzyme close to the entrance to the tunnel leading to the binding site was removed prior to simulations. The system with saturated linear C8 FA and C7 alkane was constructed by removing the 'C3-tail' and the covalent bond from C18 FA in the crystal structure and adding two hydrogens to complete the methyl groups of C8 FA and C7 alkane. To construct the system with C8 FA and C10 alkane, a similar approach was used, but no carbon atoms were removed. Three simulations of CvFAP with C8 FA alone were performed with the initial structure taken from the equilibrated trajectory of the simulation with C8 FA and C10 alkane (with C10 alkane removed). The results of the latter were used in discussion, as the results of all three simulations were very similar.

The CHARMM36 force field was used for the protein residues(22), the modified version of the TIP3P model for water molecules(23), and a recently developed force field for flavins for the FAD cofactor(24). The parameters for C8, C15 and C18 FAs and for C7 and C10 alkanes were taken from the CHARMM General Force Field (CGenFF, version 2.2.0)(25). The protonation states of all titratable residues were assigned based on a PROPKA 3.1 analysis(26) and verified by ideal stereochemistry taking into account steric effects and potential hydrogen-bonding interactions.

The systems were then centered in a cubic box of aqueous solvent with an appropriate size, at least 12 Å away from each of the box edges; thus, the final systems in addition to the CvFAP protein, FAD, and substrates contained around 30,000 water molecules. MD simulations were performed using the NAMD program (version 2.13)(27). Periodic boundary conditions were

assumed with long-range electrostatic interactions computed using the particle mesh Ewald method(28), and an appropriate number of potassium counterions was included to neutralize the net charge of the systems. The integration time step was set to 2 fs. After energy minimization, the system was equilibrated first in an NVT ensemble for 50 ps, followed by a 500 ps simulation in the NPT ensemble, at 295 K and 1.0 atm pressure. The Berendsen thermostat and barostat were employed, with a relaxation time of 500 fs and four timesteps between position rescalings for constant pressure simulations(29). The production runs were then performed, and coordinates of the systems were collected every 100 ps.

### ***In vitro* photoenzymatic production of *n*-alkanes**

*In vitro* assay was performed using the purified CvFAP. Teorell-Stenhagen buffer of the following composition: 33 mM citric acid monohydrate, 33 mM phosphoric acid and 16.7 mM boric acid was used for the enzymatic reaction. The buffer was mixed with pure FA solution before adjusting the pH to the desired value using NaOH (10M) or HCl (6N) solutions. The protein solution was then added in the dark from a 1000-fold concentrated stock to obtain 70 nM of protein in 5 ml of final reaction volume (contained in 10 ml vials, *i.e.*, 5 ml liquid and 5 ml gas phase). The samples were then exposed to light at 360  $\mu\text{moles photons/m}^2/\text{s}$  of blue LED light (450 nm; FWHM: 20 nm) at 25°C for 20 min to induce photodecarboxylation.

### **Analysis and quantification of C7 alkane**

After C7 alkane production under light, the reaction was stopped by the enzyme denaturation at 100°C for 10 min. The samples were then cooled at room temperature for 30 min to recondense all volatile alkanes. Then, the samples were re-heated to 40°C and kept at this temperature for 5 min. 1 ml of gas phase was collected with a syringe heated at 80°C and injected into a GC-MS/FID (AGILENT Technologies; 5977B; MSD) for analysis using helium as carrier gas. The analysis parameters were as follows: oven initial temperature: 50°C for 1 min, ramp: 20°C/min to 260°C for 5 min (Column reference: 19091P-Q04PT, PH-PLOT/Q+RT, 30 m  $\times$  0.320 mm, 20 micron). The absolute amount of C7 alkane was calculated from an external standard curve obtained using pure C7 alkane solutions. The curves were processed under the same conditions as those used for the test reactions.

### **Analysis and quantification of C15 alkane**

After C15 alkane production under light, 100  $\mu\text{l}$  of NaOH (10M) were added before heating at 100°C during 10 min. 4 ml of hexane and 10  $\mu\text{g}$  of *n*-hexadecane (C16 alkane, internal standard for C15 alkane quantification) were added. The samples were then vortexed and centrifuged

(3000 rpm, 5 min). The organic phase containing the extracted C15 alkane was collected and 1  $\mu$ L was injected in GC-MS/FID for analysis according to the following program: Oven initial temperature: 60°C for 1 min, ramp: 20°C/min to 150°C; 10°C/min to 260°C for 2 min. Column reference : OPTIMA WAXplus, 30 m  $\times$  0.25 mm. The absolute amount of C15 alkane was calculated based on the internal standard added before the sample treatment.

### ***In vivo* photoenzymatic production of *n*-alkanes**

The production of *n*-alkanes was performed using the *E. coli* BL21 strain containing pLIC07FAPv2 plasmid. A pre-culture was grown overnight in Luria-Bertani (LB) medium at 37°C, 180 rpm. For protein production, the strain was cultured in Terrific Broth (TB) medium supplemented with 0.5% (w/v) glycerol at 37°C, 180 rpm, to an absorbance of 1 before induction with 500  $\mu$ M IPTG and addition of 1-<sup>13</sup>C-labeled FA substrates at 2 mM (The concentrated stock solution of C8 and C16 FAs was previously prepared in ethanol). The temperature was then lowered to 18°C and the cultures were incubated in the dark for 24 h. 5 mL of each culture were then put in sealed 10 mL airtight vials and illuminated at 360  $\mu$ mol photons/m<sup>2</sup>/s of blue LED light (450 nm) at 25°C for 1 h to induce photodecarboxylation.

### **Analysis and quantification of <sup>13</sup>CO<sub>2</sub>**

For <sup>13</sup>CO<sub>2</sub> analysis after the illumination, the reaction was stopped by heating the samples to 100 °C for 10 min (denaturing the enzymes and lysing the cells to release all trapped volatiles). After cooling the samples, hydrochloric acid was added to the vials to lower the pH to 1 and convert all dissolved bicarbonate to CO<sub>2</sub>. Then the samples were re-heated to 40°C and kept at this temperature for 5 min. 1 mL of gas was collected and injected into a GC-MS/FID (AGILENT Technologies; 5977B; MSD) for analysis. The analysis parameters were as follows: oven initial temperature: 50°C for 1 min, ramp: 20°C/min to 260°C for 5 min (Column reference: 19091P-Q04PT, PH-PLOT/Q+RT, 30 m  $\times$  0.320 mm, 20 micron) using helium as carrier gas.

### **<sup>13</sup>C-labeled substrate quantification**

To quantify the amount of the remaining substrates which were not incorporated into the cells after 24 hours of culture incubation, 1 mL of each were centrifuged. Then, 10  $\mu$ g of saturated C17 FA methyl ester (standard) were added to 50  $\mu$ L of the supernatant (note that the cells were also washed to dissolve any remaining FAs at the membrane surface). To perform the transmethylation of the samples, 2 mL of sulfuric acid / methanol (5/100, v/v) were added before



heating at 85°C for 90 minutes. The samples were then cooled to room temperature, 2 ml of NaCl (0.9%) and 500 µl of hexane were added successively. To extract the transmethyated fatty acids, the samples were shaken during 10 min and centrifuged at 3000 rpm for 5 min. 1 µl of the hexane phase was injected into a GC-MS/FID (AGILENT Technologies; 5977B; MSD) using helium as carrier gas. The analysis parameters were as follows: oven initial temperature 20°C for 1 min, ramp 20°C/min to 300°C for 1 min using helium as gas carrier. Column reference: 190915-433UI, HP-5MS UI, 30 m × 0.250 mm, 0.25 micron). The results shown in Fig. 5b are averages of values obtained in three independent samples normalized by the absorbance. The absolute amount of C8 and C16 FA methyl ester was calculated on the basis of the internal standard added before the sample treatment.

## References

1. D. Sorigué, B. Légeret, S. Cuiné, S. Blangy, S. Moulin, E. Billon, P. Richaud, S. Brugière, Y. Couté, D. Nurizzo, P. Müller, K. Brettel, D. Pignol, P. Arnoux, Y. Li-Beisson, G. Peltier, F. Beisson, An algal photoenzyme converts fatty acids to hydrocarbons. *Science* **357**, 903-907 (2017).
2. D. Ramírez-Gamboa, A. L. Díaz-Zamorano, E. R. Meléndez-Sánchez, H. Reyes-Pardo, K. R. Villaseñor-Zepeda, M. E. López-Arellanes, J. E. Sosa-Hernández, K. G. Coronado-Apodaca, A. Gámez-Méndez, S. Afewerki, H. M. N. Iqbal, R. Parra-Saldivar, M. Martínez-Ruiz, Photolyase Production and Current Applications: A Review. *Molecules (Basel, Switzerland)* **27**, (2022).
3. D. J. Heyes, S. Zhang, A. Taylor, L. O. Johannissen, S. J. O. Hardman, S. Hay, N. S. Scrutton, Photocatalysis as the ‘master switch’ of photomorphogenesis in early plant development. *Nat. Plants* **7**, 268-276 (2021).
4. M. M. E. Huijbers, W. Zhang, F. Tonin, F. Hollmann, Light-Driven Enzymatic Decarboxylation of Fatty Acids. *Angew. Chem. Int. Ed.* **57**, 13648-13651 (2018).
5. I. S. Yunus, J. Wichmann, R. Wördenweber, K. J. Lauersen, O. Kruse, P. R. Jones, Synthetic metabolic pathways for photobiological conversion of CO<sub>2</sub> into hydrocarbon fuel. *Metab. Eng.* **49**, 201-211 (2018).
6. S. Moulin, B. Légeret, S. Blangy, D. Sorigué, A. Burlacot, P. Auroy, Y. Li-Beisson, G. Peltier, F. Beisson, Continuous photoproduction of hydrocarbon drop-in fuel by microbial cell factories. *Sci. Rep.* **9**, 13713 (2019).
7. S. Bruder, E. J. Moldenhauer, R. D. Lemke, R. Ledesma-Amaro, J. Kabisch, Drop-in biofuel production using fatty acid photodecarboxylase from *Chlorella variabilis* in the oleaginous yeast *Yarrowia lipolytica*. *Biotechnol. Biofuels* **12**, 202 (2019).
8. J. Xu, J. Fan, Y. Lou, W. Xu, Z. Wang, D. Li, H. Zhou, X. Lin, Q. Wu, Light-driven decarboxylative deuteration enabled by a divergently engineered photodecarboxylase. *Nat. Commun.* **12**, 3983 (2021).
9. J. Xu, Y. Hu, J. Fan, M. Arkin, D. Li, Y. Peng, W. Xu, X. Lin, Q. Wu, Light-Driven Kinetic Resolution of  $\alpha$ -Functionalized Carboxylic Acids Enabled by an Engineered Fatty Acid Photodecarboxylase. *Angew. Chem. Int. Ed.* **58**, 8474-8478 (2019).

10. W. Zhang, J. H. Lee, S. H. H. Younes, F. Tonin, P. L. Hagedoorn, H. Pichler, Y. Baeg, J. B. Park, R. Kourist, F. Hollmann, Photobiocatalytic synthesis of chiral secondary fatty alcohols from renewable unsaturated fatty acids. *Nat. Commun.* **11**, 2258 (2020).
11. H. J. Cha, S. Y. Hwang, D. S. Lee, A. R. Kumar, Y. U. Kwon, M. Voß, E. Schuiten, U. T. Bornscheuer, F. Hollmann, D. K. Oh, J. B. Park, Whole-Cell Photoenzymatic Cascades to Synthesize Long-Chain Aliphatic Amines and Esters from Renewable Fatty Acids. *Angew. Chem. Int. Ed.* **59**, 7024-7028 (2020).
12. D. Sorigué, B. Légeret, S. Cuiné, P. Morales, B. Mirabella, G. Guédeney, Y. Li-Beisson, R. Jetter, G. Peltier, F. Beisson, Microalgae Synthesize Hydrocarbons from Long-Chain Fatty Acids via a Light-Dependent Pathway. *Plant Physiol.* **171**, 2393-2405 (2016).
13. D. Sorigué, K. Hadjidemetriou, S. Blangy, G. Gotthard, A. Bonvalet, N. Coquelle, P. Samire, A. Aleksandrov, L. Antonucci, A. Benachir, S. Boutet, M. Byrdin, M. Cammarata, S. Carbajo, S. Cuiné, R. B. Doak, L. Foucar, A. Gorel, M. Grünbein, E. Hartmann, R. Hienerwadel, M. Hilpert, M. Kloos, T. J. Lane, B. Légeret, P. Legrand, Y. Li-Beisson, S. L. Y. Moulin, D. Nurizzo, G. Peltier, G. Schirò, R. L. Shoeman, M. Sliwa, X. Solinas, B. Zhuang, T. R. M. Barends, J.-P. Colletier, M. Joffre, A. Royant, C. Berthomieu, M. Weik, T. Domratcheva, K. Brettel, M. H. Vos, I. Schlichting, P. Arnoux, P. Müller, F. Beisson, Mechanism and dynamics of fatty acid photodecarboxylase. *Science* **372**, eabd5687 (2021).
14. S. L. Y. Moulin, A. Beyly-Adriano, S. Cuiné, S. Blangy, B. Légeret, M. Floriani, A. Burlacot, D. Sorigué, P. P. Samire, Y. Li-Beisson, G. Peltier, F. Beisson, Fatty acid photodecarboxylase is an ancient photoenzyme that forms hydrocarbons in the thylakoids of algae. *Plant Physiol.* **186**, 1455-1472 (2021).
15. W. Zhang, M. Ma, M. M. E. Huijbers, G. A. Filonenko, E. A. Pidko, M. van Schie, S. de Boer, B. O. Burek, J. Z. Bloh, W. J. H. van Berkel, W. A. Smith, F. Hollmann, Hydrocarbon Synthesis via Photoenzymatic Decarboxylation of Carboxylic Acids. *J. Am. Chem. Soc.* **141**, 3116-3120 (2019).
16. C. Aselmeyer, B. Légeret, A. Bénarouche, D. Sorigué, G. Parsiegla, F. Beisson, F. Carrière, Fatty Acid Photodecarboxylase Is an Interfacial Enzyme That Binds to Lipid-Water Interfaces to Access Its Insoluble Substrate. *Biochemistry* **60**, 3200-3212 (2021).
17. D. P. Cistola, J. A. Hamilton, D. Jackson, D. M. Small, Ionization and phase behavior of fatty acids in water: application of the Gibbs phase rule. *Biochemistry* **27**, 1881-1888 (1988).
18. J. R. Kanicky, A. F. Poniatowski, N. R. Mehta, D. O. Shah, Cooperativity among Molecules at Interfaces in Relation to Various Technological Processes: Effect of Chain Length on the  $pK_a$  of Fatty Acid Salt Solutions. *Langmuir* **16**, 172-177 (2000).
19. Y. B. Vysotsky, E. S. Kartashynska, D. Vollhardt, V. B. Fainerman, Surface  $pK_a$  of Saturated Carboxylic Acids at the Air/Water Interface: A Quantum Chemical Approach. *J. Phys. Chem. C* **124**, 13809-13818 (2020).
20. Y. Wu, C. E. Paul, F. Hollmann, Stabilisation of the Fatty Acid Decarboxylase from *Chlorella variabilis* by Caprylic Acid. *ChemBioChem* **22**, 2420-2423 (2021).
21. P. Macheroux, UV-visible spectroscopy as a tool to study flavoproteins. *Meth. Mol. Biol.* **131**, 1-7 (1999).
22. J. Huang, A. D. MacKerell, Jr., CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J. Comput. Chem.* **34**, 2135-2145 (2013).
23. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926-935 (1983).

24. A. Aleksandrov, A Molecular Mechanics Model for Flavins. *J. Comput. Chem.* **40**, 2834-2842 (2019).
25. K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, A. D. Mackerell, Jr., CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **31**, 671-690 (2010).
26. M. H. Olsson, C. R. Søndergaard, M. Rostkowski, J. H. Jensen, PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical  $pK_a$  Predictions. *J. Chem. Theory Comput.* **7**, 525-537 (2011).
27. J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, K. Schulten, Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781-1802 (2005).
28. T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An  $N \cdot \log(N)$  method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089-10092 (1993).
29. H. J. C. Berendsen, J. P. M. Postma, W. F. v. Gunsteren, A. DiNola, J. R. Haak, Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684-3690 (1984).

## Acknowledgments

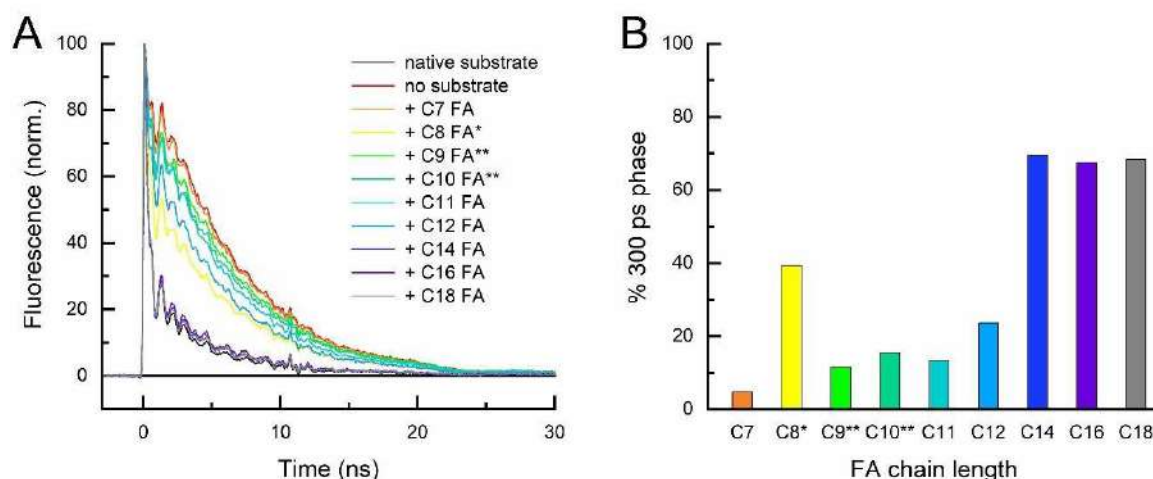
**Funding:** The present work has benefited from the platform of Biophysics of I2BC supported by the French Infrastructure for Integrated Structural Biology (FRISBI): ANR-10-INBS-05, and from the technical support of the Heliobiotec platform (BIAM Cadarache). We acknowledge ANR grants 18-CE11-0021 (FB and PM), 18-CE43-0008 (GP) and 18-CE44-0002 (AA). PPS and BZ acknowledge the French Alternative Energies and Atomic Energy Commission and the China Scholarship Council, respectively, for providing their PhD scholarships. We are grateful to Drs. Klaus Brettel and Marten Vos for stimulating discussions.

**Author contributions:** Conceptualization: PM, FB, AA. Methodology: BL, GP, AA, PM. Investigation: PPS, BZ, ÁBP, BL, GP, DS, AA, PM. Visualization: PPS, BZ, AA, PM. Supervision: BL, GP, DS. Writing original draft: PM, PPS, FB and AA with input from all co-authors.

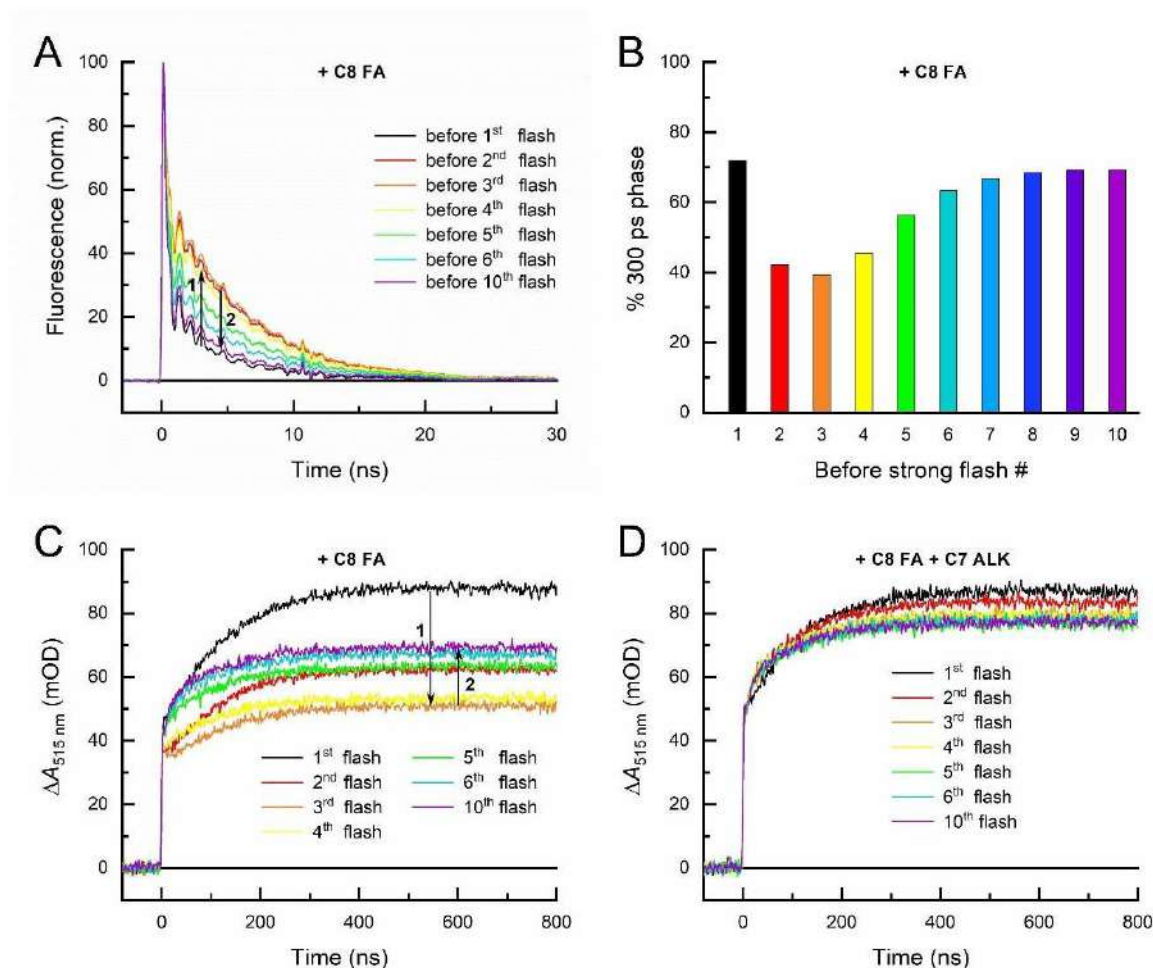
**Competing interests:** Authors declare that they have no competing interests.

**Data and materials availability:** All data are available in the main text or the Supplementary Materials.

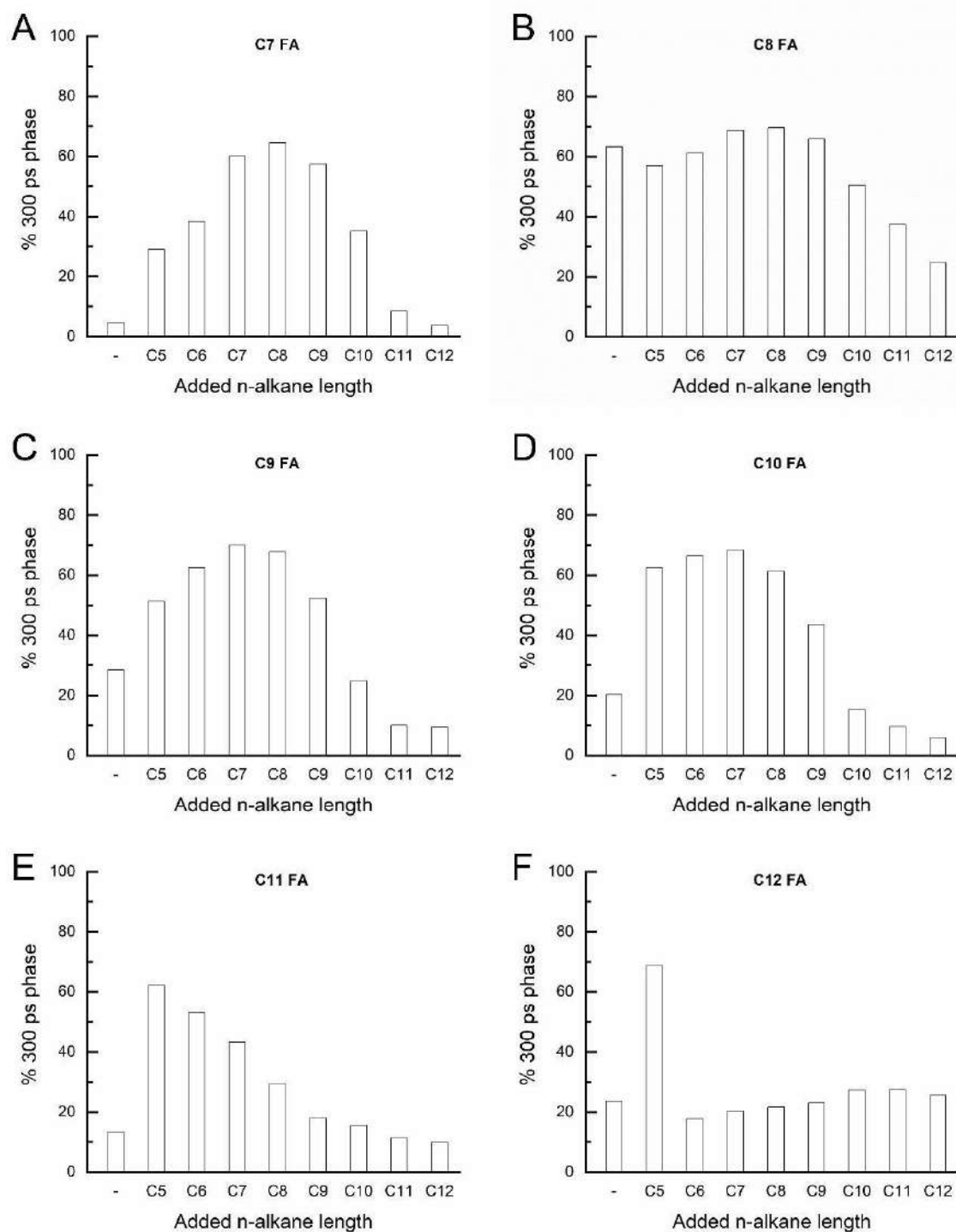
## Figures



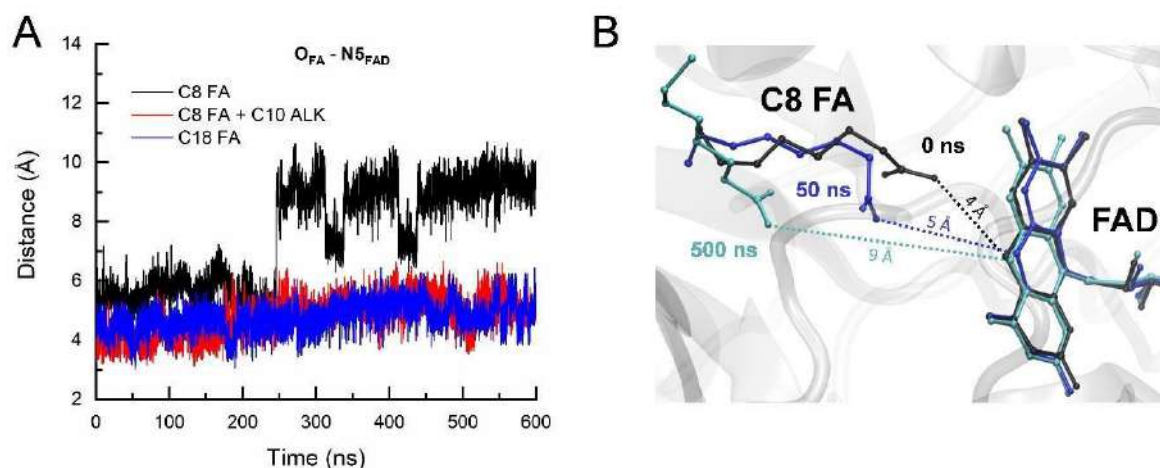
**Fig. 1. Photodecarboxylation of saturated linear C7-C18 fatty acids by CvFAP.** (A) Normalized time-resolved fluorescence (TRF) signals at 560 nm recorded for ~30  $\mu$ M CvFAP with the native substrate before its consumption, after its consumption (no substrate), and with added ~300  $\mu$ M C7 - C18 FA substrates (see legend for color code). Most of the latter signals were recorded after five strong (~10 mJ/cm<sup>2</sup>) 470 nm flashes, *i.e.*, after almost complete consumption of the strongly bound native substrate (see Figure S1). The signals for C8, C9 and C10 FAs are those for which the amplitude of the fast (~300 ps) phase was smallest (at the beginning of the onset of the autocatalytic effect, *i.e.*, before the third [\*] or the fourth [\*\*] strong flash; see Figure 2 for C8 FA). All traces are averages of 64 signals recorded upon excitation by weak (~20  $\mu$ J/cm<sup>2</sup>) flashes at 355 nm. (B) Share of the ~300 ps phase in the TRF signals with added substrates shown in panel 'A', corresponding to forward electron transfer from the substrate to the photoexcited FAD followed by the quasi-instantaneous CO<sub>2</sub> cleavage (13). The relatively high minimum share of the fast phase in the C8 FA TRF signal (recorded before the third strong flash) indicates that C8 FA itself probably also acts to some extent as a co-catalyst for the decarboxylation of another C8 FA molecule, although not as efficiently as its C7 alkane product.



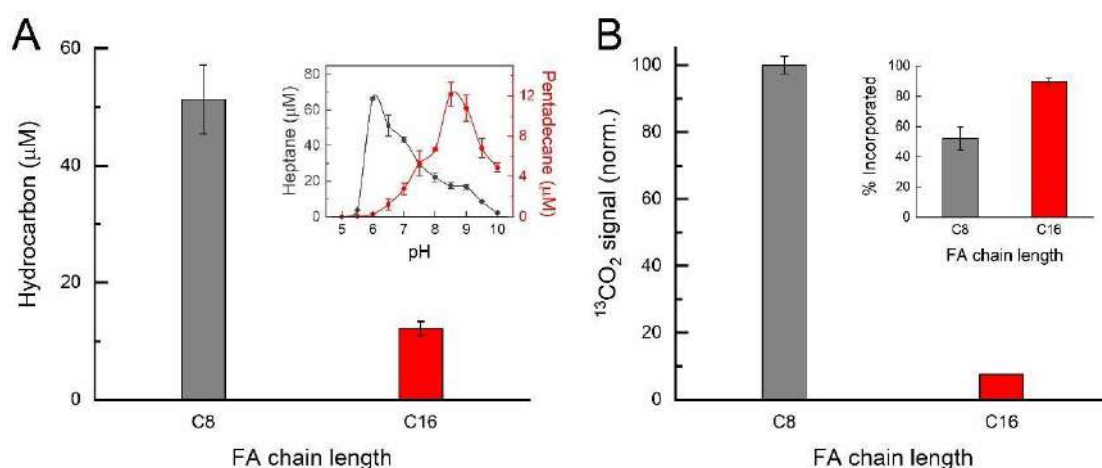
**Fig. 2. Auto-/Co-catalysis of saturated linear C8 FA photodecarboxylation by its C7 *n*-alkane product.** (A) Normalized TRF signals at 560 nm recorded for  $\sim 30\ \mu\text{M}$  CvFAP with added  $\sim 300\ \mu\text{M}$  C8 FA prior to ten strong 470 nm flashes progressively consuming the native C18 FA substrate and the added C8 FA. The first two flashes led to the decrease of the amplitude of the 300 ps phase but the subsequent flashes led to its gradual recovery (“yo-yo” effect). (B) Share of the  $\sim 300$  ps phase in the TRF signals with added substrates shown in panel ‘A’ reflecting the quantum yield of photodecarboxylation. The initial flashes consuming most of the native substrate and also small amounts of C8 FA result in a gradual onset of autocatalysis by the formed C7 alkane product. (C) Transient absorption changes at 515 nm on the sub-microsecond timescale recorded upon the strong 470 nm flashes. The step-like initial increase (reflecting the reduction of FAD to the  $\text{FAD}^{\bullet-}$  radical in  $\sim 300$  ps) is followed by further growth with a time constant of  $\sim 100$  ns, in line with the back ET and the formation of red-shifted reoxidized flavin ( $\text{FAD}_{\text{RS}}$ ) (1, 13). The “yo-yo” effect is clearly visible also in the TAS signals, although the amplitude recovery seems less complete than in the case of TRF signals. This could, in principle, be either due to a slight difference in the magnitude of the  $\text{FAD}_{\text{RS}}$  spectral shift depending on the substrate/product length, or due to partial recombination of the alkyl radical with  $\text{CO}_2$  upon the back ET from  $\text{FAD}^{\bullet-}$  (or a combination of both). The TRF/TAS data recorded for CvFAP with added C16 FA tend to support the substrate-/product-dependent redshift magnitude (see fig. S2). (D) TAS signals at 515 nm recorded for a sample with the same CvFAP and C8 FA concentrations as the one used for experiments shown in panels ‘A’ and ‘C’, but in the presence of additional  $\sim 3\ \text{mM}$  C7 *n*-alkane. With C7 alkane present from the start, the “yo-yo” effect disappears.



**Fig. 3. Effects of C5-C12 *n*-alkane co-catalysts on the photodecarboxylation of saturated linear C7-C12 FAs by CvFAP.** (A–F) Share of the ~300 ps phase in the TRF signals at 560 nm recorded for ~30  $\mu$ M CvFAP with added ~300  $\mu$ M C7-C12 FAs in the absence (-) and in the presence of ~3 mM C5-C12 *n*-alkanes as co-catalysts after five strong 470 nm flashes (*i.e.*, after the consumption of the native substrate).



**Fig. 4. MD Simulations.** (A) Dynamics of the distances between the carboxyl O atoms of FAs and the N5 atom of the FAD isoalloxazine ring during the MD simulations of FAP in complex with C8 FA alone (black trace), C8 FA together with C10 alkane (red trace), and C18 FA (blue trace), respectively. (B) Snapshots of the active-site structure from the MD simulation of FAP with C8 FA alone taken at selected simulation time points. The dotted lines show the evolution of the distance followed in panel 'A'. See fig. S3 for more data.



**Fig. 5. *In vitro* and *in vivo* decarboxylation of saturated linear C8 & C16 FAs by CvFAP under continuous light.** (A) CvFAP activity on C8 and C16 FAs *in vitro* expressed in terms of concentrations of the corresponding hydrocarbon products obtained at concentrations close to their respective decarboxylation efficiency saturation levels (~25 mM C8 and ~0.5 mM C16 FA; see fig. S5) and at the optimal pH (6 for C8 and 8.5 for C16 FA; see inset) under identical illumination conditions. Inset: pH dependence of CvFAP activity on C8 FA (dark gray curve) and on C16 FA (red curve). (B) CvFAP activity on C8 and C16 FAs *in vivo* (in living *E. coli* bacteria) based on amplitudes of the  $^{13}\text{CO}_2$  signals obtained by GC-MS analysis and normalized to the C8 signal. Inset: Incorporation of C8 and C16 FAs into the *E. coli* cells after 24 h dark incubation at 18°C. The error bars show the standard deviation obtained performing the experiments with three independent samples.



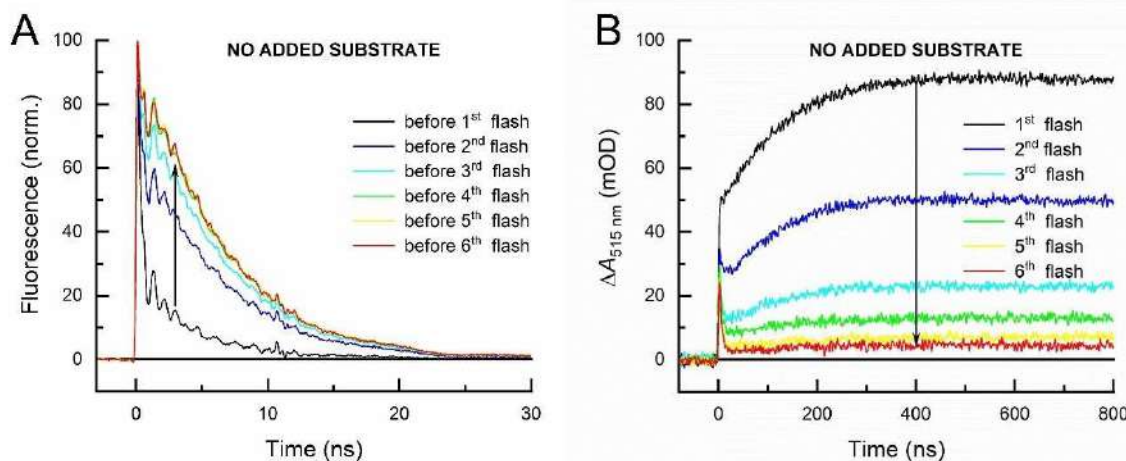


## Supplementary Materials for

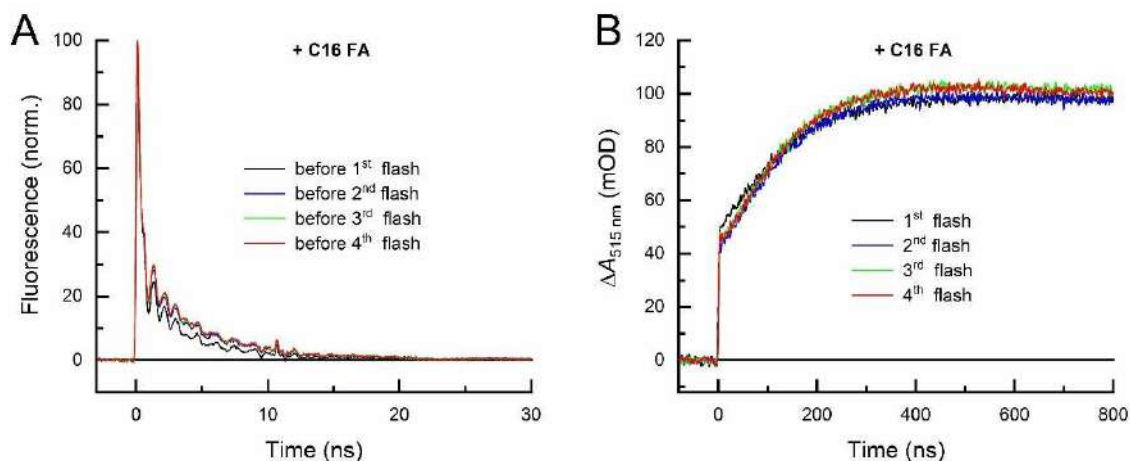
### **Autocatalytic Effect Boosts the Production of Medium-Chain Hydrocarbons by Fatty Acid Photodecarboxylase**

Poutoum P. Samire *et al.*

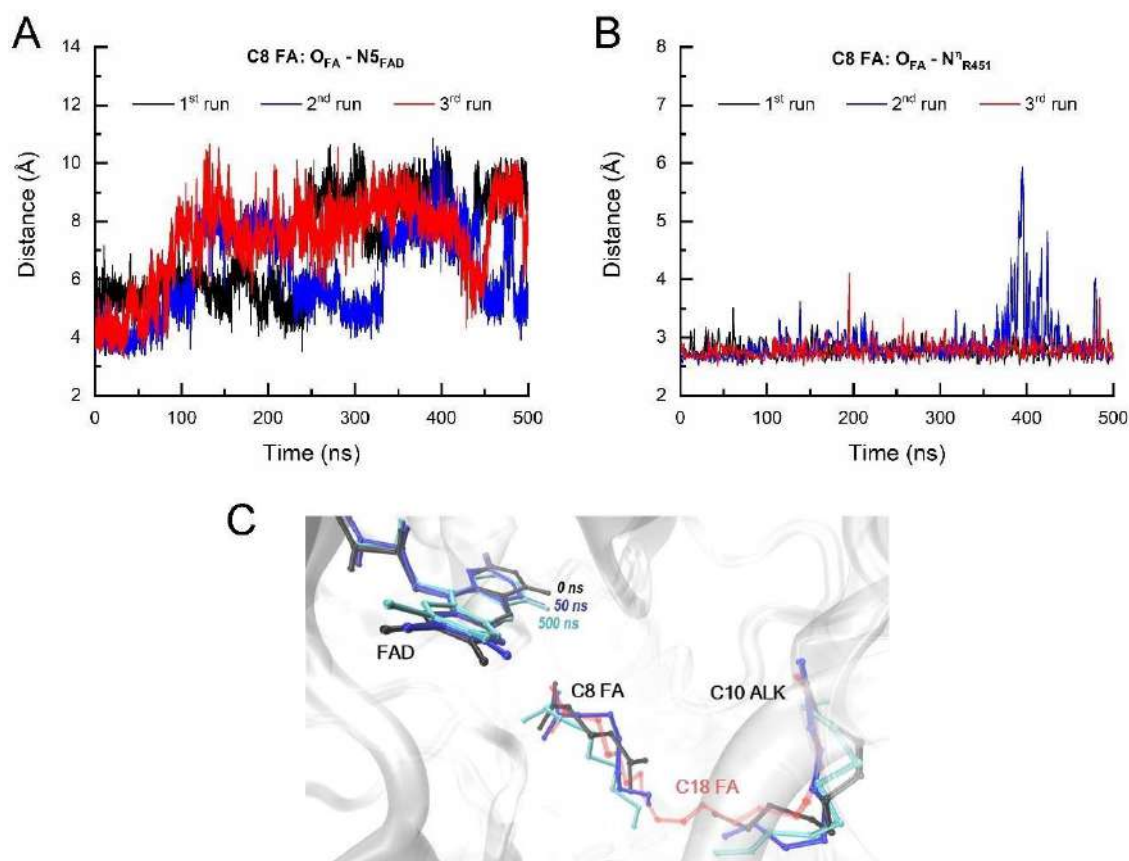
Corresponding authors: Pavel Müller, [pavel.muller@i2bc.paris-saclay.fr](mailto:pavel.muller@i2bc.paris-saclay.fr); Frédéric Beisson, [frederic.beisson@cea.fr](mailto:frederic.beisson@cea.fr); Alexey Aleksandrov, [alexey.aleksandrov@polytechnique.edu](mailto:alexey.aleksandrov@polytechnique.edu).



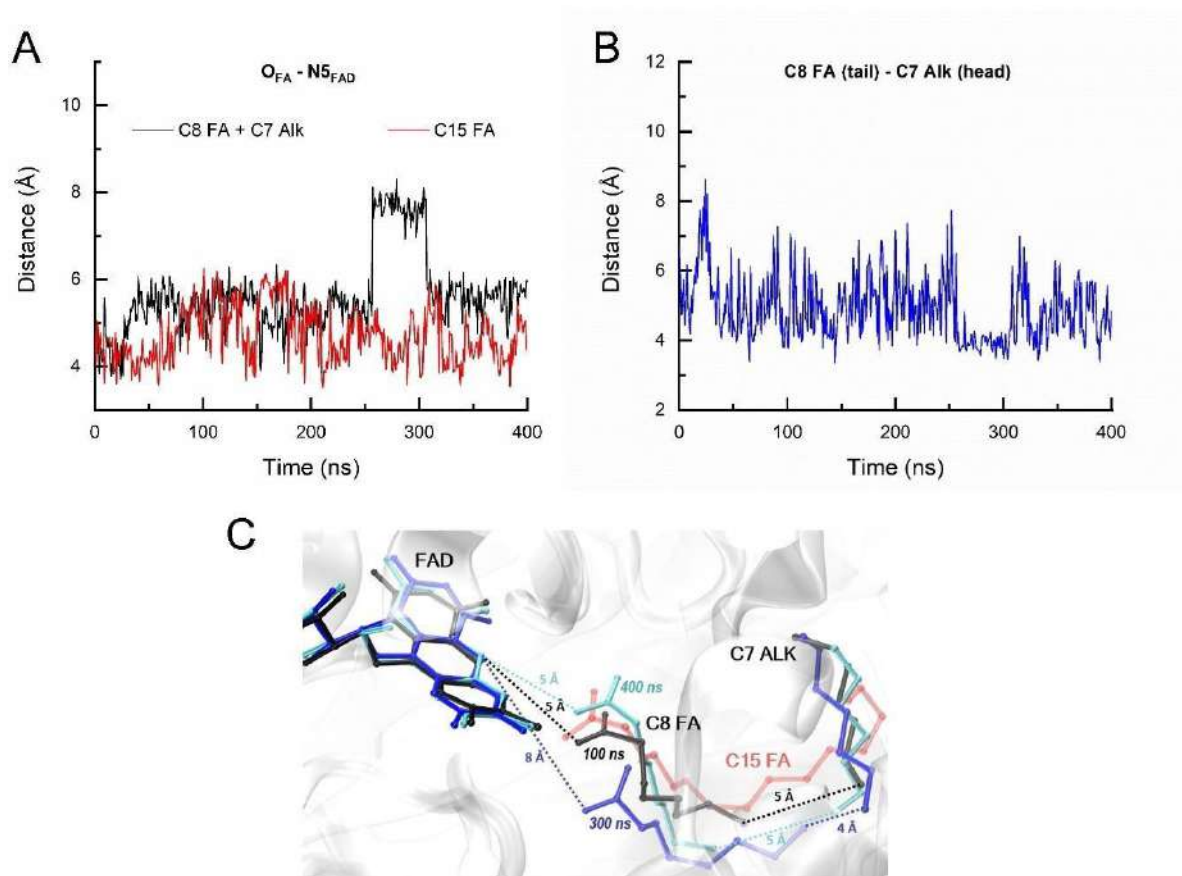
**Fig. S1. TRF and TAS signals of CvFAP with the native C18 substrate.** (A) Normalized TRF signals at 560 nm recorded for  $\sim 30 \mu\text{M}$  CvFAP without an added substrate, *i.e.*, containing maximum two native C18 FAs per enzyme (likely 1 to 1.5 FAs per FAP) prior to six strong ( $\sim 10 \text{ mJ/cm}^2$ ) 470 nm flashes progressively consuming the native substrate(s). (B) Transient absorption changes at 515 nm on the sub-microsecond timescale recorded for  $\sim 30 \mu\text{M}$  CvFAP without added substrate upon the individual strong excitation flashes. Note that the consumption of the native substrate in the TRF experiment is somewhat accelerated compared to the TAS experiment. This is because some substrate is also consumed upon the 64 weak ( $\sim 20 \mu\text{J/cm}^2$ ) flashes at 355 nm used to obtain the fluorescence signals (between the strong 470 nm flashes).



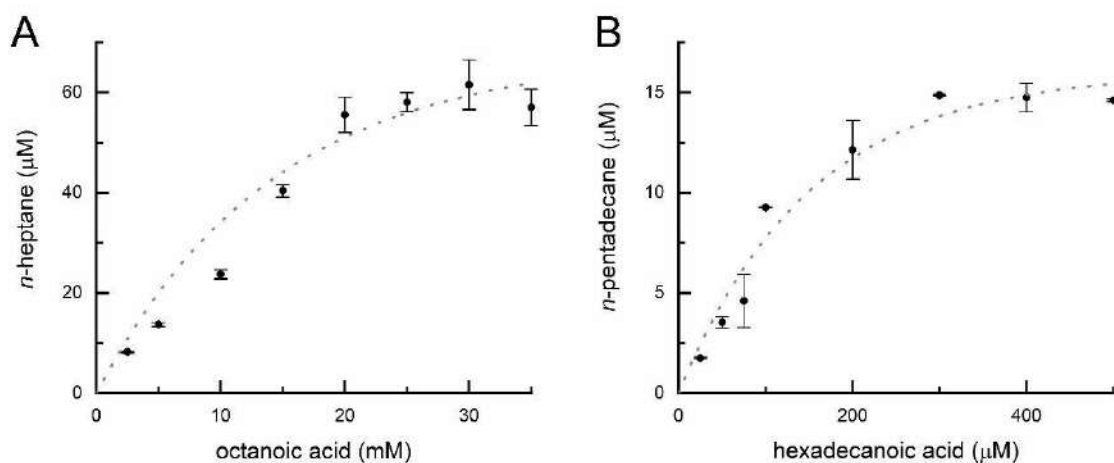
**Fig. S2. TRF and TAS signals of CvFAP with added C16 substrate.** (A) Normalized TRF signals at 560 nm recorded for  $\sim 30 \mu\text{M}$  CvFAP in the presence of  $\sim 300 \mu\text{M}$  C16 FA prior to four strong ( $\sim 10 \text{ mJ/cm}^2$ ) 470 nm flashes consuming the native substrate. The share of the  $\sim 300 \text{ ps}$  phase and hence the yield of C16 decarboxylation appear to be only slightly lower than for the native substrate. (B) Transient absorption changes at 515 nm on the sub-microsecond timescale recorded for  $\sim 30 \mu\text{M}$  CvFAP with  $\sim 300 \mu\text{M}$  C16 FA upon the strong 470 nm flashes. Note that although the TRF signals indicate a slightly lower yield of forward ET (consistent with the slightly lower initial stepwise growth of the TAS signal reflecting the  $\text{FAD}^{\bullet-}$  radical formation), the amplitude of the subsequent  $\sim 100 \text{ ns}$  growth phase (reflecting  $\text{FAD}_{\text{RS}}$  formation) is larger for C16 than for the native substrate, which is indicative of the dependence of the magnitude of the  $\text{FAD}_{\text{RS}}$  spectral shift on the nature (saturation and/or length) of the substrate/product.



**Fig. S3. MD Simulations of CvFAP containing C8 FA alone, C8 FA + C10 alkane, or C18 FA.** Dynamics of the distances between: (A) the carboxyl O atoms of C8 FA and the N5 atom of the FAD isoalloxazine ring, and (B) the carboxyl O atoms of C8 FA and the N<sub>η</sub> atom of the arginine R451, during the MD simulations of FAP in complex with C8 FA alone (without a co-catalyst) in three independent runs starting with different initial conditions. The 1<sup>st</sup> run is identical to the one shown in Fig. 4A of the main text (black trace). (C) Snapshots from the simulation of C8 FA + C10 alkane (red trace in Fig. 4A) showing their relative positions at selected simulation time points. Native C18 FA after relaxation of the crystal structure (PDB entry: 6ZH7; resolution: 2.0 Å) is shown for comparison (light red).



**Fig. S4. MD Simulations of CvFAP containing C8 FA + C7 alkane vs. C15 FA.** Dynamics of the distances between: (A) the carboxyl O atoms of C8 FA and/or C15 FA and the N5 atom of the FAD isoalloxazine ring, and (B) the tail (the last carbon atom) of C8 FA and the head (the first carbon atom) of C7 alkane in the course of the corresponding MD simulations. (C) Snapshots from the simulation of C8 FA + C7 alkane. The corresponding C15 after relaxation is shown for comparison (light red).



**Fig. S5. Search for the *in-vitro* concentration optima/saturation levels for C8 and C16 FAs.** (A) Yield of *n*-heptane as a function of added C8 substrate in the *in-vitro* activity test at pH 6 (optimum pH

for C8 FA). **(B)** Yield of *n*-hexadecane as a function of added C16 substrate in the *in-vitro* activity test at pH 8.5 (optimum pH for C16 FA). CvFAP concentration in these tests was ~70 nM. The samples were exposed to continuous light from a blue LED (emission maximum centered at 450 nm) of the intensity of 360  $\mu\text{moles photons/m}^2/\text{s}$ . Near-saturation concentrations of the individual substrates were used for the final comparison experiment shown in Fig. 5A. The error bars show the standard deviation obtained performing the experiments using three independent samples.



## **Chapter 5: Optimization of short-chain fatty acids photoconversion**





# Introduction: context and objectives

FAP represents a promising enzyme for a bio-based alkane production. Unfortunately, the use of this enzyme in industrial processes faces some limitations. Some are directly linked to the enzyme, such as its relatively low turnover on long-chain fatty acids-about 0.9 per second on palmitic acid- (Sorigué et al. [2017](#))- or its sensitivity to light-induced aggregation (Moulin et al. [2019](#)) and to photoinactivation (Lakavath et al. [2020](#)). Other limitations may come from the fact that CvFAP is an interfacial enzyme with a potential interface recognition site (IRS) and is more active on LC fatty acids present in liposomes and at the surface of microemulsions than in random bulk dispersions (Aselmeyer et al. [2021](#)).

In order to improve the photostability and/or the turnover of FAP, enzyme engineering approaches may be used. Several studies report the use of directed mutagenesis and rational design techniques to produce CvFAP variants that are more active on SC (Amer et al. [2020](#)) or on LC (Santner et al. [2021](#)) fatty acids. These approaches may be useful, but we should keep in mind that they first require basic preliminary studies aiming at characterizing better the non-mutated CvFAP and at optimizing the conditions of activity. Optimal activity conditions may be particularly useful to isolate mutants that are really more active than the WT enzyme. For example, in our previous study (Samire et al. [2023](#)), we have shown that under specific conditions, *n*-octanoic acid (C8 fatty acid) is more efficiently converted to hydrocarbons by CvFAP than palmitic acid (*n*-hexadecanoic acid, C16 fatty acid), which was described so far as one of the best substrates of FAP. This shows that engineering CvFAP to be more active on MC fatty acids may not be a priority or should use screening procedures that consider optimal activity conditions.

Depending on chain length, concentration and pH, monomeric fatty acids can self-organize into various forms in solution, such as vesicles, micelles, lamellar phases, or oil microdroplets. Since CvFAP has been shown to be an interfacial enzyme (Aselmeyer et al. [2021](#)), it is clear that the form of substrate organization in solution can have a major impact on the enzymatic activity of CvFAP. The relatively high activity we previously measured on *n*-octanoic acid encouraged us to characterize CvFAP activity on SC fatty acids, mainly *n*-hexanoic acid (C6 fatty acid) and butyric acid (*n*-butanoic acid, C4 fatty acid). Besides, *n*-butanoic acid and *n*-hexanoic acid being much more soluble than *n*-octanoic acid, the two former substrates may be a good choice to provide evidence for an activity of CvFAP on true soluble monomers, which is not clear so

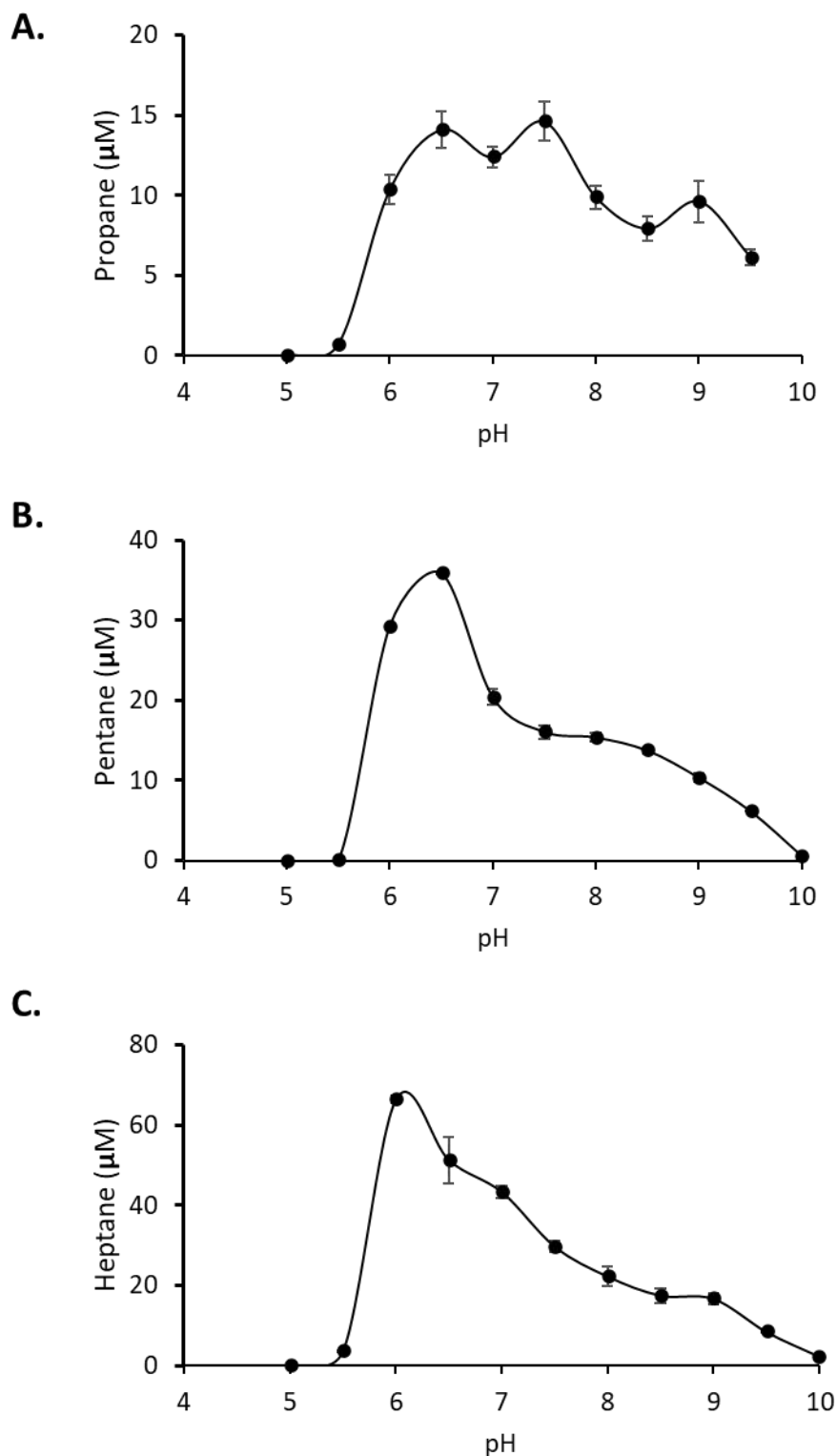
far. We therefore manage to determine the influence of substrate concentration, pH and light intensity on the conversion of these fatty acids to hydrocarbons using the purified recombinant enzyme.

## Results and Discussion

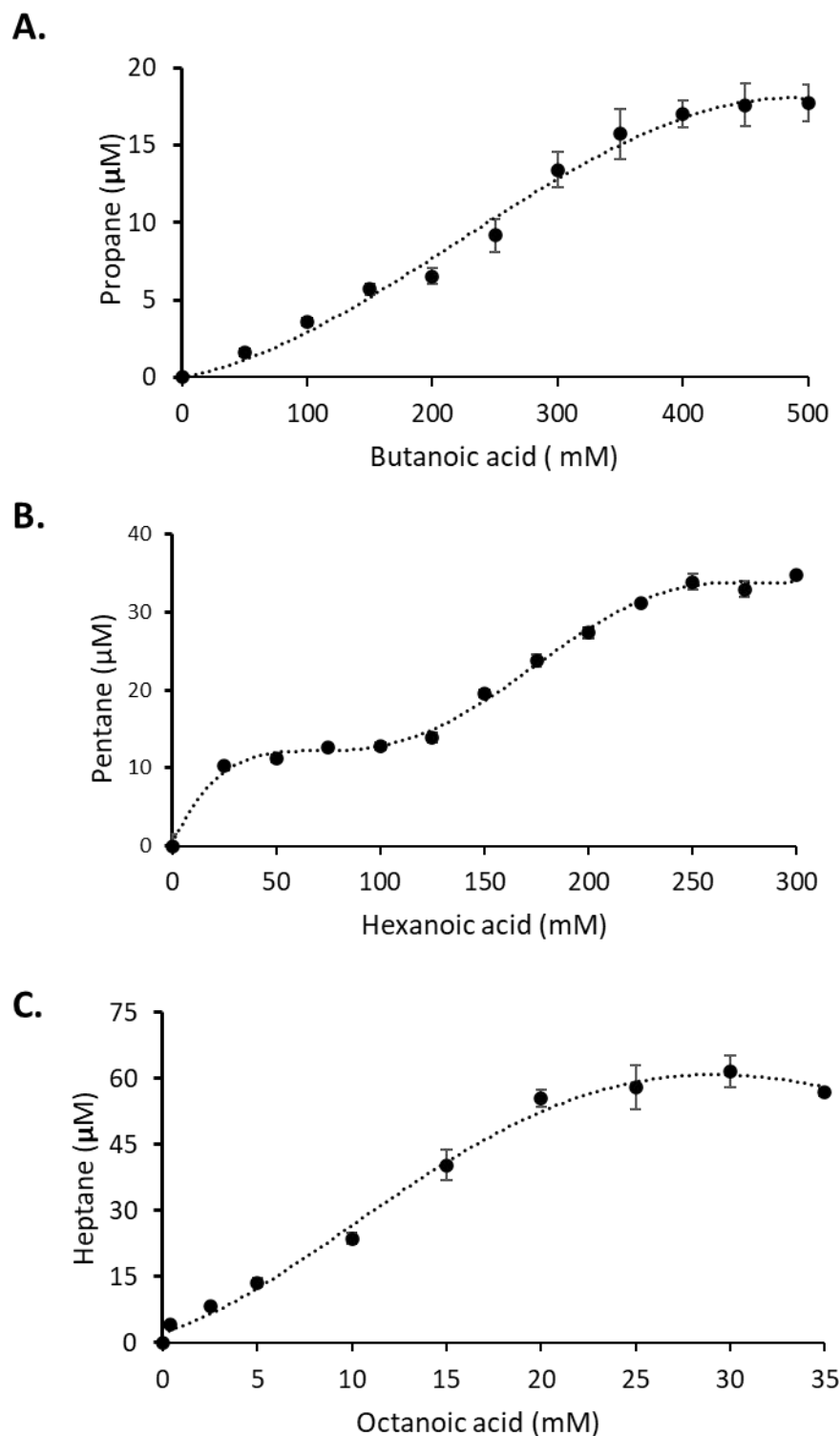
### Effect of pH and substrate concentration

We first determined the optimal pH for the C6 and C4 fatty acids substrates. We began by making a pH range from 5 to 10 in Teorell-stenhagen buffer in the presence of a fixed FAP concentration (70 nM) (**Figure 1**). The highest activity for C6 fatty acid was found at pH 6.5 and for C4 fatty acid at pH 6.5-7.5.

Having determined the pH optima, we proceeded with substrate saturation curves - for each substrate at its respective optimal pH (**Figure 2**). Between the lowest and the highest fatty acid concentration, the photoconversion of C4, C6 and C8 fatty acids was increased by a factor of 3.5, 10 and 6 respectively. The curves had a flattened sigmoid shape. It is close to the one observed for a lipase on paranitrophenylpalmitate, a substrate also partially soluble (Demera et al. [2019](#)). After a rather flat section for low concentrations, a sudden rise in activity is noted around 135 mM and 5 mM for C6 and C8 fatty acids respectively. Interestingly, these values are in the same range as their solubility limit in water (83.4 mM and 4.8 mM respectively: see table 2 in chapter 1). But it should be noted that even below the solubility limit, activity was detectable, indicating that FAP is active on fatty acid monomers (this is obvious for butyrate whose solubility limit is > 200 mM). One can thus hypothesize that the rise in activity around the solubility limit reflects the preference of CvFAP for a lipid-water interface. To verify that the rising point is intrinsic to the substrate and does not depend on enzyme/substrate ratio, we have repeated the experiment for C8 fatty acid with twice as much enzyme (140 nM). As shown in **Figure S1**, the product concentrations have doubled but the shape of the saturation curve and the rising point have remained unchanged, confirming that this feature is related to a particular substrate concentration.



**Figure 1: Effect of pH on decarboxylation of linear saturated C4, C6 and C8 fatty acids.** **A.** For *n*-butanoic acid; **B.** For *n*-hexanoic acid. **C.** For *n*-octanoic acid. Note that data in C are from Samire et al. 2023 and is reported here for comparison. The error bars show the standard deviation obtained in three independent experiments. Inset: For all panels 70 nM of purified CvFAP was mixed with the substrate. The samples were exposed to continuous light from a blue LED (emission maximum centered at 450 nm) at the intensity of 360  $\mu\text{moles photons/m}^2/\text{s}$  at 25°C during 20 min.



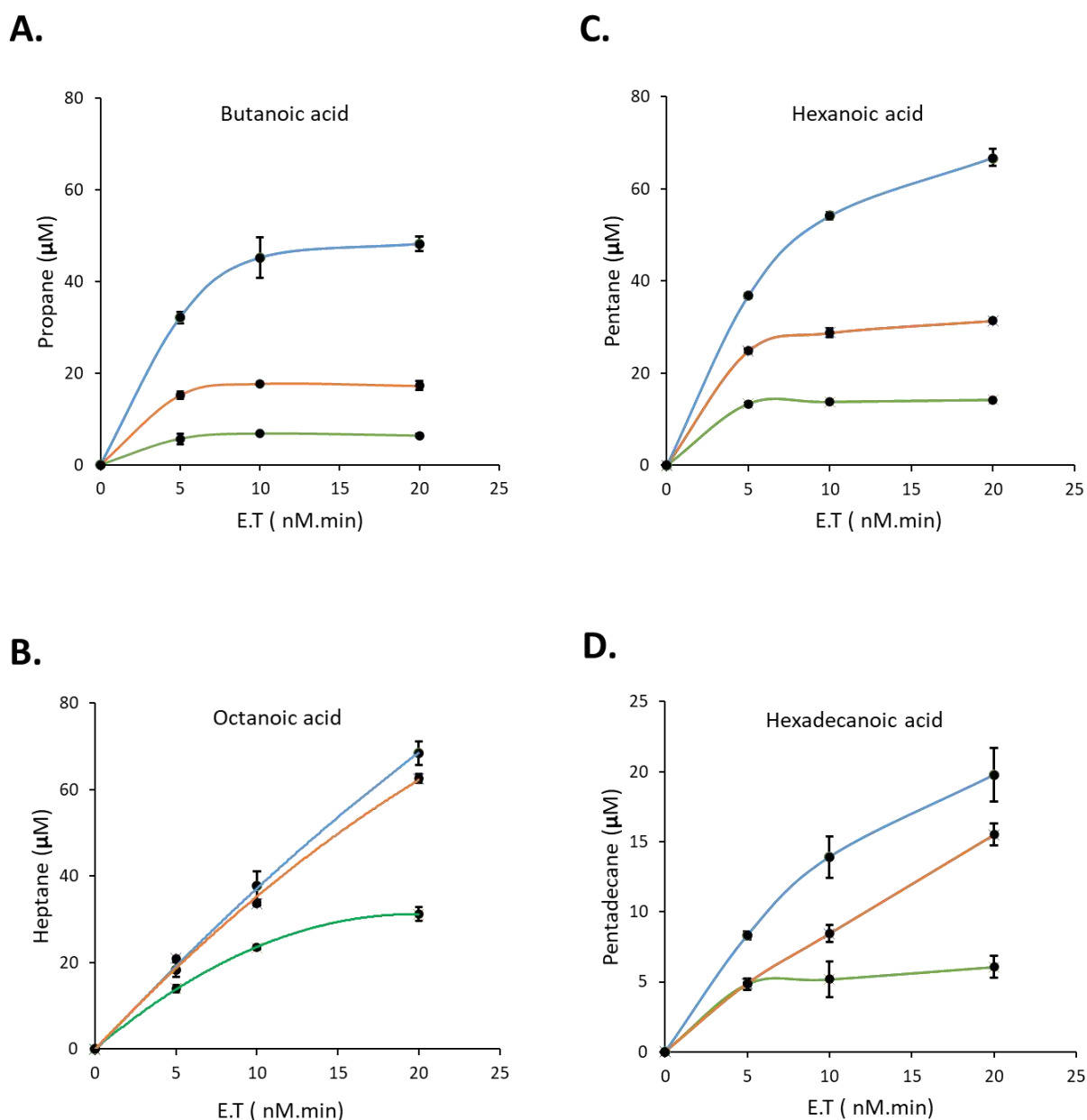
**Figure 2: Substrate saturation curves of CvFAP on C4, C6 and C8 fatty acids.** **A.** For *n*-butanoic acid; **B.** For *n*-hexanoic acid. **C.** For *n*-octanoic acid. Note that data in C are from Samire et al. 2023 and is reported here for comparison. For each fatty acid, optimum pH was used for C4 (6.5) and C6 (6.5). 70 nM of protein were used in all conditions. The samples were exposed to continuous light from a blue LED (emission maximum centered at 450 nm) at the intensity of 360  $\mu\text{moles photons/m}^2/\text{s}$  at 25°C during 20 min before the reaction were stopped and *n*-alkanes (pentane and propane) quantified. The error bars show the standard deviations obtained in three independent experiments.

## Evidence for CvFAP inactivation on some substrates

We have shown in our previous study (Samire et al [2023](#)) that *n*-octanoic acid could be converted to heptane in higher yield than *n*-hexadecanoic acid to *n*-pentadecane. This high conversion efficiency was supported by the fact that heptane, the product of the decarboxylation of *n*-octanoic acid, is able to assist in the positioning of the latter in the enzyme pocket, thus improving its photodecarboxylation yield; we then concluded that there was an autocatalytic effect, which was observed for substrates with a chain length between 8 and 10. (Typically, for an optimal autocatalytic effect, the total sum of the carbon atoms of the substrate plus its product must be  $16 \pm 1$ ). This effect has been demonstrated by kinetic spectroscopy and supported by MD simulations. However, we concluded in this previous article that the autocatalytic effect alone could not explain the better chemical yields obtained on *n*-octanoic acid in vitro (activity measurements) and in vivo (bioconversion approach). A large part may be due to an interfacial effect. We also hypothesized that the protein photostability is potentially higher with *n*-octanoic acid, increasing the total turnover numbers of the enzyme in a continuous system. Since *n*-hexanoic and butyric acids show a high activity (without being capable of autocatalytic effect), our goal was to check here the stability of the protein under continuous light condition in presence of these substrates to see what correlation there could be with the conversion yield. To test for enzyme inactivation, we have carried out a Selwyn test (Selwyn [1965](#)) on *n*-butanoic, *n*-hexanoic, *n*-octanoic and *n*-hexadecanoic acids. As stated in the original publication: “Considering the parameters: [E] Initial concentration of enzyme, [P] concentration of product formed, and [T] reaction time; the Selwyn test is based on the equation of MICHAELIS and DAVIDSHON which states that  $f([P]) = [E] \times [T]$  i.e. if all other kinetic parameters are fixed (temperature pH etc...), the concentration of product is only a function of the product of the enzyme concentration and the reaction time. Thus, if we vary [E] and [T] so that  $[E] \times [T]$  is constant, we should have points aligned on the same curve by plotting [P] against  $[E] \times [T]$  if the enzyme is not degraded over time.”

We therefore chose to test the conditions: [FAP=70 nM (1x)]. [T=10, 20, 30 min]; [FAP=35 nM (0.5x)]. [T=20, 40, 60 min] and [FAP=140 nM (2x)]. [T=5, 10, 15 min] on butyric, *n*-hexanoic, *n*-octanoic and hexadecanoic acids. **Figure 3** shows that, using 140 nM of CvFAP, the curve is linear for *n*-octanoic acid compared to butyric, *n*-hexanoic and *n*-hexadecanoic acids for which

there is a decrease of the corresponding alkanes production. This drop in activity is more pronounced when 70 and 35 nM of protein is used suggesting that an irreversible degradation (and/or inhibition) occurs when using *n*-butanoic, *n*-hexanoic and *n*-hexadecanoic acids.



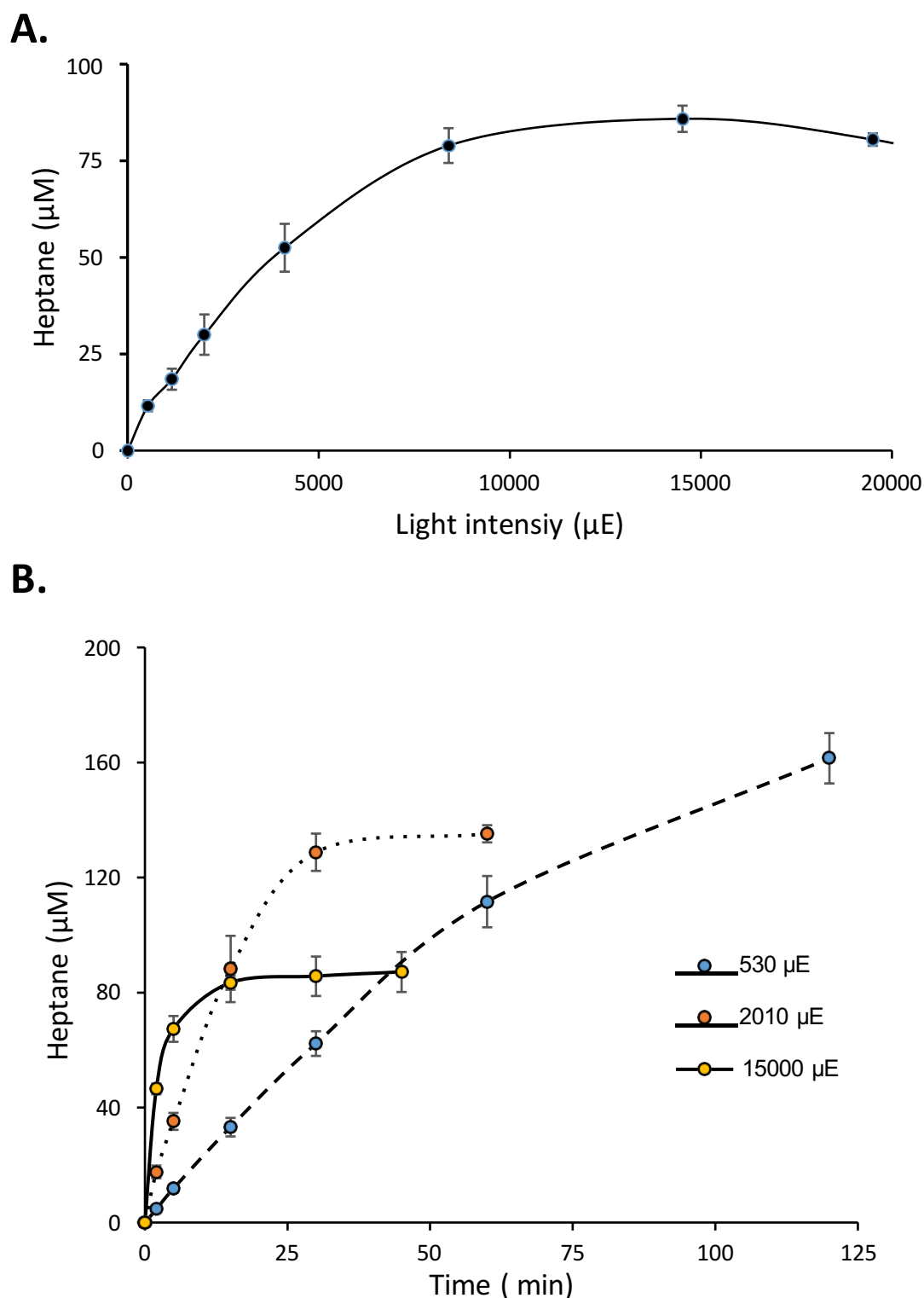
**Figure 3:** Selwyn test of CvFAP on C4, C6, C8 and C16 saturated fatty acids. Optimized conditions were used for enzymatic test i.e., pH6.5, 500mM for C4; pH 6.5, 250mM for C6; pH6 ,30mM for C8, 0.4Mm ;pH8.5 for C16.

**A.B.C.D** correspond to Selwyn test respectively on *n*-butanoic, *n*-octanoic, *n*-hexanoic and *n*-hexadecenoic acids. Blue curve corresponds to 140nM of enzyme, orange for 70nM and green for 35nM. The samples were exposed to continuous light from a blue LED (emission maximum centered at 450 nm) at the intensity of 360 μmoles photons/m<sup>2</sup>/s at 25°C before the reaction were stopped and alkanes quantified. The error bars show the standard deviation obtained in three independent experiments.

# Light-dependent inactivation of FAP

FAP is a photoenzyme, meaning that its activity depends on photons at each cycle. Indeed, knowing that to carry out a complete photocycle the enzyme needs a photon, one can imagine that enzyme activity should increase with photon fluxes until reaching a maximum. It has been shown previously that CvFAP activity increases linearly with white light intensity in the range 50-2000  $\mu\text{mol photons/m}^2/\text{s}$  under our conditions using palmitic acid as substrate (Sorigué et al. 2017). We wanted to know here if the activity of the enzyme could be saturated by light intensity. To do so, we measured over 5 min of lighting time the amount of alkane produced using *n*-octanoic acid as substrate. We tested different blue light intensities ranging from 530 to 19500  $\mu\text{mol photons/m}^2/\text{s}$ . As can be seen in **Figure 4A**, the activity of the enzyme increases with the light intensity until reaching a plateau from 8500  $\mu\text{mol photons/m}^2/\text{s}$  onwards.

Since a major goal is to be able to use the enzyme for continuous production of hydrocarbons, we wanted to evaluate how the activity of the enzyme would progress during continuous production conditions using different light intensities. To make this observation, we chose 3 light intensities: 530, 2010 and 15000  $\mu\text{mol photons/m}^2/\text{s}$  and made kinetics of heptane production using *n*-octanoic acid as substrate (**Figure 4B**). At illumination times of less than 15 min, the amount of heptane produced is greater the higher the light intensity: we produced 11, 35 and 67  $\mu\text{M}$  respectively for 530, 2010 and 15000  $\mu\text{mol photons/m}^2/\text{s}$ . Using the light intensity of 15000  $\mu\text{mol photons/m}^2/\text{s}$ , heptane production reaches 85  $\mu\text{M}$  but remains constant from 15 min onwards, suggesting that the enzyme is irreversibly degraded. Interestingly, using 30 times less light (530  $\mu\text{mol photons/m}^2/\text{s}$ ), heptane production increases to 160  $\mu\text{M}$  after 2h without saturation. At the intermediate light intensity of 2010  $\mu\text{E}$ , we observe a saturation at 135  $\mu\text{M}$  of heptane (**Figure 4B**).



**Figure 4: Light intensity effect on CvFAP activity in continuous illumination.** **A.** *n*-octanoic acid (30 mM, pH 6) were mixed with 70 nM of purified CvFAP. An increasing light intensity were tested (530-19500  $\mu\text{mol photons/m}^2/\text{s}$ ) and the corresponding heptane produced were quantified and plotted against light intensity. **B.** Samples were illuminated with three different light intensities (color code indicated on the figure) during variable duration, each dot represents illumination duration in min. Error bars show the standard deviation obtained in three independent experiments.



# Conclusions

We aimed here to determine optimal photoconversion conditions for *n*-butanoic and *n*-hexanoic acids. The pH and the concentration of the fatty acids being the key parameters that can influence the structuring of the fatty acids in solution, we have first studied the influence of these parameters on the enzyme activity on *n*-hexanoic and *n*-butanoic acids. We show how substrate concentration and pH modulation can drastically affect the activity of the purified wild type enzyme on *n*-hexanoic and *n*-butanoic acids. Since the enzyme is sensitive to light, we also evaluated in optimal photoconversion conditions its stability in the presence of substrates under different light intensities. We show here that the CvFAP stability can strongly be affected depending on the ratios light intensity / reaction time; enzyme concentration / substrate concentration. These data should guide the use of FAP for biocatalytic applications.

# Material and methods

**FAP expression** ([see](#) Samire et al. 2023, chapter 4)

**FAP purification and quantification** ([see](#) Samire et al. 2023, chapter 4 )

## **In vitro photoenzymatic production of alkanes**

The production of *n*-alkanes was performed using the *E. coli* BL21 strain containing pLIC07FAPv2 plasmid. A pre-culture was grown overnight in Luria-Bertani (LB) medium at 37°C, 180 rpm. For protein production, the strain was cultured in Terrific Broth (TB) medium supplemented with 0.5% (w/v) glycerol at 37°C, 180 rpm, to an optical density of 1 before induction with 500 µM IPTG and addition of 1-<sup>13</sup>C-labeled FA substrates at 2 mM (The concentrated stock solution of *n*-octanoic and *n*-hexadecanoic acids was previously prepared in ethanol). The temperature was then lowered to 18°C and the cultures were incubated in the dark for 24 h. 5 mL of each culture were then put in sealed 10 mL airtight vials and illuminated at 360 µmol photons/m<sup>2</sup>/s of blue LED light (450 nm) at 25°C for 1 h to induce photodecarboxylation.

### **Short chains alkanes quantification (propane, pentane, heptane)**

After C7 alkane production under light, the reaction was stopped by the enzyme denaturation at 100°C for 10 min. The samples were then cooled at room temperature for 30 min to recondense all volatile alkanes. Then, the samples were re-heated to 40°C and kept at this temperature for 5 min. 1 ml of gas phase was collected with a syringe heated at 80°C and injected into a GC-MS/FID (AGILENT Technologies; 5977B; MSD) for analysis using helium as carrier gas. The analysis parameters were as follows: oven initial temperature: 50°C for 1 min, ramp: 20°C/min to 260°C for 5 min (Column reference: 19091P-Q04PT, PH-PLOT/Q+RT, 30 m × 0.320 mm, 20 micron). The absolute amount of C7, C5 and C3 alkanes was calculated from an external standard curve ( See S2 A,B,C) obtained using pure alkane solutions. The curves were processed under the same conditions as those used for the test reactions.

### **Analysis and quantification of C15 alkane**

After C15 alkane production under light, 100 µl of NaOH (10M) were added before heating at 100°C during 10 min. 4 ml of hexane and 10 µg of *n*-hexadecane (C16 alkane, internal standard for C15 alkane quantification) were added. The samples were then vortexed and centrifuged (3000 rpm, 5 min). The organic phase containing the extracted C15 alkane was collected and 1 µl was injected in GC-MS/FID for analysis according to the following program: Oven initial temperature: 60°C for 1 min, ramp: 20°C/min to 150°C; 10°C/min to 260°C for 2 min. Column reference : OPTIMA WAXplus, 30 m × 0.25 mm. The absolute amount of C15 alkane was calculated based on the internal standard added before the sample treatment.

## **References**

- Amer, M., Wojcik, E. Z., Sun, C., Hoeven, R., Hughes, J. M., Faulkner, M., ... & Scrutton, N. S. (2020). Low carbon strategies for sustainable bio-alkane gas production and renewable energy. *Energy & Environmental Science*, 13(6), 1818-1831.
- Aselmeyer, C., Légeret, B., Bénarouche, A., Sorigué, D., Parsiegla, G., Beisson, F., & Carrière, F. (2021). Fatty Acid Photodecarboxylase Is an Interfacial Enzyme That Binds to Lipid–Water Interfaces to Access Its Insoluble Substrate. *Biochemistry*, 60(42), 3200-3212.

Demera, L. L., Barahona, P. P., & Barriga, E. J. C. (2019). Production, extraction and characterization of lipases from the antarctic yeast *Guehomyces pullulans*. *Am J Biochem Biotechnol*, 15(2), 75-82.

Lakavath, B., Hedison, T. M., Heyes, D. J., Shanmugam, M., Sakuma, M., Hoeven, R., ... & Scrutton, N. S. (2020). Radical-based photoinactivation of fatty acid photodecarboxylases. *Analytical biochemistry*, 600, 113749.

Poutoum P. Samire, Bo Zhuang, Bertrand Légeret, Ángel Baca-Porcel, Gilles Peltier, Damien Sorigué, Alexey Aleksandrov, Frédéric Beisson, Pavel Müller (2023), Autocatalytic Effect Boosts the Production of Medium-Chain Hydrocarbons by Fatty Acid Photodecarboxylase. *Science Advances*, *in press*

Santner, P., Szabó, L. K., Chanquia, S. N., Merrild, A. H., Hollmann, F., Kara, S., & Eser, B. E. (2021). Optimization and engineering of fatty acid photodecarboxylase for substrate specificity. *ChemCatChem*, 13(18), 4038-4046.

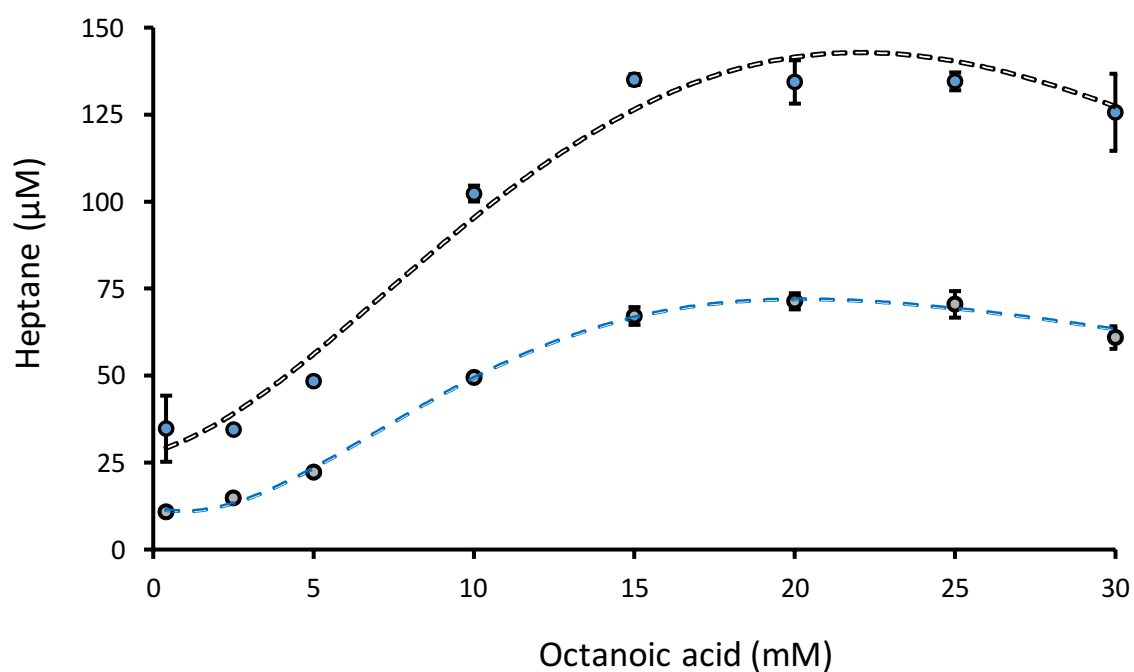
Selwyn, M. MJ. (1965) A simple test for inactivation of an enzyme during assay. *Biochim. Biophys. Acta* 105 193-195.

Sorigué, D., Légeret, B., Cuiné, S., Blangy, S., Moulin, S., Billon, E., ... & Beisson, F. (2017). An algal photoenzyme converts fatty acids to hydrocarbons. *Science*, 357(6354), 903-907.

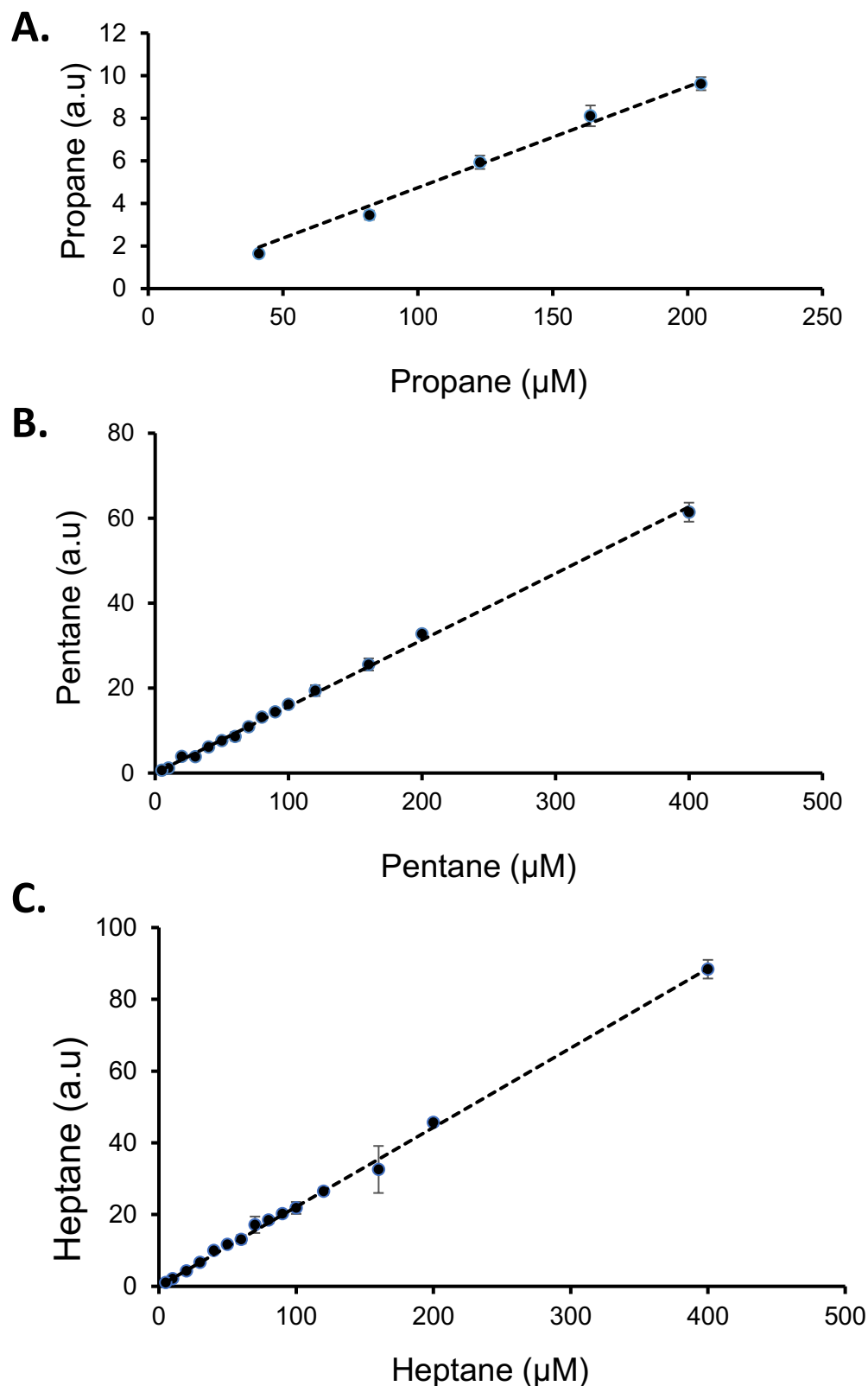
Selwyn, M. J. (1965). A simple test for inactivation of an enzyme during assay. *Biochimica et Biophysica Acta (BBA)-enzymology and biological oxidation*, 105(1), 193-195.



# Additional figures



**Figure S1: Effect of enzyme concentration on the shape of C8 fatty acid saturation curve.** Concentration dependence of CvFAP activity on C8 fatty acid. (Black curve) represents the curve for 70 nM of protein and blue curve 140 nM. The samples were exposed to continuous light from a blue LED (emission maximum centered at 450 nm) at the intensity of 360  $\mu\text{moles photons/m}^2/\text{s}$  at 25°C during 20 min before the reaction were stopped and heptane quantified. The error bars show the standard deviations obtained in three independent experiments.



**Figure S2: Calibration curves for alkane quantification.** Pure alkanes were putted into 10 ml vial containing 5 ml of water. Samples were heated for 10 min before cooled at room temperature. Alkanes were then analyzed by headspace GC-MS like described in MM (Short chains alkanes quantification (propane, pentane, heptane). Corresponding signal area obtained (designated by a.u) were plotted against the know initial concentration putted in the vial. **A; B; C** corresponds respectively to calibration curves of propane, pentane, and heptane.





# General conclusion and Perspectives

During my thesis I explored the conservation of FAP photochemical activity in the biodiversity of FAPs and performed structural and spectroscopic studies on CvFAP that contributed to a more detailed model of its mechanism. I also investigated CvFAP substrate specificity, which led to the discovery of an autocatalytic mechanism on medium-chain fatty acids and to the realization that CvFAP can in fact be fairly active on C4-C10 fatty acids under certain conditions.

By culturing various algae, we have demonstrated that they are able to produce hydrocarbons with a varied composition. The heterologous expression of some homologous FAP sequences in *E. coli* revealed that these genes are sufficient to perform light-induced hydrocarbon production. By observing the production profiles, we noted variable proportions of hydrocarbons. Indeed, some of these homologous proteins, in this case *Ectocarpus silicosus*, showed a higher production of shorter chain hydrocarbons. This shows that it is worth screening more FAPs from the biodiversity (shorter, and thus more volatile hydrocarbons are very interesting for the development of continuous production systems since they can be directly recovered in the gas phase of cells without the need to lyse the cells to recover them, or to perform purification steps which are very costly). It would also be interesting to purify the FAPs with the most promising features in order to study more precisely their substrate specificity in vitro. During this thesis, I also had the chance to use recent advanced biophysical techniques such as XFEL (European X-ray Free Electron Laser) which allowed us to obtain a structure of better resolution (1.8-Å-resolution). Moreover, all the steps of the photocycle have been elucidated thanks to the study of variants from important residues involved in the mechanism of the enzyme (R451, C432). The main information that we have been able to bring to light with respect to the first proposed mechanism are the following: (i) residues R451 and C432 are involved in the positioning of the deprotonated fatty acid in the enzyme pocket thus favoring its oxidation by the excited FAD molecule i.e.  $^1\text{FAD}^*$ . (step 2) (ii) the back ET of the  $\text{FAD}^{\bullet-}$  group towards the alkyl is concomitant to a proton transfer. A good part of the  $\text{CO}_2$  formed during the decarboxylation is transformed into bicarbonate (step 4). We were able to elucidate this very complex mechanism.

In this thesis, we also performed the first in vitro characterization of the enzymatic activity of FAP. Despite a large number of studies using FAP as a model and/or tool, few have

worked on the characterization using pure enzyme. We were able to demonstrate here that the optimization of the catalysis conditions, in particular the pH and the substrate concentration, are as well as the various mutagenesis techniques, a reliable alternative to improve the activity of the enzyme. Indeed, *n*-octanoic acid has proven to be a much better substrate than palmitic acid which has always been described as the preferred substrate for FAP. The efficient conversion of short chains not only allows to produce more hydrocarbon material but also to have a technical advantage in the recovery of hydrocarbons as described in the previous paragraph. This work should be expanded to other substrates but also on mutant with high efficiency; hopefully my work will be useful to guide the engineering and/or the screening of the FAPs found in the biodiversity, by taking into account both extrinsic and intrinsic parameters governing the efficiency of the enzyme.

# Abstract

Hydrocarbons (alkanes, alkenes) are molecules composed only of hydrogen and carbon atoms that are an integral part of daily human needs. Indeed, they are the basic compounds of fuels and are used in chemistry, as solvents, lubricants and in cosmetics. Almost all the hydrocarbons we use today are of fossil origin. However, pathways of hydrocarbon biosynthesis from fatty acids exist in many organisms and involve various enzymes. Unfortunately, these enzymes often have low turnovers and require catalytic conditions that are difficult to implement in industry. Thus, the understanding of the mechanisms of hydrocarbon-forming enzymes and their engineering has gained significant interest over the last 10 years. The last hydrocarbon-forming enzyme discovered (in 2017) is fatty acid photodecarboxylase (FAP), a protein found only in algae. This enzyme is particularly interesting because its reaction does not require electron donors but only a photon at each catalytic cycle. It is thus a photoenzyme, a rare type of enzyme. FAP has then attracted a lot of interest, especially for the photoconversion of fatty acids to hydrocarbons. My PhD work had three objectives: (i) Determine whether FAP activity was conserved in algae beyond the model green alga *Chlorella variabilis* NC64A where it was discovered, and see if some other putative FAP were likely to have a fatty acid specificity different than the FAP from *Chlorella variabilis* (CvFAP); (ii) Participate in a large multidisciplinary study involving an international consortium of laboratories which aimed at gaining insights into the structure and mechanism of CvFAP; (iii) Study the fatty acid specificity of CvFAP with a focus on short- and medium-chain fatty acids, which were thought to be poor substrates for FAP. In the first part, I have expressed in *E. coli* seven homologs of CvFAP from various algal groups and been able to obtain soluble active FAP for four of them: FAPs from *Ectocarpus siliculosus* (a brown macroalga), *Chondrus crispus* (a red macroalga), *Nannochloropsis gaditana* (a microalga close to brown algae) and *Galdieria sulphuraria* (a red microalga) were found to have conserved FAP activity but with distinct fatty acid specificities. In the second part, I have participated in the production and purification of CvFAP on a large scale for different biophysical approaches, performed biochemical, static crystallography and transient absorption spectroscopy experiments on CvFAP and on a mutant (R451K) important for the substrate stabilization. The whole study allowed us to elucidate the complete CvFAP mechanism. In the third part, I have studied the activity of CvFAP in vitro and shown that it can convert *n*-octanoic acid four times faster than *n*-hexadecanoic acid, its best substrate reported to date. I have also shown that in vivo this translates into a CvFAP-based production rate over ten-fold higher for *n*-heptane than for *n*-pentadecane. Experiments of time-resolved spectroscopy I have performed in collaboration have provided evidence that the high catalytic activity of FAP on *n*-octanoic acid is in part due to an autocatalytic effect of its *n*-heptane product, which fills the rest of the binding pocket of CvFAP. Finally, I have determined the effect of substrate concentration, pH (and light intensity) on the activity of CvFAP on *n*-hexanoic and *n*-butanoic acids. These results should guide future strategies of FAP selection, improvement and use for a light-driven production of medium- and short-chain hydrocarbons.

**Key words:** hydrocarbons, fatty acids, photoenzymes, biocatalysis, biofuels.